MEDIA PSYCHOLOGY, *9,* 389–409

Two-Year-Olds' Object Retrieval Based on Television: Testing a Perceptual Account

Marie Evans Schmidt
Center on Media and Child Health, Children's Hospital Boston

Alisha M. Crawley-Davis and Daniel R. Anderson
University of Massachusetts at Amherst

Two-year-olds frequently fail to use information provided by video to find objects hidden in an adjacent room. Schmitt and Anderson (2002) hypothesized that they fail to map the 2-dimensional (2D) video image onto the 3D layout of the search space. Two experiments tested whether 2-year-olds can successfully use information from video when the search space is 2D or when the information is provided verbally (by telling the child where the toy is hidden). In both experiments, children performed poorly in the video conditions but performed well in direct live experience comparison conditions, contradicting Schmitt and Anderson's hypothesis. Performance was above chance on the first trial in the video conditions, suggesting that 2-year-olds do have a memory of the hiding location, albeit one that is easily disrupted by perseverative errors on subsequent trials. Overall, the results are most consistent with the hypothesis that very young children give priority to direct experience over mediated information.

Despite the proliferation of TV shows and videos designed for very young children, there is, as yet, very little research showing that infants and toddlers learn anything of value from them (for a review, see Anderson & Pempek, 2004). This is in contrast to children aged 3 years and older, where there is a substantial body of evidence that they readily learn from video (for a review, see Huston & Wright, 1997).

Recent research with children 2 years and younger, in fact, suggests a *video deficit* in early learning; that is, very young children do not appear to learn as much

Correspondence should be sent to Marie Evans Schmidt, Ph.D., Center on Media and Child Health, Children's Hospital Boston/Harvard Medical School, 300 Longwood Ave., Boston, MA 02115. E-mail: marie.evans@childrens.harvard.edu

from television as they do from equivalent real life experiences (Anderson & Pempek, 2004). Although 2-year-olds can sustain attention to television programs (Schmitt, Anderson, & Collins, 1999) and differentiate comprehensible and incomprehensible TV content (Anderson, Lorch, Field, & Sanders, 1981), they imitate simple actions seen on television less accurately than actions seen live (Hayne, Herbert, & Simcock, 2003) and learn verbal labels for objects less well from television than from live presentations (Grela, Lin, & Krcmar, 2003).

This video deficit is apparent when 2-year-old children are asked to use information from video to retrieve a real object. In the standard video task, children watch on video as a toy is hidden in an adjacent room. Troseth and DeLoache (1998) found that 2-year-olds had difficulty finding the toy after watching it being hidden on video in one of six locations (44% correct over 4 trials). Similarly, Schmitt and Anderson (2002) found that 2-year-olds had great difficulty finding a toy when they watched the hiding event on video in one of four locations (25% correct over 4 trials). Thus, 2-year-olds in both studies were much less likely to find a toy if they watched it being hidden on video.

Despite the similarities in their findings, these two sets of authors offer very different explanations for them. Troseth and DeLoache (1998) hypothesized that 2-year-olds do not consider television relevant to reality and thus fail to appreciate the TV as showing a representation of the hiding space. They argued that 2-year-old children have learned from experience that television does not affect them directly and that objects and events on television are pretend.

Schmitt and Anderson (2002), in contrast, hypothesized that 2-year-olds may have difficulty because perceptual information on TV concerning three-dimensional (3D) space is impoverished relative to perceptual information presented in reality. For example, the perceptual information provided by motion parallax is absent on television. That is, if a child moves his or her head when looking at a real 3D display, nearer objects appear to move with a greater angular displacement than more distant objects. Motion parallax is an important cue in judging the relative distance of objects and the layout of space. No such perceptual phenomenon occurs when looking at a video image. Because of this and other perceptual differences between television and reality (for example, the poor resolution of TV reduces perception of texture gradients which also provide cues to depth), Schmitt and Anderson suggested that very young children might have difficulties mapping the television image onto the 3D space of the adjacent room. As a result of these mapping difficulties, young children may fail to see any connection between the TV screen and the search space.

In addition, prior research has suggested that the television display is interpreted by young children as consisting of small objects contained within the space of the television cabinet (Lyle & Hoffman, 1972; Nikken & Peters, 1988). Therefore, Schmitt and Anderson (2002) hypothesized that young children represent objects on television as small graspable items contained in a space not navigable by

their own bodies. To succeed on the standard video task, then, 2-year-olds would have to rely on an allocentric or objective representation of the search space, ignoring their egocentric representation based on video information. According to Schmitt and Anderson (2002), all of these difficulties place information processing burdens on the child, reducing encoding and performance on the retrieval task. Note that this theory does not assume that the 2-year-old child has a fundamental problem understanding that television can represent a real state of affairs or that the child assumes that events on television are not real, in contrast to Troseth and DeLoache (1998).

The video information, nevertheless, appears to help some of the children guide retrieval on their first trial. Schmitt and Anderson (2002) found that first trial performance was superior to performance on subsequent trials (60% on the first trial, 33% on trial 2, 0% on trial 3, and 9% on trial 4). While watching the video, children's attention is likely drawn to the piece of furniture hiding the object. This can create a memory of that piece of furniture that primes the subsequent search. In terms of Schmitt and Anderson's theory, having such a recent and probably transient memory is not the same as having representational insight that leads to knowledge about where the object is hidden. Schmitt and Anderson thus hypothesized that, on all trials after the first, the relatively poor representation of the hiding place is overwhelmed by the rich 3D experience of actually having found the toy on the prior trial, a form of proactive interference. Children are thus likely to make a perseveration error by going back to the location where they found the toy previously. The children's most common mistakes on trials 2 through 4 were, in fact, perseveration errors.

The goal of our experiments was to test when and how young children can use information from television to guide behavior in a search task. In the first experiment, we substantially reduced the perceptual differences between the display space as depicted on video and the actual search space. This was accomplished by using a task, previously suggested by Schmitt and Anderson (2002), in which 2-year-olds searched for a sticker on a felt board that was the same size as the TV screen. By using a two-dimensional (2D) search space that is the same size as the 2D video screen, there is no need for the child to form a 3D representation of the search space, thus eliminating the 2D to 3D mapping perceptual problem theorized by Schmitt and Anderson (2002). Also, by using objects on the search space that are the same size as the objects displayed on the video screen, the representations that the child forms with respect to the video objects are consistent with those necessary to retrieve the actual object. Because the perceptual problems of 3D mapping and coordinating allocentric and egocentric representations have been eliminated, search performance on the basis of video should be good.

An additional experimental condition used a felt board contained in a box with a transparent plastic screen as the display space. This condition was included as a partial test of Troseth and DeLoache's (1998) hypothesis that video performance is

poor because 2-year-olds consider objects and events on TV as pretend and irrelevant to reality. If children perform well in this condition and poorly in the video condition, their hypothesis is supported insofar as the display felt board is patently real. On the other hand, if 2-year-olds perform poorly in this condition, the result would be inconsistent with the Schmitt and Anderson (2002) theory that predicts good performance. A third, unmediated, control condition involved children searching for a sticker on the same felt board on which it was hidden. Good performance was expected in this condition.

In Experiment 2, an experimenter verbally told the children where to search for the hidden toy, either live, standing directly in front of the child (unmediated condition) or via closed circuit video (mediated video condition). If 2-year-olds' difficulty using information from video is primarily visual and perceptual in nature, as Schmitt and Anderson (2002) hypothesized, both groups would be expected to succeed on this task, as the need for visually mapping the 2D image of the toy's location onto the 3D search space (as in Experiment 1) was completely eliminated. Alternatively, if, as Troseth and Deloache (1998) hypothesized, young children consider information from video of little or no relevance to reality, children in the mediated video condition should fail to find the hidden object, whereas children should succeed in the unmediated condition.

EXPERIMENT 1

Participants

Fifty 2-year-olds (Range = 23 to 25 months) were randomly assigned to the live unmediated experience group (8 boys, 9 girls), felt board in a box group (8 boys, 8 girls), or the video group (8 boys, 9 girls). Children were recruited from state birth records via a letter and a follow up telephone call. Eight additional children were dropped from the data analysis (5 due to failure to complete the task and 3 due to experimenter error).

Materials and Setting

One room (4.26 m × 3.81 m) was used for testing. A floor to ceiling curtain, extending from one end of the room to the other, divided the test room into 2 rectangular spaces. A table (1.98 m × .76 m) was placed in the center of the room; the table's back edge rested against the curtains. Another floor to ceiling curtain ran along the front edge of this table, extending half its length (about 1 m; see Figure 1).

Two identical, rectangular felt boards (31.5 cm × 37 cm) were used in the study. Black borders were placed along the edges of the felt boards to make the hiding

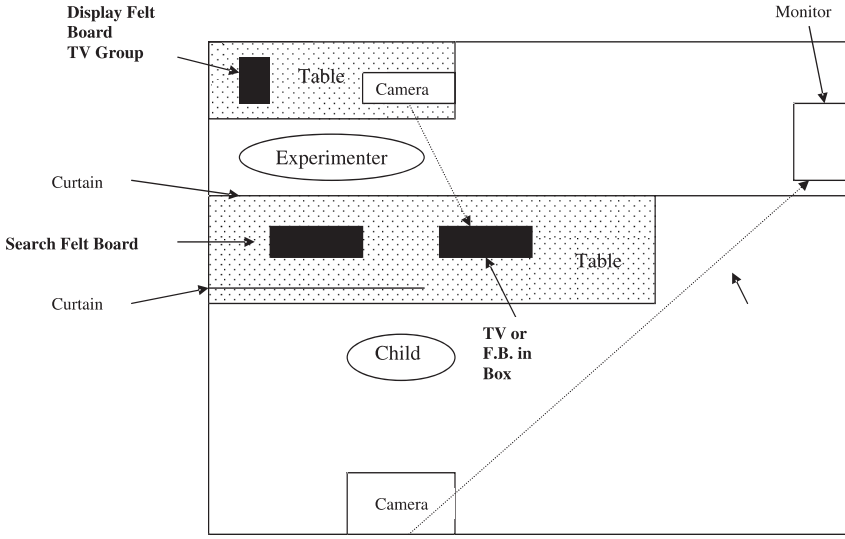

FIGURE 1 Layout of room for Experiment 1.

space on each the same size as the video screen. Four felt hiding objects, representing a balloon (6 cm × 4.5 cm), two different-colored gift-wrapped boxes (7.5 cm × 4.5 cm and 5 cm × 4.5 cm), and a birthday cake (5 cm × 7.5 cm), were used as the hiding places. A small teddy bear sticker (1.5 cm × 1.5 cm) was used as the hiding object. See Figure 2.

Two easels were used to hold the felt boards, except in the felt board in a box condition. In this condition, a rectangular box (41 cm × 41 cm × 78 cm) with a clear Plexiglas front, wooden sides, and no back was used to hold one of the felt boards.

A 13-inch color television set was used for the child to watch the hiding event in the video condition. A video camera, on an additional table behind the curtains, was used to show the hiding event via real-time closed circuit TV, also in the video condition.

Another camera facing the felt board displays was used to record the child; this same camera, because it was also connected to a small video monitor behind the curtains, allowed the experimenter to guide her hands when hiding the sticker. Only the experimenter's hands were visible to the children in all conditions because there was no way in the video condition to record the hiding event so that the experimenter's face would be visible (as it would have been in the other conditions) and the felt board on video would still be the same size as the search felt board.

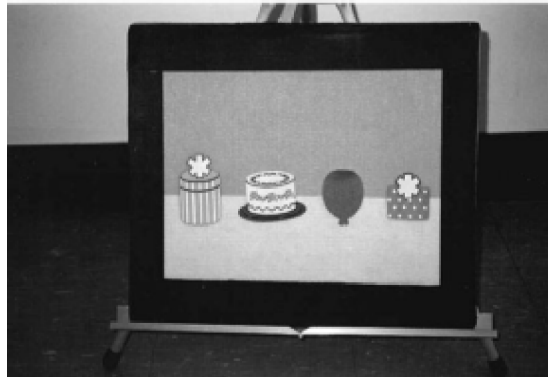

FIGURE 2 Search Felt Board for Experiment 1.

Procedure

Orientation. After a brief warm up in the waiting room, the child entered the testing room, where he or she sat on the parent's lap in front of the table. An orientation followed. The experimenter's assistant pointed out all four hiding places on the search felt board and the sticker, saying "We're going to play a hiding game. This is my felt board, and all these things are on it. [Experimenter] likes to hide this teddy bear [the sticker] on the felt board." She placed the sticker behind each felt object in turn, saying "Sometimes she hides him here," and then removed the felt object, displaying the sticker underneath.

The children in the video group additionally participated in video correspondence training. The child turned towards the television set, and the assistant pointed out the correspondence between the video display and the felt board. She said, "This is my special TV that shows my felt board. Look, it's my felt board on TV!" The assistant then asked the child to touch the felt object (on the search felt board) that she pointed to on the television set. She said: "I'm touching this one. Can you touch this one on the felt board?" The child was encouraged to touch the correct object on the felt board if he or she did not do so initially. This continued for each location.

Children in the felt board in a box group also participated in correspondence training. The assistant showed the child both felt boards, saying about the felt board in the box, "This is my special box. It has a felt board inside it just like this one. See, it looks the same as this felt board." For each felt object, the experimenter pointed to it on the felt board in the box and asked the child to point to it on the search felt board, saying "I'm touching this one. Can you touch this one on that felt board?"

Test Trials. Once training was complete, each child participated in 4 test trials. At the beginning of each trial, the assistant showed the child a bear sticker. The child was told that he or she was going to help find the sticker but that first the experimenter would demonstrate where to look for it. As the experimenter walked behind the curtain, the assistant said: "Now [experimenter] is going to go behind the curtain. There she goes." In the unmediated condition, the assistant then said: "She's going to hide the teddy bear sticker, and we're going to watch her." In the video condition, the assistant said "She's going to hide the teddy bear sticker, and we're going to watch her on television." In the felt board in a box condition, the assistant said "[Experimenter] is going to go behind the curtain. She's going to hide the sticker, but we're not going to watch her. She's going to show us where she hid it."

In the unmediated condition, the experimenter walked behind the curtains and hid the sticker on the search felt board, using the TV monitor to her left (behind the curtains) to guide her hands. A short delay between trials was deliberately introduced equate times used by the experimenter to hide the sticker on the search felt board in the other two experimental conditions. During this delay, the assistant blocked the child's view of the search felt board by pulling the curtain in front of it.

In the video condition, the experimenter walked behind the curtains and immediately hid the sticker on the search felt board. During this time, the front curtains remained closed, so the child could not see the experimenter hiding the sticker on the search felt board. She then turned and hid another sticker on the felt board on the back table (facing the video camera). The camera recorded this hiding event. Immediately after the experimenter hid the sticker on camera, the television set was turned off, and the assistant pulled back the curtain covering the search felt board, allowing the child to search for the sticker.

For the felt board in a box group, the experimenter went behind the curtain and hid the sticker on the felt board in the box. Again, she used the video monitor to her left to guide her hands. After the experimenter hid the sticker on the felt board in the box, the curtain was pulled back from the search felt board, and the child was immediately allowed to search on it. As the child searched, the experimenter held black construction paper against the front of the box, so that the child could not look back at the felt board inside it.

The assistant followed an identical script for all the groups. When the experimenter's hands emerged from behind the curtain (in the unmediated and felt board in the box conditions), or on the TV screen (in the video condition), the assistant said "Look, [experimenter] is hiding the bear now. Look where she's hiding it…Remember where it is." The location was not labeled.

In all conditions, the assistant said, "Find the teddy bear on the felt board," and pointed to the search felt board. If the child was not correct on the first attempt, the assistant said: "Try a different place," until the child found the sticker. One prompt, "It's in the same place that she hid it on [TV, that felt board]" was used after the first unsuccessful search.

Two independent observers scored each videotaped test session for whether the child found the sticker on the first attempt, number of attempts to find the sticker, and whether the child made perseverative errors on trials 2 through 4. They agreed on 98% of the trials as to whether the child's first reach was successful as well as the targets of the reaches. Disagreements were resolved through discussion as both observers viewed the videotape of the child again.

Results

The main dependent variable was the percentage of errorless retrievals. An errorless retrieval occurred when the child found the sticker in the first place he or she searched. The percentage of errorless retrievals was calculated as the number of correct retrievals divided by the number of complete trials times 100 for each child.

A 3 (conditions: unmediated, video, box) by 2 (gender) ANOVA revealed a main effect of condition, $F(1, 44) = 25.76$, mse = 1.402, $p < .001$. The 2-year-olds had little difficulty retrieving the sticker when they watched it hidden directly, in the unmediated condition (82% errorless retrievals, $SD = 23\%$) and this performance was clearly superior to the video group (30%, $SD = 20\%$) and the felt board in a box group (33%, $SD = 28\%$). In the unmediated condition performance over the four trials combined was significantly better than chance (25 %), $t(16) = 10.29$, $p < .001$, but in the video condition, performance was not significantly better than chance, $t(16) = 1.00$, $p > .10$. In the felt board in a box condition, performance was also not better than chance, $t(15) = 1.17$, $p < .10$.

To determine whether Trial 1 performance was superior to that of later trials, binomial tests were used, with the mean proportion of errorless retrievals on Trials 2 through 4 (for each condition) as the null hypothesis probability against which Trial 1 performance could be compared. The first trial performance (53% correct) in the video condition was significantly different (binomial test) from the mean percentage of errorless retrievals on Trials 2 through 4 in that group (25%, $p < .01$). Performance was significantly above chance on the first trial (53% errorless re-

TABLE 1
Mean Percent Errorless Retrievals in Experiments 1 and 2

Condition	*All trials*	*Trial 1 only*	*Trials 2–4*
Experiment 1			
Unmediated	82	70	86
Box	33	27	34
Video	30	53	25
Experiment 2			
Unmediated	64	60	62
Video	20	38	15

trievals, $p < .01$), but was not greater than chance on trials 2, 3, and 4 (23%, 25%, and 18%, respectively).

First trial performance in the felt board in a box condition (27%) was not significantly different from the mean percentage of errorless retrievals on Trials 2 through 4 in that condition (34% errorless retrievals). Performance on trials 2, 3, and 4 was 43%, 31%, and 29% respectively.

First trial performance in the unmediated condition (70%) was not significantly different from the mean percentage of errorless retrievals on Trials 2 through 4 (86%). Performance was significantly above chance on the first trial, and remained above chance on all subsequent trials (88%, 77%, and 94%, respectively).

First trial performance in the unmediated condition (70 %) was superior to the felt board in a box condition (26%), $\chi^2(1) = 6.87$, $p < .01$. First trial performance in the video condition was at an intermediate level (52%) but was not significantly different from either the unmediated or 2 felt board conditions.

Errors on Trials 2,3, and 4 in the video condition were predominantly perseverative (53% of errors) as were errors on Trials 2, 3, and 4 in the felt board in a box condition (55 % of errors). None of the errors in the unmediated condition were perseverative.

Discussion

Experiment 1 tested the hypothesis taken from Schmitt and Anderson (2002) that the poor performance of 2-year-olds on the standard video task is due to their difficulties in mapping 2D space with egocentric coordinates onto a much larger 3D space requiring allocentric coordinates. Experiment 1 used a felt board as the hiding space so that the 2D video image closely matched the 2D search space of the felt board, thus reducing if not eliminating mapping problems. Consequently, the theory predicts that differences between the mediated (video) and unmediated condition would be minimal. Performance in the video condition and felt board in a box condition was poor, whereas performance in the unmediated condition was good; therefore, the prediction was not confirmed.

Performance was above chance on the first trial in the video condition and deteriorated to chance levels on subsequent trials, primarily due to perseveration errors. This result adds to similar observations from earlier studies (Schmitt & Anderson, 2002; Suddendorf, 2003).

EXPERIMENT 2

In Experiment 1, 2-year-olds did not successfully retrieve a hidden object based on information from video, even when the perceptual differences between TV and reality were substantially reduced. In Experiment 2, the children were verbally told

where to search for a toy hidden in an adjacent room, so there was less need for visual interpretation. In the unmediated condition, an experimenter standing directly in front of the child told the child where the toy was hidden. In the mediated video condition, the experimenter told the child via closed circuit video. If 2-year-olds' difficulty using information from video is primarily visual and perceptual in nature, as Schmitt and Anderson (2002) hypothesized, both groups would be expected to succeed on this task, insofar as DeLoache and Burns (1994) had previously shown that 2-year-olds were successful at object retrieval if they were simply told where the object was hidden.

There are two other theories that suggest predictions to be tested by this experiment. It is possible that 2-year-olds have difficulty using video to guide their search because they perceive the images as being real objects residing inside the TV cabinet (Lyle & Hoffman, 1972; Nikken & Peters, 1988). In this case, they may have problems with dual representations, as described by DeLoache (1987) in her classic research with scale models of the search space. In other words, the children may view the video objects as things in and of themselves and not be able to simultaneously view them as iconic representations of the real room. If their problem is of this nature, then providing verbal information about each object's location should eliminate the problem with video. The children should do as well as they do when the verbal message is unmediated. If, on the other hand, their difficulty is due to their tendency to discount information gained from video as relevant to reality, as Troseth and DeLoache (1998) hypothesized, performance in the video condition should be poor.

Participants

Thirty-two 2-year-olds (Range = 23 to 25 months), recruited from state birth records, participated (16 girls). Half were randomly assigned to the unmediated condition and half to the video condition. Parents were sent a letter explaining the study and then later contacted by telephone. Four additional children were dropped from the data analysis because they refused to complete the task.

Materials

Three adjoining rooms were required for testing. There was a hiding room where the experimenter hid the toy, a video room where she recorded the verbal cue live on camera, and an information room, where the child received the verbal cue, either directly from the experimenter or on the TV set. The hiding room (3.5 m × 2.5 m × 2.3 m) contained a chair, a table, a pillow, and a box. These 4 pieces of furniture functioned as the hiding places. The hiding object was a plush Snoopy dog (22 cm high).

A video camera was mounted on a tripod in the video room. It was connected to a television monitor (22 cm × 29 cm) on a rolling cart in the information room. In the information room, the experimenter told the child in the unmediated group where the toy was hidden. In the video group the child watched in the information room via closed circuit video as the experimenter in the video room told them where the toy was hidden.

A window, covered by black poster board, separated the hiding room and the information room. Another stationary video camera was located in the information room. A small hole was cut in the poster board, and the camera was pointed through it, toward the hiding room. This camera recorded the children's behavior on the test trials.

Procedure

Before participating in the test trials, each child first interacted with the assistant. After this brief warm-up, there was also retrieval and placement training. At the beginning of training (which took place in the hiding room), the experimenter showed the child the Snoopy doll and the four hiding places in the hiding room. A verbal label was provided for Snoopy and each of the hiding places. During retrieval training, the experimenter hid the Snoopy doll in each of the hiding places and asked the child to immediately retrieve it. During placement training, the child was told to place the Snoopy doll on each of the four hiding places.

After training was completed, there were four test trials. During the test trials, the experimenter either told the child in person, while standing in the information room, about the location of the Snoopy doll in the hiding room (unmediated group), or she told the child, on closed circuit video, about the location of the doll in the hiding room (video group). On video the experimenter was shown from the waist up and looking directly into the camera, giving the impression of making eye contact. The order of the four hiding places was balanced, with the constraint that each of the possible hiding places was first for 2 boys and 2 girls from each condition. The order of the hiding places for the last three trials was randomly selected without replacement from the remaining six possible orders.

An identical script, which verbally conveyed where the doll was hidden, was followed in both conditions. The experimenter initially told the child, "I'm going to hide Snoopy in his room and then come back and tell you where to find him. Then, you can go find Snoopy! I'll be right back!" While the experimenter was hiding Snoopy, the assistant said, "Right now, [experimenter] is hiding Snoopy. I wonder where she's hiding him." The experimenter then either returned to the information room (in the live unmediated condition) or went into the video room and stood in front of the video camera (video condition). In both conditions, she said, "I put Snoopy [in location]. Snoopy's hiding [in location]. Can you find him? Remember,

he's [in location]." In both conditions, after a delay of 3 sec, the child was allowed to search in the hiding room.

RESULTS

The main dependent variable was the percentage of errorless retrievals. Children in the video group (20% errorless retrievals, $SD = 26\%$) did not successfully retrieve the toy as compared to children in the unmediated group (64% errorless retrievals, $SD = 34\%$). A 2 (condition: video vs. unmediated) × 2 (gender) ANOVA revealed a conditions main effect, $F(1, 31) = 17.26$, mean squared error = 15312.5, $p < .001$. In the unmediated condition, performance over the four trials combined was significantly better than chance (25 %), $t(15) = 7.48$, $p < .001$ whereas in the video condition performance over the four trials combined was significantly worse than chance, $t(15) = 3.07$, $p < .01$.

Binomial tests using the mean proportion of errorless retrievals on Trials 2 through 4 (for each condition) as the null hypothesis probability of a correct response were used to determine whether Trial 1 performance was superior to that of later trials. The first trial performance (38% correct) in the video condition was significantly different from the mean percentage of errorless retrievals on Trials 2 through 4 (15 %, $p < .05$). Performance was not significantly above chance on any of the individual trials (38%, 13%, 25%, 6%, respectively). In the unmediated (real person talking) condition, first trial performance (60%) was not significantly different from the mean percentage of errorless retrievals on Trials 2 through 4 (62%). Trial 1 performance (62%) was significantly above chance ($p < .01$), as was Trial 2 (63%), Trial 3 (73%), and Trial 4 (56 %).

Errors on Trials 2, 3, and 4 in the video condition were mostly perseverative (78 %). Fifty-nine percent of errors on Trials 2 through 4 in the unmediated condition were perseverative.

GENERAL DISCUSSION

Two experiments tested the theory (Schmitt & Anderson, 2002) that 2-year-olds have perceptual difficulties mapping a 2D video image onto a 3D space in an adjacent room. The theory predicts that these difficulties would disappear in search tasks that do not require 2D to 3D mapping. Experiment 1 utilized a search space that was 2D in nature (a sticker hidden on a felt board). In contrast to the prediction, performance in the video condition was poor, whereas performance in the unmediated condition was good. Performance was also poor in a comparison condition not involving video. In this condition, the child was shown where the target

object was hidden on a separate felt board inside a box with a transparent window. We comment on the significance of this finding later.

In Experiment 2, the information was presented verbally, bypassing the issue of visual perception of the search space as depicted on video. Again, performance in the video condition was poor compared to the unmediated condition. This latter experiment also shows that the deficit in object retrieval produced by televised mediation is not limited to the visual aspect of television; it applies to the auditory aspect as well. Taken together, the experiments provide results that are inconsistent with predictions based on the Schmitt and Anderson (2002) theory.

Although these experiments eliminate the Schmitt and Anderson (2002) theory, the question nevertheless remains: why do 2-year-olds do poorly when they receive information about the location of a hidden object via video? Two other theories have been advanced: dual representation and the theory that 2-year-olds do not realize that video can inform them about reality. Each of these theories has problems accounting for the data as we note in the following. We suggest an additional theory. We propose that children do not succeed on Trials 2, 3, and 4 in video conditions because they lack true representational insight and prioritize their direct experience in the room over mediated information.

The Dual Representation Hypothesis

In Experiment 1, 2-year-olds were able, on the first trial, to find a sticker on a felt board based on visual information from video. The dual representation hypothesis (DeLoache, 1989; 2000) can explain this good first trial performance. Iconic representations—e.g., pictures or scale models—possess a dual reality, in that they are tangible things, as well as symbolic representations (DeLoache, 2000). For a child to use an iconic representation as a source of information to guide behavior, he or she must represent it not only as a real object, but also as part of an abstract relation between itself and what it stands for. This requires representational flexibility, or "the ability to represent, to think about, one and the same thing in two different ways" (DeLoache, 1989, p.32). In tasks where toddlers are asked to retrieve objects based on information from symbols, the dual representation hypothesis has been supported by their poor performance.

Although all external representations demand dual representation (if they are to be understood as symbols), DeLoache (2000) argued that 2D representations—e.g., pictures or video images—are less salient than concrete 3D objects; therefore, dual representation is easier to achieve with them. In support of this, 2.5-year-olds are better able to access information from a 2D video image or photograph than a 3D scale model (known as the picture superiority effect, DeLoache, 2000). In the video condition of Experiment 1, 2-year-olds responded correctly 53% of the time, or above chance, on their first trial.

Superior first trial performance on video or picture tasks was also found in two experiments in Schmitt and Anderson (2002), by Suddendorf (2003), and in the data collected by Troseth and DeLoache (1998). To elaborate on the latter point, we obtained the data reported by Troseth and DeLoache and reanalyzed it for trial effects. Like the other studies, performance in the video condition was superior on the first trial as compared to the other 3 trials (66% correct retrieval on the first trial compared to 36% correct on Trials 2, 3, and 4, $p < .05$, binomial test). Trial 1 performance was significantly better than chance (66 % vs. a chance level of 17 %, $p < .01$, binomial test). In fact, 2-year-olds overall performance in the video condition was better than chance (44 % errorless retrievals). We suggest, in line with DeLoache's (1987) theory, that it is precisely the lack of salience of 2D images that makes it possible for 2-year-olds to succeed on Trial 1, as described. The children do not perceive the video image as a thing in and of itself, thereby facilitating use of the information it provides about the adjacent room. It should be noted, however, that this interpretation, nevertheless, contrasts with claims by earlier researchers that children perceive TV images as showing objects within the TV cabinet (Lyle & Hoffmann, 1972; Nikken & Peters, 1988). The contradictions remain to be resolved.

There are two additional points to be made about the first trial advantage. The first is that video first trial performance is still worse than unmediated performance (Schmitt & Anderson, 2002; Suddendorf, 2003). The second is that this performance is severely disrupted on subsequent trials by perseveration errors. Perseverative responding was common in Experiment 1 and Experiment 2.

Suddendorf (2003) tested the possibility that perseverative responding was the principal source of interference with performance on Trials 2 through 4 in the object retrieval task. He used photographs rather than video to depict the location of the object. Instead of searching in one space four times, each child searched one time in each of four different rooms. This procedure eliminates the possibility of perseverative errors. Performance was above chance (averaging 53% correct) on all four trials. This suggests that when perseveration is not possible, photos and video can guide search at above chance levels, although not at the high levels of performance found with unmediated direct views of the hiding events. When a perseverative error is possible, however, photos and video are extremely poor guides for search. Although some believe that perseveration reflects a difficulty inhibiting prior motor responses (Diamond, 1985; 1991), others have suggested that children perseverate because they can't inhibit a prior representation (Jacques, Zelazo, Kirkham, & Semcesen, 1999). Along with Schmitt and Anderson (2002) and Suddendorf (2003), we suggest that superior first trial performance following a video presentation may be due to proactive interference, or the child's inability to inhibit previous representations on Trials 2, 3, and 4. On every trial after the first, the child has to overcome the tendency to act on their representation of the object's location in the room from the previous trial (which has been reinforced by direct experience) to choose the correct location of the object.

Given the drop in performance after Trial 1, we can only assume that 2-year-olds do not have an enduring understanding of the relationship between information on video and reality. We suggest instead that on Trial 1 the child has formed only a fragile memory of the location of the hidden object, which she uses to guide her search. However, this memory is not the same as the child having either conscious knowledge of the object's location or representational insight. This interpretation could be testable in future research by manipulations designed to disrupt recent or short-term memory. Such manipulations should eliminate the first trial advantage, producing poor performance on all trials, even using Suddendorf's (2003) procedure that eliminates perseverative errors.

In our research, there was no first trial effect in the felt board in a box condition of Experiment 1. Rather, performance was poor on all four trials. This is consistent with DeLoache's (1987) dual representation hypothesis insofar as the felt board in the box appears to be tangibly real and salient. Thus, 2-year-olds may have formed an initial strong expectation that the sticker was to be found on the felt board in the box, not on the search felt board. This was evident in pilot testing of earlier versions of the task, where the felt board was not enclosed in a box. The children, convinced that the sticker was on the felt board on which the experimenter hid it, would try to find it there, even if it required them to move a physical barrier, such as a curtain or a large piece of Plexiglas. Two-year-olds represented the location of the sticker in the box as an accessible real space, whereas in the video condition the location of the sticker was not perceived as accessible and most likely not represented as such.

The standard dual representation hypothesis does not, however, explain why 2-year-olds in Experiment 2 failed to use the verbal information presented on video. The majority of children clearly understood the language used by the experimenter, and they were able to act upon that understanding as evidenced by the performance in the unmediated group, which heard the same script as the video group. It should be noted that Experiment 2 might have been more difficult for children overall because they are required to use at least two symbol systems simultaneously (language conveyed through video). The child watching the experimenter on video not only has to understand what is being said, but also must understand that the video provides an iconic representation of the experimenter. In this sense, TV as a representation of reality always provides an additional processing burden.

The Reality Hypothesis

Yet another theory is that 2-year-olds discount TV as a plausible source of information about reality (Troseth & DeLoache, 1998). This is based on their experience with television as a medium that in fact has no connection to their daily lives in the real world. Troseth and DeLoache tested this theory by placing a TV set flush behind the window to an adjacent room. They told the children looking at the

television set that it was, in fact, a window. In this case, the children's performance improved from 44% in the TV condition to 66% in the TV as window condition. Although this did not match the 100% performance in the actual window condition, the improvement is consistent with the reality theory. The reality theory is also consistent with the overall poor performance in the video conditions in Experiments 1 and 2.

There are two problems with the theory. First, it does not explain why older children have little difficulty with object retrieval tasks based on video (Schmitt & Anderson, 2002; Troseth & DeLoache, 1998). Three-year-olds have much more experience than 2-year-olds with television as a medium of fantasy. One possible explanation is that older children may have had greater opportunity to experience live, contingent TV that is directly related to reality (as with camcorders). Troseth (2003) has shown that when 2-year-olds are given extensive experience with closed-circuit video images of their own behavior, they succeed on the standard video task (77% errorless retrievals compared to 23% by controls). Even so, it is not clear that the superior performance on the standard video task by 30- and 36-month-olds is due to such experience with video. Questionnaires filled out by parents in the Schmitt and Anderson (2002) studies indicated that few parents owned video cameras, so it is unlikely that experience with contingent video accounted for their superior performance on the standard video task.

The second problem is that the reality theory does not explain the relatively superior first trial performance in the standard video task. After all, if 2-year-olds believe television is not connected to reality, it is just as disconnected from reality on the first trial as on later trials.

The Conflicting Representations Account

None of the extant theories can, of themselves, fully account for the findings of Experiments 1 and 2. It is clear that on Trial 1, in both experiments, some 2-year-olds are able to apply the information from TV to guide their search. After they find the object on the first trial, however, they now have a competing representation of the toy's location based on their experience in the room. What representation do they choose for Trial 2? They now use the representation of the location of the object based on their real experience in Trial 1. Thus, we suggest that 2-year-olds only use video as a source of information when it does not conflict with their real life experience. This preference for real life experience in making choices may even extend to adults when the choices are based on incomplete or uncertain information. Research with adults shows that information from media sources is less likely to be considered than information from real life sources when making political and other decisions (Bryant & Thompson, 2001; Jeffres, 1997; Klapper, 1960; Lazarsfeld, Berelson, & Gaudet, 1948).

We suggest that 2-year-olds are unable to update their representations of the location of the object because they perceive the new information from video as conflicting with information garnered from real life experience. Several studies have revealed that young children have great difficulty retrieving a hidden object when they have to choose from among conflicting alternatives. In Povinelli, Landau, and Perilloux (1996), 2-, 3-, and 4-year-olds had a sticker placed on their head during a game. The children did not know about the sticker. After several minutes, they watched a video of the game, showing the sticker being placed on their head. None of the 2-year-olds reached up to the sticker, compared to 25% of the 3-year-olds and 75% of the 4-year-olds. In a similar study using Polaroid photographs instead of video, equivalent results were found. However, when the children who didn't try to touch the sticker looked in the mirror, nearly all of them reached up to touch the sticker. According to Povinelli et al., the younger children had difficulty in their task because it was a "situation of conflicting representations, and they may simply defer to what they currently believe to be true" (Povinelli et al., 1996, p. 1553). Lacking true representational insight, the children acted on the representation based on their direct unmediated experience rather than on information from video or photographs. In contrast, 2- and 3-year-old children do appear to have representational insight about mirrors. Simcock, Suddendorf, and Neilsen (2004) found that 90% of 2-year-olds removed a sticker that had been surreptitiously placed on their head when they saw themselves in the mirror, whereas only 35% of 2-year-olds removed the sticker when they saw themselves on video. This was the case even though the video image appeared on a life-sized screen of high resolution and the image was "left-right symmetrically contingent," like a mirror image would be (Simcock et al., 2004, p. 1)

Zelazo, Somerville, and Nichols (1999) also tested whether 3-year-olds and 4-year-olds could use external representations that conflicted with what they expected. Half of the children first directly watched an experimenter hide an object in a room. After the hiding event, they were told the object would be in a different place, and were then either told or shown (on video) where the object was located. The other half of the children did not first see the object being hidden; they were either told or shown the location of the object, with no conflicting information provided. Three-year-olds in the first group, where the verbal or video information conflicted with the hiding event they had previously seen, often could not find the hidden object. Even though they were told or shown that the object would be in a different place, younger children did not choose the location that they had been told about or seen when it conflicted with their direct experience. In this case, 3-year-olds could not use information from video, whereas 3-year-olds in Schmitt and Anderson (2001) and Troseth and DeLoache (1998) could. We submit that the difference between these studies and the study by Zelazo et al. was whether or not the child participated in the removal of the toy. In the Zelazo et al. study, the child passively observed the toy being removed from its original location. Thus, we sug-

gest 3-year-olds need to have real life experience removing the toy from a location before they are able to use televised information to update their representation of the toy's location.

According to Zelazo et al. (1999), young children often cannot use a higher order rule to select from among multiple representations. Lacking this ability, they default to the representation that is most salient to them. In Experiments 1 and 2 here, we suggest children succeed on the first trial because they use their initial representation of the object's location to guide behavior. The direct experience of finding the object in the room strengthens that representation (presumably by reinforcing the neural connections underlying it). On the second trial, the stronger representation remains present and overcomes the new representation acquired from watching video. This is a low-level representation because the child still does not have symbolic insight about the function of video and thus does not have a rule or basis for preferring the most recent video event over his or her direct experience in the room when the two conflict.

The two test conditions in Experiment 1 provide further support for the notion that conflicting representations can inhibit search task performance. In Experiment 1, in the felt board in a box condition, the 2-year-olds had a strong expectation that the sticker was on the felt board in the box, which competed with their ability to form a symbolic representation of the correct location of the sticker on the target felt board. As in the video condition, perseverative responding in the box condition was very common on Trials 2 through 4. Two-year-olds in this condition had to inhibit representations formed from their most recent experience with the search felt board, as well as the strong expectation that the sticker was on the felt board in the box. In the video condition, the felt board on TV was not a solid object like the felt board in the box, so 2-year-olds did not have a strong expectation on Trial 1 that the sticker was on a felt board in the TV. Lacking this strong conflicting expectation on Trial 1, the majority of children in the video condition were able to choose the correct location on this trial only.

Why, then, can children only 6 months older succeed in the standard video task? We suggest that a combination of experience and maturation are responsible. Between 24 and 30 months, children acquire information about symbols through such experiences as picture book reading, language, symbolic play, and interactions with parents. Troseth (2003) suggests that as children's knowledge base about symbols grows, so does their ability to detect symbol-referent relations. If the experience is specific to the test symbol system, in this case video, experience with video cameras can improve children's understanding of how video and reality can be related (Troseth, 2003). Troseth's research indicates that the overall poor performance of the 24-month-olds on the standard video task can be overcome by training, so it does not appear due to a purely biologically-based maturational deficit. However, maturational developments are probably partly responsible for improvements on the task with age and in the absence of specific training. Research

on analogical reasoning has shown that children, as they age, are increasingly able to rely on deep structural relations between entities, as opposed to superficial relations. They can see past differences in the appearances of things to see how they are related on a more abstract level (Gentner & Medina, 1998, Troseth, 2003). Furthermore, inhibitory skills improve between the ages of 24 and 30 months (Ruff & Rothbart, 1996; Vaughn, Kopp, & Krakow, 1984), perhaps allowing children to inhibit responding based on salient representations in favor of abstract, higher-order rules. Prefrontal cortical development is the maturational factor most likely responsible for improved inhibition (Diamond & Gilbert, 1989), although language development may be implicated as well (Vaughn et al., 1984).

Whatever the reasons, our research adds to a small body of evidence showing that infants and toddlers are less likely to learn from, or at least act on, the basis of video than equivalent unmediated displays. For example, Barr and Hayne (1999) reported that infant imitation from video is inferior to imitation from a real display, especially if the actions involved two or more discrete steps. Hayne et al. (2003) showed that this video deficit extends up to 36 months.

These findings stand in interesting contrast to the recent great increase in video production for infants and toddlers. These videos and TV programs, such as *Baby Einstein* and *Teletubbies*, are marketed in such a way that parents believe their babies will engage in important learning from them. Given that infants and toddlers may be spending increasing amounts of time watching video, understanding the reasons for their learning difficulty, and understanding the impact of time spent watching video without associated learning, appears to be of considerable urgency.

ACKNOWLEDGMENTS

This research was supported in part by a training grant from the National Institutes of Health and from the Center on Media and Child Health at Children's Hospital Boston to Marie Evans Schmidt and by grant BCS0111811 from the National Science Foundation to Daniel R. Anderson. The National Science Foundation is not responsible for the findings and opinions reported in this article. The authors thank Georgene Troseth and Judy DeLoache for giving us access to the Troseth and DeLoache (1998) data as well as for their cogent comments.

Publication of article paper was delayed. Troseth, Saylor, and Archer (2006) have since published a paper based on Experiment 2. The authors replicated our results and conducted an additional experiment, in which an actor, via video, established a personal, interactive relationship with each participating child before telling them where the object was hidden; under these circumstances, 2-year-old children were able to use verbal information from video to find the object.

REFERENCES

Anderson, D. R., Lorch, E. P., Field, D., & Sanders, J. (1981). The effects of TV program comprehensibility on preschool children's visual attention to television. *Child Development*, *52*, 151–157.

Anderson, D. R., & Pempek, T. (2004). Television and very young children. *American Behavioral Scientist, 48*, 505–522.

Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development, 70*, 1067–1081.

Bryant, J., & Thompson, S. (2001). *Fundamentals of media effects.* New York: McGraw Hill.

Crawley, A. M. (2002). Two-year-olds' comprehension of television: Do they believe their eyes or their ears? Unpublished doctoral dissertation, University of Massachusetts, Amherst.

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

DeLoache, J. S. (1989). Young children's understanding of the correspondence between a scale model and a larger space. *Child Development, 62,* 121–129.

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

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

DeLoache, J. S., Kolstad, D. V., & Anderson, K. (1991). Physical similarity and young children's understanding of scale models. *Child Development, 62,* 111–126.

DeLoache, J. S., Peralta de Mendoza, O. A., & Anderson, K. (1999). Multiple factors in early symbol use. Instructions, similarity, and age in understanding a symbol-referent relation. *Cognitive Development, 14,* 299–312.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56*(4), 868–883.

Diamond, A. (1991). Frontal lobe involvement in cognitive changes during the first year of life. In K. R. Gibson & A. C. Petersen (Eds.), *Brain maturation and cognitive development: Comparative and cross-cultural perspectives* (pp. 127–180). Hawthorne, NY: Aldine de Gruyter.

Diamond, A., & Gilbert, J. (1989). Development as progressive inhibitory control of action: Retrieval of a contiguous object. *Cognitive Development, 4,* 223–249.

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

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65,* 263–297.

Grela, B., Lin, Y., & Krcmar, M. (2003, XXMONTH). *Can television be used to teach vocabulary to toddlers?* Presented at annual meeting of the American Speech Language Hearing Association, April, Chicago.

Hawkins, R. P. (1977). The dimensional structure of children's perceptions of television reality. *Communication Research, 4*, 299–320.

Hayne, H., Herbert, J., & Simcock, G. (2003). Imitation from television by 24- and 30-month olds. *Developmental Science, 6*(3), 254–261.

Huston, A., & Wright, J. (1997). Mass media and children's development. In W. Damon, I. E. Sigel, & K. A. Renninger (Eds.), *Handbook of child psychology, Volume 4: Child psychology in practice* (pp. 999–1058). New York: Wiley.

Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology, 35,* 770–780.

Jaglom, L. M. & Gardner, H. (1981). The preschool television viewer as anthropologist. In H. Kelly & H. Gardner (Vol. Eds.), *New directions in child development: No. 13, Viewing children through television* (pp. 9–30). San Francisco: Jossey-Bass.

Jeffres, L. (1997). *Mass media effects* (2nd Ed.). Prospect Heights, IL: Waveland Press.

Klapper, J. T. (1960). *The effects of mass communication.* New York: Free Press.

Krakow, J. B., Kopp, C. B., & Vaughn, B. E. (1982). *Sustained attention during the second year: Age trends, individual differences, and implications for development*. Unpublished manuscript.

Lazarsfeld, P. F., Berelson, B. R., & Gaudet, H. (1948). *The people's choice*. New York: Columbia University Press.

Lyle, J., & Hoffman, H. R. (1972). Children's use of television and other media. In E. A. Rubinstein, G. A. Comstock, & J. P. Murray (Eds.), *Television and social behavior: Television in day-to-day life. Patterns of use* (DHEW Publication No. HSM–72–9059, Vol. 4, pp.129–256). Washington, DC: U.S. Government Printing Office.

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.

Munakata, Y., McClelland, J. L., Johnson, M., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104*(4), 686–713.

Nikken, P., & Peters, A. L. (1988). Children's perceptions of television reality. *Journal of Broadcasting and Electronic Media, 32*, 441–452.

Peralta de Mendoza, O. A., & Salsa, A. M. (2003). Instruction in early comprehension and use of a symbol-referent relation. *Cognitive Development, 18,* 269–284.

Povinelli, D. J., Landau, K. R., & Perilloux, H. K. (1996). Self-recognition in young children using delayed versus live feedback: Evidence for developmental asynchrony. *Child Development, 67,* 1540–1554.

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development*. New York: Oxford University Press.

Schmitt, K. L. (1995). *Infant's visual attention to form and content features of television*. Unpublished Masters thesis, University of Massachusetts, Amherst.

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

Schmitt, K. L., Anderson, D. R., & Collins, P. A. (1999). Form and content: Looking at visual features of television. *Developmental Psychology, 35*, 1156–1167.

Simcock, G., Suddendorf, T., & Nielsen, M. (2004, April). *Developmental differences in the emergence of mirror vs. live video self-recognition*. Poster presented at the biannual meeting of the International Conference on Infant Studies, Chicago, IL.

Suddendorf, T. (2003). Early representational insight: Twenty-four-month olds can use a photo to find an object in the world. *Child Development, 74*(3), 896–904.

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

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., 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.

Vaughn, B. E., Kopp, C. B., & Krakow, J. B. (1984). The emergence and consolidation of self control from eighteen to thirty months of age: Normative trends and individual differences. *Developmental Psychology, 55,* 990–1004.

Wright, J. C., Huston, A. C., Reitz, A. L., & Piemyat, S. (1994). Young children's perceptions of television reality: Determinants and developmental differences. *Developmental Psychology, 30,* 229–329.

Zelazo, P. D., Sommerville, J. A., & Nichols, S. (1999). Age-related changes in children's use of external representations. *Developmental Psychology, 35*(4), 1059–1071,