CHAPTER 2

Early Digital Literacy: Learning to Watch, Watching to Learn

Georgene L. Troseth*, Gabrielle A. Strouse†, Colleen E. Russo Johnson*
*Vanderbilt University, Nashville, TN, United States
†University of South Dakota, Vermillion, SD, United States

Since the 1960s, educational researchers have argued that in the early stages of literacy, teachers should focus on teaching children *how* to read before expecting them to read for meaning (Snow, 2015). For example, according to Chall (1983), when children first begin to master aspects of this complex skill, "reading is not for learning;" children must develop awareness of the sounds that make up words and learn to decode written letters before they can read to learn. In current thinking, this is a gradual progression: as the component skills develop, children become better able to learn independently from text (Soden et al., 2015). Individual differences in reading ability, largely determined by stable genetic differences, are found in children's earliest mastery of components and in reading comprehension throughout formal schooling; differences in home literacy environment contribute to reading comprehension most strongly in the early stages (Soden et al., 2015). In this chapter, we show that coming to understand and learn from video—a form of *digital literacy*—is similar in several respects.

The idea that children need to *learn how to learn* from video may be surprising, because watching video seems so much like watching and learning from real events. Although digital literacy develops earlier than mastering written language, research indicates that very young children do not learn efficiently from video. We will examine research suggesting why this might be and the experiences that help children to learn from video. These studies reveal information about what has been termed "early symbolic development," the beginning of representational skills that both help toddlers use video and help children of school age master a wide range of symbolic media, including writing, numerals, diagrams, and maps (Troseth & DeLoache, 1998).

Cognitive Development in Digital Contexts
http://dx.doi.org/10.1016/B978-0-12-809481-5.00002-X

PERCEPTION OF VIDEO IMAGES

One component of digital literacy develops early in infancy: the ability to make sense of two-dimensional (2D) still and moving images and see their similarity to what they depict. For instance, after 5-month-olds habituate to the sight of a doll, they transfer their habituation to a picture of the doll (DeLoache, Strauss, & Maynard, 1979), showing that they recognize it as the same doll. Between 2 and 5 months, infants use a video image of their otherwise out-of-sight legs to direct their kicks at a noisy toy (Rochat & Morgan, 1995) and respond to video of an adult as they would to the actual adult, with smiles and increased movement (Bigelow, 1996; Hayes & Watson, 1981; Muir, Hains, Cao, & D'Entremont, 1996; Murray & Trevarthen, 1985). At 9 months, infants express similar emotions to videos of various entities (a person, an interesting toy, a spooky mask) as they do to the real things (Diener, Pierroutsakos, Troseth, & Roberts, 2008). By 1 year, the emotional responses that adults on video make to toys influence infants' willingness to engage with the real toys (Mumme & Fernald, 2003).

Besides making sense of pictorial images, infants also discriminate images from reality. Newborns both see the similarities between and differentiate pictures of objects from real objects (Slater, Rose, & Morison, 1984). By 4–6 months, infants smile more at a real person than at live video of the same (equally responsive) person (Hains & Muir, 1996), and 9-month-olds look longer at actual people, objects, and events than at videos of these entities (Diener et al., 2008). Thus, perception of the contents of video appears to be relatively automatic, but infants *do* attend to perceptual cues that differentiate images from reality.

An initial challenge for children's digital literacy is posed by the conflicting cues that make pictorial images both similar to and different from the three-dimensional (3D) world. Realistic video images and pictures retain much of the information in their real-world referents, including color, shadows, relative size, and (in the case of video) movement. At the same time, infants can perceive the flatness of images. Sensitivity to depth cues based on binocular disparity and motion parallax develops by 4 months (Held, Birch, & Gwiazda, 1980; Nawrot, Mayo, & Nawrot, 2009): babies detect that the image of a 3D object to their two eyes is slightly different, and changes as they move their head.

Children's natural exploratory behavior gradually helps them determine the affordances of 2D images. For instance, young infants often attempt to grasp at objects depicted in realistic pictures (DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998) and video. When 9-month-olds were seated within reach of a video screen on which a series of toys appeared, each child manually investigated by rubbing and patting the pictured objects and attempting to pluck them off the screen (Pierroutsakos & Troseth, 2003). As shown in Fig. 1, infants were especially persistent with moving

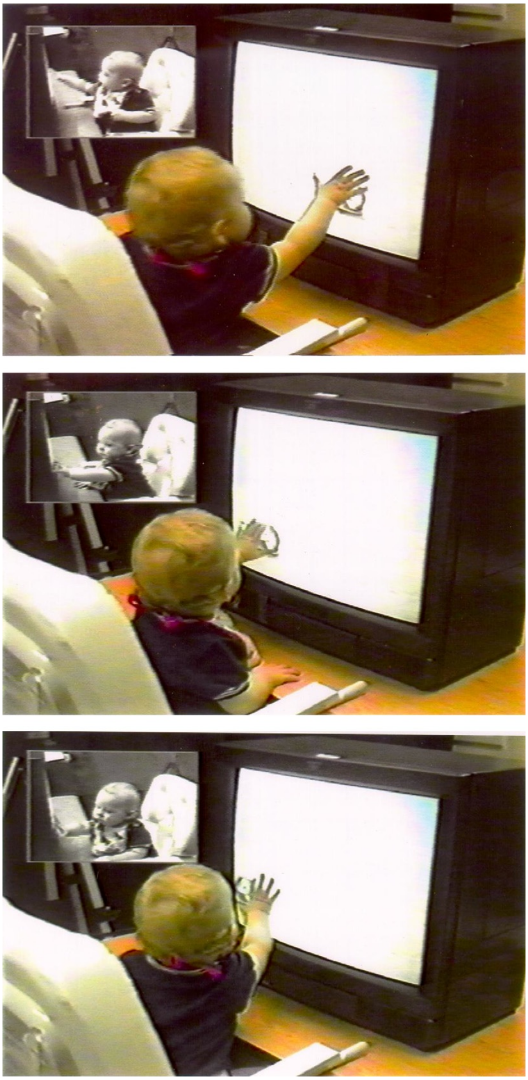

Fig. 1 A 9-month-old trying to grasp a moving object pictured on video.

toys, following them across the screen, and repeatedly trying to access them. Although 9-month-olds perceive information about the lack of depth in the images, they may not know the implications of these cues for their behavior, failing to realize that depictions do not afford the same set of responses as do the actual depicted objects.

LEARNING ABOUT PICTURES

Over the course of their second year, children "rapidly progress to treating pictures referentially" (Walker, Walker, & Ganea, 2013, p. 1321). Part of digital literacy is regarding images as related to, but distinct from, what they represent. Video images and pictures bring to mind stored knowledge about whatever is depicted. For instance, a picture of an apple will bring to mind conceptual information related to fruit, food, and possibly apple trees, but the affordances of a real apple and a picture are different. For those with mature "pictorial competence" (i.e., understanding of the picture-referent relation; Troseth, Pierroutsakos, & DeLoache, 2004), seeing a picture or a video automatically elicits conceptual knowledge about the contents and about the representational medium itself that blocks responding to the contents directly. For example, seeing a picture of an apple evokes knowledge of apples as food that can be held and eaten. But for those with mature pictorial competence, flatness cues that identify the apple as "*a picture of…*" elicit representational knowledge that blocks inappropriate responses such as trying to eat the picture.

In studies with both pictures (DeLoache et al., 1998) and video (Pierroutsakos & Troseth, 2003), 15-month-olds rarely tried to grasp at depicted objects, and 19-month-olds never did. Rather, the older infants pointed at the depictions, looking to adults in the room and attempting to talk about the images. By 15 months, infants realize that a verbal label used to describe a picture (e.g., "This is a whisk") refers to the depicted object, not just to the picture of the object (Allen Preissler & Carey, 2004; Ganea, Allen, Butler, Carey, & DeLoache, 2009). The realization that an image stands for and directs attention to something other than itself (DeLoache & Burns, 1994; Troseth et al., 2004) is a major development in digital literacy.

Young children's more mature stance toward pictures appears to stem from experience with adults' responses to them. Middle-class parents in western cultures spend substantial time reading and discussing picture books with their young children (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Ninio & Bruner, 1978). This experience appears to promote children's representational understanding of pictures: US and Canadian children learn the novel names of pictured objects and generalize those labels to the real, depicted object at a younger age than children do in cultures without easy access to pictures (Callaghan et al., 2011; Walker et al., 2013). The ways that adults differentiate pictured objects during conversation support children's learning to treat them as representations. In one study, when parents and children in the US were given replica toys (animals, artifacts) and pictures of those toys one at a time, parents referred to pictured objects generically, as kinds of things ("Look, a dog"), whereas they referred to the replicas by name ("Hello, Spot"), treating them as imaginary conversational partners (Gelman, Chesnick, & Waxman, 2005). Of interest, the children (2.5- to 3-year-olds) also marked this distinction by how they referred to the toys and pictures. In this early aspect of digital literacy, social input plays an important role in helping children take a referential stance toward 2D images. Additionally, children's initial concepts may reflect the activities and parental behavior they have experienced with particular symbolic media. For instance, at an age when children recognize that a picture and its label (e.g., "This is a whisk") refer to the real, depicted entity (an experimental task very similar to the naming game that Western, middle-class parents and children play with picture books), children fail to use a picture in a novel way: as information about the location of a toy hidden during a "hide and seek" game (DeLoache & Burns, 1994).

Because children's inclination to treat a symbolic object referentially sometimes appears to be context bound, it is important to consider that children's concepts about pictures and about video may develop separately, based on different experiences with these media. For instance, most parents rarely if ever co-view video with their children in the same supportive way that they co-read picture books (Strouse, O'Doherty, & Troseth, 2013).

LEARNING TO USE INFORMATION FROM PEOPLE ON VIDEO

Starting in infancy, children perceive various matches and mismatches between video and reality. The very fact that video can (but does not always) depict ongoing events may challenge young children to work out the correct relation between a video and real life and affect their learning from the medium. At 2 months of age, babies respond the same to live and pretaped (delayed) video of their talking and smiling mothers (Marian, Neisser, & Rochat, 1996). Between 4 and 8 months, they become less responsive to noncontingent video (Bigelow, MacLean, & MacDonald, 1996; Hains & Muir, 1996). After watching people on pretaped video, infants of this age were less attentive and responsive to the same people on live video and in face-to-face interactions, compared to infants who saw the responsive partners first (Hains & Muir, 1996). Expectations of noncontingency apparently endured, affecting infants' responses to people on video as long as a week later (Bigelow & Birch, 2000).

In another study, missing social cues affected infants' learning of speech information from people on video. After interacting with a Mandarin speaker face-to-face during 12 sessions, infants (whose parents spoke English) could distinguish Mandarin sounds; however, infants who watched the same speaker on pretaped video for equivalent time showed no evidence of exposure to Mandarin (Kuhl, Tsao, & Liu, 2003). Although the speaker on video made apparent eye contact through the camera while she talked in Mandarin about picture books and toys, the 9-month-olds who watched were much less attentive than those who observed the real person facing them. The authors concluded that social cues in the direct interaction kept infants' attention and indicated what the speaker was talking about, which promoted infants' learning of speech information. The repeated sessions also gave infants ample time to form expectations about the nonresponsiveness of the on-screen person.

A large body of research examines young children's imitation of behavior that has been modeled on video. Typically, the target behavior is a simple series of actions to assemble a toy. To imitate, children must store a memory of the objects and event and then retrieve that memory when they later are given access to the real objects. In Barr and colleagues' research, the person demonstrating the series of three actions either was present in the children's living room, or appeared on the family television set. The youngest children that Barr, Muentener, and Garcia (2007) were able to test (6-month-olds) imitated just as much after watching the person on video or "face to face;" they required six repetitions of the demonstration to show significant imitation over a baseline control group, regardless of presentation mode. Infants from 12 to 30 months learned after just three repetitions from a modeler who was present (Barr, Dowden, & Hayne, 1996; Barr & Hayne, 1999; Hayne, Herbert, & Simcock, 2003). However, the same three repetitions of the person's actions on video elicited significantly less imitation (Barr, 2010; Barr & Hayne, 1999; Hayne et al., 2003; McCall et al., 1977). Across studies using this paradigm, 18-month-olds produced more target actions after watching them demonstrated on TV than a no-demonstration control group, but they imitated only half as many actions as infants who saw an in-person demonstration, and remembered these actions only half as long (Barr, 2013). Letting 24-month-old children see more repetitions of the target behaviors on video increased their imitation (Strouse & Troseth, 2008). Barr (2013) reasoned that repeated demonstrations allowed children to notice different aspects of the depicted event across repetitions, which strengthens the memory and enhances retrieval cues.

Several ideas have been proposed to explain why children imitate less from video. According to the *perceptual impoverishment* hypothesis (Barr, 2010), 2D images lack some perceptual features present in a 3D event (e.g., depth cues, matching size), whereby the child encodes a memory that is less rich and detailed and provides fewer retrieval cues—thus, a memory that is more difficult to access during the test than a memory of the actual event (Barr, 2013). However, it is difficult to use perceptual impoverishment to explain why the youngest viewers, but not the older infants, imitated equivalently from video and reality. Another explanation involves *context differences* between an event seen on a TV screen and reality. The idea is that young children encode details of an event's context (such as the frame around the TV screen) *with* the content, leading to a retrieval mismatch that hinders access to the memory in the testing situation (Barr, 2013). (A similar case was reflected by Butler and Rovee-Collier's (1989) infant participants failing to kick to get a mobile to move because the crib bumper present when they formed the memory had changed.) Barr's recent research with touchscreens reveals the difficulty of transferring a memory across contexts (Moser et al., 2015; Zack, Barr, Gerhardstein, Dickerson, & Meltzoff, 2009). Children saw an adult demonstrate pushing a button on a toy fire truck, or a virtual button on a 2D fire truck on a touchscreen, to make a siren sound. Children imitated the button push twice as often if they were given the object on which they saw the demonstration than they did when they had to transfer to the new item. Of interest, they did as poorly transferring from the real, 3D ("perceptually richer") toy to the touchscreen as they did from the "perceptually impoverished" touchscreen to the toy, suggesting that something other than diminished perceptual richness in the encoded memory blocked children's access to it. A mismatch between memories created in the different contexts may have compromised transfer in either direction.

More than *perceptual* details of the context (e.g., the frame of the TV) may contribute to a context mismatch between demonstration memory and testing situation—specifically, children's *concepts* about images on TV may impact whether they see the situations as different. Using Barr and Hayne's (1999) imitation task and replicas of their stimuli with 24-month-olds, we found an interesting contextual effect of children seeing the demonstration on their home TV (Strouse & Troseth, 2008). Children saw the demonstration at home or in the lab and were tested in the same context a day later. Children who saw the demonstration on their home TV produced only half as many target behaviors as children who watched the same video on an unfamiliar monitor in the lab; children who saw an in-person demonstration imitated at a high rate regardless of context. Seeing the video demonstration in their typical TV-viewing context seemed to impede children's access to the memory when they were presented with the real toy at home a day later. Presumably the video screen in both settings had a frame around it, which should have made it equally likely that children encoded context features. One possibility (discussed below) is that children's *concept* that events "on TV" are distinct from reality was heightened in the setting where they typically watched unrealistic video such as cartoons.

SOLVING A PROBLEM USING VIDEO

To imitate an adult's demonstration of toy assembly seen on video, children's memory of the event needs to be triggered when they are presented with the real, 3D components of the toy. However, to succeed, they do not need to know *where* they learned the information; in fact, contextual features about the source of the information may make the memory harder to access (Barr, 2013). In another frequently used task (Troseth & DeLoache, 1998), success may *require* a degree of awareness that video can provide information about current reality. Children are introduced to a toy (Snoopy) who, they are informed, likes to hide in a room. A researcher points out and labels each piece of furniture in the room. Next, she turns on a TV set, identifies a video camera connected to the TV, and demonstrates that children can see themselves, the researcher, their parent, Snoopy, and the furniture on the TV. She points out the correspondence between each real item of furniture and its image on the screen. Then she wheels the TV into an adjoining control room and spatially aligns it with the main room. The researcher tells the child that she will go hide Snoopy in his room, and "you can watch me on the TV." She leaves the control room, closes the door, appears on the screen, and deliberately hides Snoopy. An assistant with the child points out what is happening (making sure the child attends) without labeling the hiding place. The researcher returns to the control room, reminds the child that Snoopy is hiding "where you saw him on TV," and takes the child to the room to search.

Children who are 30 months of age usually succeed at this task; most find Snoopy using the information from video as he is hidden in four different places across trials (Troseth & DeLoache, 1998). To do so, they must update their mental representation of his current location on each trial using information from the video, rather than using outdated information from their direct experience (where they found Snoopy last time). What seems crucial for children of this age is getting the brief orientation to live video—seeing themselves, their parent, the toy, and the furniture on the TV "in real time." In a preliminary study in which children saw a stranger hide Snoopy in a pretaped video, they were unsuccessful (Troseth & DeLoache, 1998). Although the video of the hiding event showed them Snoopy's location in the familiar room, the children did not seem to realize that they had been given information relevant to the finding game.

For 24-month-olds, however, even the live video orientation was not sufficient. During the orientation phase, children appeared to see the similarity between the video image and what was depicted, identifying themselves and their parent, holding up the toy and commenting on its appearance on the screen, and identifying furniture items in the room when the researcher pointed to them on the TV. Nevertheless, once the TV was moved to the control room and the only way to know where Snoopy was hiding was to use information presented on screen, the children searched inconsistently (see Fig. 2). Quite frequently, children would find the toy on the first trial, showing some memory for an event seen on video. However, on subsequent trials, they would not update their knowledge of the toy's current location using information from the video. Rather, they would return to a location where they had previously found the toy and search there (Troseth & DeLoache, 1998).

One proposed explanation for 24-month-olds' poor performance was that memory limitations or difficulty inhibiting a previously successful response made the task challenging for this age group (Schmitt & Anderson, 2002; Suddendorf, 2003). Alternatively, the need to apply information from a symbolic medium (video) to reality was the challenge. To test this, we had 24-month-olds participate in the same search task, except they saw the hiding events directly through a window (the same size as the TV screen) between the rooms. All of the children found Snoopy on every trial. Rather than perseverating, they used what they saw through the window to update their mental representation of the toy's current location, and then used that

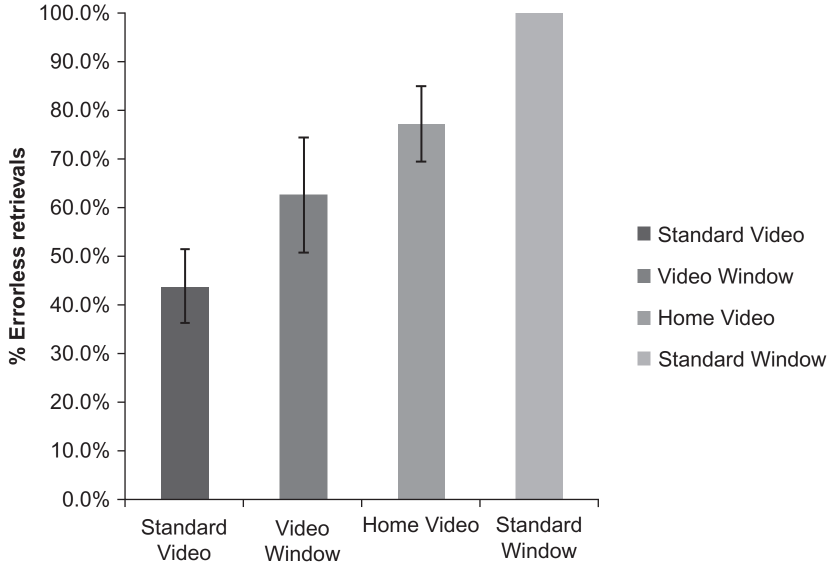

Fig. 2 Percentage of errorless retrievals 24-month-olds achieved when searching for a toy after seeing live video of hiding events (Standard Video); live video visible through a window, with the monitor hidden (Video Window); live video experience at home before the search task (Home Video); and a direct view through a window (Standard Window). Vertical lines depict standard errors of the means.

memory to search correctly (Troseth & DeLoache, 1998). Thus, searching across multiple trials was an age-appropriate task that did not strain children's memories or inhibitory capacity.

In follow-up studies, we gave 24-month-olds more support to see that the video represented the real event. In one study, the door between the control room and hiding room was left open on four trials, so children saw hiding events directly (through the doorway) and on the TV simultaneously (Troseth, 2003a). They were highly successful on these trials. Next there were four normal trials; the door was closed and children watched on video as the researcher hid the toy and were no more successful than the 24-month-olds were in the original study.

A "perceptual difference" explanation was proposed—searching in a 3D room based on a 2D video image of the hiding event was perceptually challenging (Schmidt, Crawley-Davis, & Anderson, 2007). To explore this hypothesis, Schmidt and colleagues invented two clever tasks. In one, rather than children *seeing* a researcher hiding a toy in a room, the researcher appeared on video and *verbally told* children where she had hidden the toy (also see Troseth, Saylor, & Archer, 2006). In the other task variation, the search space was a flat feltboard containing four distinctive flat shapes made of felt (e.g., a present, a cake) behind which a sticker could hide. Here, the 2D video image of a researcher's hand hiding a sticker behind a flat object on the feltboard was perceptually similar to the actual feltboard on which they searched. Despite the decreased need to map visually the toy's location across 2D-3D perceptual differences, children in both studies experienced great difficulty using information from video to find a hidden object.

CHILDREN'S CONCEPT OF VIDEO

Earlier we mentioned that babies developed enduring expectations about nonresponsive people on video (Bigelow & Birch, 2000; Hains & Muir, 1996) and that children had particular trouble imitating when they saw a demonstration on their home TV screen (Strouse & Troseth, 2008). Could toddlers' expectations about video—their concept about events appearing "on TV"—make using information from video a challenge?

One way to test this hypothesis was to convince children that they were watching directly through a window, when they were actually watching video—that is, interfere with children conceiving of the event as being "on TV." To do this, we ensured that the child never saw the television or video camera (Troseth & DeLoache, 1998). After they were introduced to Snoopy and the furniture, the researcher pointed out the window between the rooms and suggested that the child and parent go with her to the other room, and then "watch through the window" as she hid the toy. While the child was climbing up on a chair in the control room to "look through the window" (which was covered by a curtain), the assistant in the hiding room moved the video monitor into a place behind the window, so that only the screen appeared in the window once the curtain was opened. The illusion was quite compelling (see Fig. 3). After watching the hiding event, the child climbed down to go search; meanwhile, the sneaky assistant rolled the video equipment out of sight. Thus, when the child entered the room, no monitor was visible. A control group participated in the video condition where they saw the monitor (placed in front of the window curtain inside the control room, as usual) and knew they were watching on video.

Children in the two conditions were watching the identical 2D video image, yet when we told children in one condition that they were directly

Fig. 3 A 24-month-old watching through the Video Window.

watching the hiding event, and obscured the fact that they were watching on TV, significantly more children (9 of 16) succeeded on every trial (vs. 3 of 16 who saw the monitor). This result suggests that children's concept or expectations about video was interfering with their using the information.

Features that make video a versatile representational medium may actually contribute to young children's difficulty in using it for information. While looking so much like real events, video images relate to reality in a number of different ways. For instance, a realistic video can represent current reality (e.g., a video chat session with Grandma) or the past (instant replay is notoriously confusing for young children—Rice, Huston, & Wright, 1986). The sophistication of current CGI animation makes possible real-looking video of completely fictional entities (e.g., aliens or dragons). Cartoons and movies often involve events that contradict children's growing world knowledge: animals talk and wear clothes, and physical objects defy laws that babies recognize (e.g., that unsupported objects fall—Baillargeon, Needham, & Devos, 1992). Importantly, on television, people appear to be talking to the viewer, but they cannot respond if the viewer talks back (Troseth, 2010). It would not be surprising if very young children segregated events on video as *not-reality*, as events *"on TV,"* much as children are thought to mark pretend identities (e.g., banana = phone) as different from reality (banana = fruit) so as not to infect their real-world concepts (Harris & Kavanaugh, 1993; Leslie, 1987).

CONCEPTUAL DEVELOPMENT THROUGH NEW EXPERIENCE WITH VIDEO

If children's concepts about video develop from experience, we should be able to change their concepts by giving them different experience. In a study in the early 2000s, we gave 24-month-olds the most obvious evidence we could devise to show that video could represent current reality. We randomly assigned some parents to connect their video cameras to their televisions, and their children saw themselves "live" on the screen in several play sessions at home (Troseth, 2003b). We asked parents to do everything they could to help their children see the connection between video image and reality. Other children whose parents also had video equipment and willingness to participate were randomly assigned to a control group that did the play sessions but did not get live video experience. At the lab, all children received the typical brief orientation, during which they saw themselves and everything in the hiding room "live" on the TV. However, once the hiding game began, they could not see themselves on video. Now, they needed to realize that the image of the researcher hiding the toy on the TV showed them where to find it. That is, they needed to achieve what DeLoache (1987) calls *dual representation:* represent mentally both the 2D image that they saw on a TV screen, and the actual event happening behind the closed door to the room. On each subsequent trial, when they went to search for the toy, they needed to use the memory of what they saw on the screen, rather than a memory from direct experience (where they found the toy last time). The children who had received the special experience with live video were very successful at finding the toy, but a group without the live video experience was not (Troseth, 2003b).

This study also provided evidence that, although children's concepts of video and pictures may develop separately, some experiences build generalizable skills that are transferrable across symbolic media. The day after participating in the video search task, children returned to the lab and were asked to use a photo of the hiding place to find the hidden toy; that is, we asked them to transfer what they learned about using a video image from their at-home exposure and success in the video task to another symbolic medium. In a series of earlier studies, children of this age had failed to use pictures as clues in the search task (DeLoache & Burns, 1994). However, the children with live video experience were successful. The fact that the children transferred their insight across perceptual differences (from a moving image on a 20″ TV screen to small, static photos of furniture) suggested that they had gained some general understanding of how 2D images functioned as symbolic representations in this task, rather than just learning a procedure for how to respond to TV. Subsequently, in a correlational study based on parent questionnaire data, toddlers' real-world exposure to live video (defined as seeing themselves "live" on camcorder screens flipped to face them and/or on store security monitors) predicted their success in using video and pictures for information in search tasks at the lab (Troseth, Casey, Lawver, Walker, & Cole, 2007).

WHEN PEOPLE ON VIDEO RESPOND

Much has changed in the past decade regarding children's experience with digital media. For instance, children now can see themselves on live video on their parent's smartphone screen every time their parent takes a "selfie" family photo. Families are using video chat (Skype, FaceTime) with their young children on a regular basis, such as for regular visits with grandparents (McClure, Chentsova-Dutton, Barr, Holchwost, & Parrott, 2016; Tarasuik, Galligan, & Kaufman, 2011). Research indicates that if a person on video responds to them, children may be more likely to learn from the person, particularly if their parent views with them and supports their understanding. In the pre-Skype era, we used closed-circuit video to allow an on-screen researcher to talk to children (Troseth et al., 2006). The researcher conversed with the parent about the child's pet or birthday, then engaged the child in a game of "Simon Says," commenting when the child did/did not follow instructions. After singing a song with the parent and child, the researcher informed the child that there was a sticker under the child's chair, commenting on the sticker once the child (invariably) stuck it on his/her shirt. Thus, the researcher on the screen demonstrated that she could provide information that related to the child's ongoing experience. Following this 5-min contingent video interaction, children readily used verbal cues offered by the on-screen researcher to locate a hidden object, unlike in prior research in which no such interaction was offered (Schmidt et al., 2007). In more recent research, video chat experience has supported toddlers' use of video-presented information in a variety of contexts: copying a sequence of actions with a tool to open a box (Nielsen, Simcock, & Jenkins, 2008), learning verbs (Roseberry, Hirsh-Pasek, & Golinkoff, 2014), and recognizing, preferring, and learning from video chat partners (Myers, LeWitt, Gallo, & Maselli, 2016).

In the future, researchers must examine how being exposed to live video through regular video-chat visits may affect children's concepts of, and learning from, video. Tarasuik et al. (2011) have shown that young children are comforted by interacting with their parents via Skype during the separation phase of the "Strange Situation." It will be important to discover whether interacting with people by means of video chat leads to conceptual change, such that children seem to better understand video as a representation of reality at an earlier age. Another possibility is that children will respond to and learn from a person on video while the person interacts with them, but will not transfer this understanding across people or contexts. Recent work from our lab shows that while video chat increased 2.5-year-olds' engagement with an on-screen actress, the addition of scaffolding through parent participation was crucial for supporting learning (Strouse, Troseth, & Saylor, 2016). Thus, even when using live video, children may benefit when adults point out that video can function as a representational medium, similar to the way parents communicate the use of pictures as representations during picture book reading.

CONCEPTUAL DEVELOPMENT THROUGH ACTIVE CO-VIEWING

The way that adults co-view and talk about video with children may help children realize that the medium can be informative. We have suggested three possible mechanisms by which active co-viewing supports young children's learning from video (Strouse et al., 2013). First, when adults model attention to the screen, they guide children's attention. Demers, Hanson, Kirkorian, Pempek, and Anderson (2013) showed that when an infant's look to the screen follows their parent's look, the infant tends to look at the screen longer. The human tendency to follow gaze results in parent and child sharing attention to the video the same way they would jointly focus on a book or object in the environment. Because such triadic interactions are a frequent context for learning, young viewers may assume that what is on screen is informative.

In a recent study, we tested the role of parents' attention on children's learning (Strouse et al., 2016). An on-screen person led a game of "Simon Says," sang a song, and played other games (as in our earlier study), but this time, parents did not converse with the person. In one condition, the parent sat beside the child facing the TV and was asked to "play along" with the person (modeling participation), but not to instruct their child in any way. In the other condition, the parent was told we wanted to see what their child would learn on their own; the parent sat beside their child but faced away from the TV. After the games, the on-screen person labeled a novel object with a new word. Children whose parents modeled attention to the video participated more during the games and more often selected the correct object from the video as the referent of the new label. This study shows the power of the attention mechanism: parents merely "playing along" and sharing focus on the screen supported children's engagement and learning.

Active parental co-viewing may also provide cognitive support for learning, such as giving children a familiar context for encoding information from the screen (e.g., connecting it to real life), or practice retrieving, rehearsing, and applying what they have seen on screen through conversation and questioning. In a study that also used on-screen games followed by a demonstration of a new word, some 24-month-olds watched a pre-recorded demonstration without parental support (the parent was in the room but not actively participating), and the rest received a simple parent scaffold during the labeling event (Strouse & Troseth, 2014). Parents held up replicas of two on-screen objects and stated, "These are the same as the ones on TV," once before and once after the person on screen labeled one of the toys. Parents never repeated the novel word given in the video, so children had to learn it from the on-screen person. Parents simply but explicitly drew the connection between the video image of the toys and the actual toys. The simple act of aligning the real objects with their images on the screen may have helped children to encode the two mental representations of the objects together in memory (a representation of the 2D images and one of the 3D toys)—a cognitive support for learning. Even given this small amount of parental support, children scored better on a post-test of their word knowledge.

During active co-viewing, parents also may provide verbal and nonverbal social feedback to their children. In the prior study, for instance, parents provided a responsive scaffold: they placed the objects in the child's line of sight to draw their gaze and spoke directly to the child. This behavior not only provided factual information but also provided social information through responsive gaze, placement, and timing. Whereas attention direction and cognitive information may be provided by on-screen features such as sound effects, visual highlighting, or questioning by an on-screen character, social feedback is responsive and personal in nature and much more difficult to achieve without an in-person co-viewer.

Even children older than those discussed so far benefit from active parental co-viewing. In a study with commercial video storybooks (Strouse et al., 2013), 3-year-olds were assigned to either a: (1) *dialogic questioning* condition, in which parents were trained to pause the videos and engage in high-quality questioning and conversations with their children about the videos; (2) *directed attention* condition, in which parents paused the videos and labeled things on screen but did not ask questions; (3) *dialogic actress* condition, in which an on-screen actress presented dialogic-style questions; or (4) *control* condition in which children watched "as usual" (typically without parents). Children in the different conditions therefore received different levels of help (or no help) through the various mechanisms of support. After watching the video stories for 4 weeks, children were tested on their story comprehension and learning of story vocabulary. Although all of the groups did learn some information from the video (even the "watch as usual" control group), children whose parents engaged in dialogic questioning while co-watching the videos (which incorporated attention-directing, cognitive, and social feedback supports) outscored the other groups. Children in the directed attention group did not score significantly better than the control group, showing the importance of asking questions and providing feedback (cognitive and social feedback mechanisms) to support 3-year-olds children's learning from video. The scores of children in the dialogic actress group were closer to those of the dialogic parent group, indicating that an on-screen person asking cognitively challenging questions may benefit learning—even without the attention or social feedback of an in-person partner.

The three mechanisms by which parental mediation operates (directing attention, supporting cognition, and social feedback) combine to help children *learn how to learn* from digital media: the adult directs children's attention to important information, helps children to interpret what they see, and supports children's transfer of information to their own lives. By talking about video and making connections to the real world, adults show that video can be attention-worthy and that the content of video can be meaningful and relevant to the current situation. Adult support may help override young children's tendency to dismiss video as unrelated to reality, acting as a cue that the situation of co-viewing is pedagogical and video content is worth learning.

DIGITAL LITERACY IN THE FUTURE

Recent developments in digital media may incorporate some of the benefits seen with active parental co-viewing, or lessen the perceptual and conceptual mismatches between video and reality, supporting children's learning from video. For instance, children who interacted more with the on-screen questioning characters in *Dora the Explorer* demonstrated increased comprehension of a TV episode (Calvert, Strong, Jacobs, & Conger, 2007). In newer touchscreen games, characters can respond to a child's touch or voice, adding an authentic bi-directional component to what was previously only an imitation of contingency (Troseth, Russo, & Strouse, 2016). Game applications (apps) can also adapt to a child's performance: when a child answers a question incorrectly, the app provides meaningful feedback based on the child's previous responses, partially simulating one of the mechanisms of parental support. Educational technology featuring interactive artificial intelligence has the potential to provide customized feedback. For example, older infants (18–21 months) learned more from an interactive stuffed animal that was personalized to them (e.g., knew the child's name and favorite things) than from a noncustomized stuffed animal (Calvert, Richards, & Kent, 2014).

Video that responds contingently to body movement (e.g., Wii, LeapTV) and virtual reality technology are blurring the perceptual distinction between video and reality. Televisions with 3D capabilities are becoming more common, and devices such as the Oculus Rift immerse users in a fully interactive 3D virtual environment. It remains to be determined how experience with these new kinds of media might change the course of development regarding digital literacy.

LEARNING TO LEARN FROM VIDEO

From early in development, children begin to learn to use the representational systems and symbolic media of their culture. Video is one of these symbolic media. As we have shown, adults cannot take for granted that young children detect and understand the representational role of video images. As DeLoache (1995) noted, "There are no fully transparent symbols." Instead, using symbols maturely requires developing an understanding of their nature and function. Learning to learn from video depends on children's perception of similarities, as well as differences, between video and reality.

Just as learning to read is supported by both experience with and direct instruction of the component skills of literacy such as phonics, vocabulary, and background contextual knowledge, digital literacy is supported through experience with and direct instruction of the component skills of representational competence. Highlighting the ways in which particular videos match current reality helps toddlers learn: this may include aligning the content of a video with the child's ongoing experience and using the social contingency possible in live video. Similar to the way children learn to take a referential stance to pictures during picture book reading with parents, social support promotes children's understanding of video as an informative representational medium. Adult co-viewing that includes three mechanisms of support (directing attention, supporting cognition, and social feedback) helps children to learn more from video than they do on their own. Such experiences and supports help children to develop representational skills they will use to learn from symbolic media throughout life.

REFERENCES

Allen Preissler, M., & Carey, S. (2004). Do both pictures and words function as symbols for18-and 24-month-old children? *Journal of Cognition and Development*, *5*, 185–212. http://dx.doi.org/10.1207/s15327647jcd0502_2.

Baillargeon, R., Needham, A., & Devos, J. (1992). The development of young infants'intuitions about support. *Early Development and Parenting*, *1*, 69–78. http://dx.doi.org/10.1002/edp.2430010203.

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review*, *30*, 128–154. http://dx.doi.org/10.1016/j.dr.2010.03.001.

Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Development Perspectives*, *7*, 205–210. http://dx.doi.org/10.1111/cdep.12041.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior & Development*, *19*, 159–170. http://dx.doi.org/10.1016/S0163-6383(96)90015-6.

Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development*, *70*, 1067–1081. http://dx.doi.org/10.1111/1467-8624.00079.

Barr, R., Muentener, P., & Garcia, A. (2007). Age-related changes in deferred imitation from television by 6-to 18-month-olds. *Developmental Science*, *10*, 910–921. http://dx.doi.org/10.1111/j.1467-7687.2007.00641.x.

Bigelow, A. E. (1996). Infants' memory for contingently responding persons. *Infant Behavior & Development*, *19*(Special ICIS Issue), 334. http://dx.doi.org/10.1016/S0163-6383(96)90388-4.

Bigelow, A. E., & Birch, S. A. J. (2000). The effects of contingency in previous interactions on infants' preference for social partners. *Infant Behavior & Development*, *22*, 367–382. http://dx.doi.org/10.1016/S0163-6383(99)00016-8.

Bigelow, A. E., MacLean, B. K., & MacDonald, D. (1996). Infants' response to live and replay interactions with self and mother. *Merrill Palmer Quarterly*, *42*, 596–611. www.jstor.org/stable/23087472.

Butler, J., & Rovee-Collier, C. (1989). Contextual gating of memory retrieval. *Developmental Psychobiology*, *22*, 533–552. http://dx.doi.org/10.1002/dev.420220602.

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., & Behne, T. (2011). Early social cognition in three contexts. *Monographs of the Society for Research in Child Development*, *299*(76), 93–104.

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*, 148–155. http://dx.doi.org/10.1016/j.appdev.2014.03.004.

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. http://dx.doi.org/10.1080/15213260701291379.

Chall, J. S. (1983). *Learning to read: The great debate*. New York: McGraw-Hill.

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238*, 1556–1557. http://www.jstor.org/stable/1700975.

DeLoache, J. S. (1995). Early understanding and use of symbols: The model model. *Current Directions in Psychological Science*, *4*, 109–113. http://dx.doi.org/10.1111/1467-8721.ep10772408.

DeLoache, J. S., & Burns, N. M. (1994). Early understanding of the representational function of pictures. *Cognition*, *52*(2), 83–110. http://dx.doi.org/10.1016/0010-0277(94)90063-9.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science*, *9*, 205–210. http://dx.doi.org/10.1111/1467-9280.00039.

DeLoache, J. S., Strauss, M. S., & Maynard, J. (1979). Picture perception in infancy. *Infant Behavior & Development*, *2*, 77–89. http://dx.doi.org/10.1016/S0163-6383(79)80010-7.

Demers, L. B., Hanson, K. G., Kirkorian, H. L., Pempek, T. A., & Anderson, D. R. (2013). Infant gaze following during parent-infant coviewing of baby videos. *Child Development*, *84*, 591–603. http://dx.doi.org/10.1111/j.1467-8624.2012.01868.x.

Diener, M., Pierroutsakos, S. L., Troseth, G. L., & Roberts, A. (2008). Video versus reality: Infants' attention and affective responses to video and live presentations. *Media Psychology*, *11*, 418–441. http://dx.doi.org/10.1080/15213260802103003.

Ganea, P. A., Allen, M. L., Butler, L., Carey, S., & DeLoache, J. S. (2009). Toddlers' referential understanding of pictures. *Journal of Experimental Child Psychology*, *104*, 283–295. http://dx.doi.org/10.1016/j.jecp.2009.05.008.

Gelman, S. A., Chesnick, R. J., & Waxman, S. R. (2005). Mother-child conversations about pictures and objects: Referring to categories and individuals. *Child Development*, *76*, 1129–1143. http://dx.doi.org/10.1111/j.1467-8624.2005.00876.x-i1.

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development*, *63*(1) Serial No. 253. http://dx.doi.org/10.2307/1166211.

Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant-adult interactions. *Infant Behavior & Development*, *19*, 49–61. http://dx.doi.org/10.1016/S0163-6383(96)90043-0.

Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, *58*(1) Series No. 231. http://dx.doi.org/10.2307/1166074.

Hayes, L. A., & Watson, J. S. (1981). Facial orientation of parents and elicited smiling by infants. *Infant Behavior & Development*, *4*, 333–340. http://dx.doi.org/10.1016/S0163-6383(81)80035-5.

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

Held, R., Birch, E., & Gwiazda, J. (1980). Stereoacuity of human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *77*, 5572–5574. http://dx.doi.org/10.1073/pnas.77.9.5572.

Kuhl, P. K., Tsao, F. -M., & Liu, H. -M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101. http://dx.doi.org/10.1073/pnas.1532872100.

Leslie, A. M. (1987). Pretense and representations: The origins of "theory of mind" *Psychological Review*, *94*, 412–426. http://dx.doi.org/10.1037/0033-295X.94.4.412.

Marian, V., Neisser, U., & Rochat, P. (1996). Can 2-month-old infants distinguish live from videotaped interactions with their mother. *Emory Cognition Project Report # 33*. Atlanta: Emory University.

McCall, R. B., Parke, R. D., Kavanaugh, R. D., Engstrom, R., Russell, J., & Wycoff, E. (1977). Imitation of live and televised models by children one to three years of age. *Monographs of the Society for Research in Child Development*, 1–94. http://dx.doi.org/10.2307/1165913.

McClure, E. R., Chentsova-Dutton, Y. E., Barr, R. F., Holchwost, S. J., & Parrott, W. G. (2016). "Facetime doesn't count": Video chat as an exception to media restrictions for infants and toddlers. *International Journal of Child-Computer Interaction*. http://dx.doi.org/10.1016/j.ijcci.2016.02.002.

Moser, A., Zimmerman, L., Dickerson, K., Grenell, A., Barr, R., & Gerhardstein, P. (2015). They can interact but can they learn? Toddlers' transfer from touchscreens and television. *Journal of Experimental Child Psychology*, *137*, 137–155. http://dx.doi.org/10.1016/j.jecp.2015.04.002.

Muir, D., Hains, S. M. J., Cao, Y., & D'Entremont, B. (1996). Three-to six-month olds' sensitivity to adult intentionality: The role of adult contingency and eye direction in dyadic interactions. *Infant Behavior & Development*, *19*, 200. http://dx.doi.org/10.1016/S0163-6383(96)90254-4.

Mumme, D. L., & Fernald, A. (2003). The infant as onlooker: Learning from emotional reactions observed in a television scenario. *Child Development*, *74*(1), 221–237. http://dx.doi.org/10.1111/1467-8624.00532.

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. Fields & N. Fox (Eds.), *Social perception in infants* (pp. 177–197). Norwood, NJ: Ablex.

Myers, L. J., LeWitt, R. B., Gallo, R. E., & Maselli, N. M. (2016). Baby FaceTime: Can toddlers learn from online video chat? *Developmental Science*. http://dx.doi.org/10.1111/desc.12430.

Nawrot, E., Mayo, S. L., & Nawrot, M. (2009). The development of depth perception from motion parallax in infancy. *Attention, Perception, & Psychophysics*, *71*(1), 194–199. http://dx.doi.org/10.3758/APP.71.1.194.

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*(5), 722–731. http://dx.doi.org/10.1111/j.1467-7687.2008.00722.x.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language*, *5*, 1–15. http://dx.doi.org/10.1017/S0305000900001896.

Pierroutsakos, S. L., & Troseth, G. L. (2003). Video verite: Infants' manual investigation of objects on video. *Infant Behavior & Development*, *26*(2), 183–199. http://dx.doi.org/10.1016/S0163-6383(03)00016-X.

Rice, M. L., Huston, A. C., & Wright, J. C. (1986). Replays as repetitions: Young children's interpretation of television forms. *Journal of Applied Developmental Psychology*, 7(1), 61–76. http://dx.doi.org/10.1016/0193-3973(86)90019-5.

Rochat, P., & Morgan, R. (1995). Spatial determinants in the perception of self-produced leg movements in 3- to 5-month-old infants. *Developmental Psychology*, *31*(4), 626–636. http://dx.doi.org/10.1037/0012-1649.31.4.626.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, *85*, 956–970. http://dx.doi.org/10.1111/cdev.12166.

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. http://dx.doi.org/10.1207/S1532785XMEP0401_03.

Schmitt, K. L., & Anderson, D. R. (2002). Television and reality: Toddlers' use of visual information from video to guide behavior. *Media Psychology*, *4*(1), 51–76. http://dx.doi.org/10.1207/S1532785XMEP0401_03.

Slater, A., Rose, D., & Morison, V. (1984). Newborn infants' perception of similarities and differences between two- and three dimensional stimuli. *British Journal of Developmental Psychology*, *2*(4), 287–294. http://dx.doi.org/10.1111/j.2044-835X.1984.tb00936.x.

Snow, C. (2015). *50 years of research on reading. In M. J. Feuer, A. I. Berman, & R. C. Atkinson (Eds.), Past as prologue: The National Academy of Education at 50; Members reflect* (pp. 175–179). Washington, DC: National Academy of Education. http://www.naeducation.org/cs/groups/naedsite/documents/webpage/naed_169315.pdf#page=194.

Soden, B., Christopher, M. E., Hulslander, J., Olson, R. K., Cutting, L., Keenan, J. M., et al. (2015). Longitudinal stability in reading comprehension is largely heritable from grades 1 to 6. *PLoS ONE*. *10*(1).e0113807. http://dx.doi.org/10.1371/journal.pone.0113807.

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*(4), 262–280. http://dx.doi.org/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. http://dx.doi.org/10.1016/j.cogdev.2014.01.002.

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology*, *49*, 2368–2382. http://dx.doi.org/10.1371/journal.pone.0113807.

Strouse, G.A., Troseth, G.L., O′Doherty, K., & Saylor, M.M. (2016). Co-viewing and contingency support toddlers' word learning from video. Manuscript submitted for publication.

Suddendorf, T. (2003). Early representational insight: Twenty-four-month-olds can use a photo to find an object in the world. *Child Development*, *74*, 896–904. http://dx.doi.org/10.1111/1467-8624.00574.

Tarasuik, J. C., Galligan, R., & Kaufman, J. (2011). Almost being there: Video communication with young children. *PLoS ONE*. *6*(2). e17129. http://dx.doi.org/10.1371/journal.pone.0017129.

Troseth, G. L. (2003a). Getting a clear picture: Young children's understanding of a televised image. *Developmental Science*, *6*(3), 247–253. http://dx.doi.org/10.1111/1467-7687.00280.

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. http://dx.doi.org/10.1037/0012-1649.39.1.140.

Troseth, G. L. (2010). Is it life or is it Memorex? Video as a representation of reality. *Developmental Review*, *30*, 155–175. http://dx.doi.org/10.1016/j.dr.2010.03.007.

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. http://dx.doi.org/10.1080/15248370701446772.

Troseth, G. L., & DeLoache, J. S. (1998). The medium can obscure the message: Young children's understanding of video. *Child Development*, *69*, 950–965. http://dx.doi.org/10.1111/j.1467-8624.1998.tb06153.x.

Troseth, G. L., Pierroutsakos, S. L., & DeLoache, J. S. (2004). From the innocent to the intelligent eye: The early development of pictorial competence. R. Kail (Ed.), *Vol.32. Advances in child development and behavior* (pp. 1–35). Amsterdam: Elsevier. http://dx.doi.org/10.1016/s0065-2407(04)80003-x

Troseth, G. L., Russo, C. E., & Strouse, G. A. (2016). What's next for research on young children's interactive media? *Journal of Children and Media*, *10*, 54–62. http://dx.doi.org/10.1080/17482798.2015.1123166.

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. http://dx.doi.org/10.1111/j.1467-8624.2006.00903.x.

Walker, C. M., Walker, L. B., & Ganea, P. A. (2013). The role of symbol-based experience in early learning and transfer from pictures: Evidence from Tanzania. *Developmental Psychology*, *49*, 1315–1324. http://dx.doi.org/10.1037/a0029483.

Zack, E., Barr, R., Gerhardstein, P., Dickerson, K., & Meltzoff, A. N. (2009). Infant imitation from television using novel touch screen technology. *British Journal of Developmental Psychology*, *27*, 13–26. http://dx.doi.org/10.1348/026151008X334700.