Age-Related Changes in Children’s Use of External Representations

Philip David Zelazo, Jessica A. Sommerville, and Shana Nichols
University of Toronto

This study explored children’s use of external representations. Experiment 1 focused on representations of self: Self-recognition was assessed by a mark test as a function of age (3 vs. 4 years), delay (5 vs. 3 min), and media (photographs vs. drawings). Four-year-olds outperformed 3-year-olds; children performed better with photographs than drawings; and there was no effect of delay. In Experiment 2, 3- and 4-year-olds used a delayed video image to locate a sticker on themselves (self task) or a stuffed animal (other task). The 2 tasks were positively correlated with age and vocabulary parted aloud. Experiment 3 used a search task to assess whether children have particular difficulty using external representations that conflict with their expectations: 3- and 4-year-olds were informed of an object’s location verbally or through video; on half of the trials, this information conflicted with children’s initial belief. Three-year-olds performed worse than 4-year-olds on conflict trials, indicating that assessments of self and other understanding may reflect children’s ability to reason about conflicting external representations.

Assessments of cognitive development routinely require children to understand and use external representations, such as mirror images, drawings, and speech. However, children’s use of external representations is an important topic in its own right (e.g., Bialystok, 1999; DeLoache, 1991; Liben, 1999), and difficulty using external representations may contribute to poor performance on tests that were designed to assess other aspects of children’s cognition, such as their self-concept or their understanding of belief. As Troseth and DeLoache (1998) put it, “The medium can obscure the message” (p. 950). In the present study, we first used external representations to assess children’s self-concept and then explored the extent to which children’s performance on our tests reflected their developing ability to use external representations.

External Representations of Self

Reaction to a visual image of the self has been a popular measure of children’s developing self-concept since Gallup (1970) and Amsterdam (1972) independently introduced the mark test in their work with chimpanzees and children, respectively. In Amsterdam’s (1972) investigation, mothers surreptitiously applied rouge to their infants’ noses and then turned their infants to face a mirror. Mark-directed behavior (e.g., locating the rouge on one’s nose) appeared to emerge during the second half of the 2nd year, arguably reflecting the development of an objective self-concept (e.g., Lewis & Brooks-Gunn, 1979) and the understanding that what is true of the mirror image is also true of oneself (Povinelli, Landau, & Perilloux, 1996). However, Povinelli et al. (1996) revealed limitations of young children’s visual self-recognition in experiments that used a mark test with delayed video, live video, photographs, and mirrors. In one experiment (Experiment 1), 2-, 3-, and 4-year-olds played a game during which an experimenter surreptitiously placed a sticker on the children’s heads. About 3 min later, children were presented with a video image of the marking event. Whereas the majority of older children (75% of 4-year-olds) reached up to touch the sticker, few of the younger children (25% of 3-year-olds and none of the 2-year-olds) did so. Similar results were found in another experiment (Experiment 2) using the same basic delayed self-recognition paradigm but with Polaroid photographs instead of video images. Moreover, in this experiment, nearly all of the oldest children (53- to 58-month-olds) used the personal pronoun me to label their image, but a substantial percentage of the youngest children (35- to 40-month-olds) labeled the image using their proper name.

Povinelli et al. (1996) interpreted young 3-year-olds’ poor performance with delayed-media representations compared with live representations (such as those provided by mirrors) as evidence that children at this age have a temporally restricted sense of self. Younger children’s use of a proper name to label their own image was taken to reveal a sense of dissociation from the image, as if the children had failed to identify with it. On the basis of their findings, Povinelli et al. proposed that “the asynchronous development of self-recognition in live versus delayed conditions is the result of a developmental lag between the emergence of an ‘online’ self-concept and a self-concept that includes an understanding of the temporal continuity of the self” (p. 1552). On this account, young 3-year-olds possess a succession of present-oriented representations of self (termed present selves; Povinelli, 1995, p. 165), but they fail to integrate these representations into a single unifying structure, which Povinelli referred to as the
proper self. According to Povinelli (1995) and Povinelli et al. (1996), the construction of a proper self depends on domain-general changes in children's capacity for representations (e.g., Olson, 1993; Perner, 1991).

If Povinelli et al.'s (1996) results indicate that young children's self-concept is temporarily restricted, then several important questions arise (see Povinelli, 1995). First, what is the temporal span of the present self? Although Povinelli (1995) hypothesized that the present self is "very narrow" (p. 185), he suggested that it is not limited to the fleeting here-and-now (cf. James, 1895, p. 111). Rather, it must "recognize a time corridor into the immediate past and into the immediate future" (Povinelli, 1995, p. 166). Because temporal span may be an important dimension along which children's self-awareness develops (e.g., Zelazo, 1996), we attempted in Experiment 1 to gauge the span of young children's sense of self using two levels of delay (5 s vs. 3 min) between the media-recording event and the presentation of the self-image in the delayed self-recognition paradigm. In contrast to the 3-min delay used by Povinelli et al. (1996), 5 s might be expected to fall within the span of the present self.

Another important question concerns the generality of any change in children's self-concept. Povinelli et al. (1996) suggested that the development of a temporally extended self-concept depends on the same domain-general cognitive changes that underlie the emergence of a "theory of mind," as indicated, for example, by an understanding of false belief (e.g., Perner, Leekham, & Wimmer, 1987), representational change (Gopnik & Astington, 1988), and appearance reality (Flavell, Green, & Flavell, 1986). Indeed, it seems reasonable that both a temporally extended self-concept and a theory of mind may depend on the ability to integrate conflicting perspectives into a single, hierarchical system of inferences (Frye, Zelazo, & Palfai, 1995; Zelazo & Frye, 1997). Whereas a proper self may require children to integrate past and present selves into a single system, an understanding of false belief, for example, may require children to integrate their own, correct perspective and someone else's mistaken one. If delayed self-recognition is a valid measure of children's developing self-concept, and the emergence of a temporally extended self depends on the same domain-general cognitive changes as theory of mind, then individual differences in delayed self-recognition ought to be related to individual differences in theory of mind even with other sources of variation (e.g., age and vocabulary scores) partialed out. Experiment 1 was also designed to test this hypothesis.

A third question concerns the effect of different media on young children's use of delayed representations of self. An effect of media over and above any effect of delay would indicate that young children's poor performance on tests of delayed self-recognition may be due in part to difficulty reasoning about particular types of external representation (e.g., DeLoache, 1991; Liben, 1999; Troseth & DeLoache, 1998). As noted earlier, difficulty using external representations to guide behavior may mask a more mature understanding of self. In light of this masking hypothesis (cf. Mitchell, 1994), Experiment 1 compared children's performance on photographs with their performance on drawings. Moreover, this experiment included a training phase that was designed to familiarize children with the basic properties of the two types of media.

In summary, the primary aims of Experiment 1 were (a) to assess the effect of delay (5 s vs. 3 min) on children's performance on a test of delayed self-recognition, (b) to determine whether delayed self-recognition is related to performance on theory-of-mind tasks, and (c) to compare performance on two different types of media (photographs and drawings). Consistent with the masking hypothesis, Experiment 1 revealed an effect of media type. Moreover, we were surprised by 3-year-olds' poor performance on the training phase, which we found to be correlated with their performance on the mark tests. Together these findings support the suggestion that poor performance on the mark tests can be attributed to difficulty with the method of assessment rather than to a temporally restricted sense of self. Experiment 2 demonstrated that 3-year-olds also performed poorly when required to use delayed representations to locate a sticker on someone else (a stuffed animal) and that performance on tasks involving self and other was highly correlated. Finally, Experiment 3 used a variant of Troseth and DeLoache's (1998) hidden-object search task to try to determine more precisely the circumstances in which 3-year-olds have difficulty using external representations to guide their behavior. It was found that children at this age had particular difficulty using both videotaped and verbal information when this information conflicted with their initial expectations.

Experiment 1

In Experiment 1, children at two ages (3 and 4 years) received a training task, two theory-of-mind tasks, and two mark tests. The ages correspond to the period of most rapid change in performance on delayed self-recognition (Povinelli et al., 1996). The mark tests were each presented in the context of a different game. A vocabulary test was administered as a discriminant measure. On the mark tests, children were presented with either photographs or drawings of themselves with a mark, and these representations were presented after a 5-s or a 3-min delay. The primary dependent measure was when in a series of increasingly supportive prompts children reached up to touch the mark. A secondary measure was how children labeled their image. Two theory-of-mind tasks and two mark tests were administered to provide a relatively reliable assessment of performance on each task. We were also interested in assessing practice effects on the mark test, which would be more plausibly interpreted in terms of learning about the method of assessment (i.e., use of external representations) than in terms of procedurally induced changes in the temporal extension of children's self-concept.

Method

Participants

A total of 61 children, including 29 three-year-olds and 32 four-year-olds, were recruited through local day-care centers. The majority of children were Caucasian and middle class, although data about race and socioeconomic status were not systematically collected. Parents provided informed consent for their children's participation. Eight of the children were dropped from the final sample for the following reasons: (a) They refused to cooperate and did not finish the procedure (n = 5); (b) they did not understand English sufficiently to complete the task (n = 1), and (c) the mark was not visible in the photograph (n = 2; see below). Of the remaining 53 children, 13 were dropped from the final sample because they reached up to the mark on one of the mark tests prior to media presentation (see Results). As a result, the final sample comprised 20 three-year-olds (mean age = 36.2 months; range = 29.1 to 42.1) and 20 four-year-olds.
Materials

The mark tests were presented in the context of games. One of the marking games (Duck, Duck, Goose) involved two stuffed animals (a beaver and a dog) and large, bright-pink stickers (9 × 8 cm). The other game (Animal Masks) required the use of five animal masks (pig, dog, duck, rabbit, and tiger noses) with bright-pink hypoallergenic makeup applied to the inside of one of the masks (the tiger nose). In the photograph condition, Polaroid photographs were taken of each child’s head from a distance of about 30 cm; the sticker or mark was clearly visible in the image (except for two cases, as noted in Participants). In the drawing condition, prepared silhouettes were used as a point of departure for the drawings. These silhouettes included a generic outline of a child’s head and shoulders, generic facial features, and a pink sticker or mark. The primary experimenter used colored pencils to fill in each child’s actual hair, eye, lip, and skin color and to add the color and pattern of the child’s shirt and jewelry (if any). For each child, the secondary experimenter took one real media recording (real image) and pretended to take another (sham image). Two recordings were necessary to be able to provide children in the photograph condition with media feedback after a 5-s delay, which was faster than a photograph could develop. In all conditions, the real image was surreptitiously substituted for the sham image. One theory-of-mind task involved a crayon box with a doll inside; the other involved a milk carton with rubber bands inside.

Design

Children were randomly assigned to one of four conditions, created by crossing type of medium (photographs vs. drawings) with temporal delay (5 s vs. 3 min). All of the children received two training tasks, two mark tests, two theory-of-mind tasks, and the Peabody Picture Vocabulary Test—Revised (PPVT-R, Form L; Dunn & Dunn, 1981). Half of the children received the tasks in the following order: the training tasks, one mark test (involving Duck, Duck, Goose), one theory-of-mind task (involving the crayon box), the PPVT-R, the other theory-of-mind task (involving Animal Masks). For the remaining children, the order of the two mark tests was reversed. Children in the photograph condition received the training tasks with a photograph and were tested using photographs in both mark tests. Children in the drawing condition saw drawings during both training and test.

Procedure

All of the children were tested individually by two female experimenters: a primary experimenter who administered the procedure and a secondary experimenter who generated the media recordings. Unless otherwise specified, the primary experimenter spoke to the children and the secondary experimenter interacted with the stuffed animals. As soon as children were comfortable with the setting, they received the training tasks.

Training tasks. The procedure for the training tasks is illustrated using the photograph condition. (The drawing condition was identical to the photograph condition except that a drawing was used.) Children were told, “Nancy [i.e., the secondary experimenter] really likes to take photographs.” Children were then shown one of the stuffed animals (Chester the Beaver) and told, “This is Chester. Chester really likes to have his photograph taken.” The stuffed animal had a sticker on its back that children could not see because of their position in relation to the animal. At this point, the secondary experimenter told children, “What’s that, Chester? You want me to take your photograph? Well, OK.” The primary experimenter explained, “When Nancy takes a photograph she always makes photographs of things exactly the way they are. So when Nancy takes a photograph of something it looks exactly like that thing.” The secondary experimenter then said, “What’s that, Chester? You want a photograph of your back? OK,” and the secondary experimenter moved to the opposite side of the table. As the secondary experimenter pretended to take the photograph, children were told, “Here we go. 1 . . . 2 . . . 3.” After pretending to take the photograph, the secondary experimenter told children, “Wow, this photograph is taking a long time to develop. It must be a slow photograph. We’ll wait while it develops.” A previously taken photograph of the stuffed animal was surreptitiously substituted for the new one and presented to children after a 15-s delay. A previously taken photograph was used to standardize the stimulus events. Children were allowed to look at this photograph for 15 s before they were asked, “What is that?” while the primary experimenter pointed to the sticker in the photograph. Children who did not provide the correct answer were told, “It’s a sticker.” All children were then asked to find the sticker on the stuffed animal (“Can you find the sticker for me?”) and were allowed to search for it. Children who did not find the sticker in 15 s were given the following prompt, “In the photograph, there is a sticker on Chester’s back. Remember, the photograph is exactly the same as the real thing. Can you find the sticker for me?” Children were given an additional 15 s to find the sticker. After 15 s the primary experimenter took children who had not found the sticker to the opposite side of the table. The primary experimenter then juxtaposed the photograph and the stuffed animal and said, “Look, there’s a sticker on Chester’s back right now and there’s a sticker on Chester’s back in the photograph. Remember, Nancy always takes photographs of things exactly as they are.”

The second training task proceeded in exactly the same manner as the first training task, except that on this task children were required to search for a mark on Groucho the Dog’s side (the side they could not see) instead of Chester the Beaver’s back.

In light of the masking hypothesis, we decided to score performance on the training tasks to assess whether it was related to subsequent performance on the mark tests. Performance on each training task was scored out of 2. Children received 2 points for locating the sticker or mark when they were first asked to find it. They received a score of 1 if they located the sticker or mark after being provided with a verbal prompt. Otherwise, they received a score of 0.

Mark tests. The procedure for the mark tests is illustrated using the photograph condition. (The drawing condition was identical to the photograph condition except that a drawing was used.) After the training tasks, the primary experimenter introduced the game that would be used in the first mark test and told children that they would be photographed or drawn (depending on their condition) during the game. In the Duck, Duck, Goose game, children were told:

Duck, Duck, Goose is a really fun game. In this game, everyone will sit in a circle with their eyes closed, except for me because I’m going to be the one who gets chased. What I’ll do is walk around the circle and pat everyone on the head. When you hear the word “goose” you have to chase me around the circle. You’d better be quick because I’m going to try to steal your spot! Are you ready?

The children were also told about the media recording that would occur: “You know what? When you’re playing the game, Nancy will take a photograph of you. Remember, when Nancy takes a photograph of something it looks just like that thing. Later we’ll look at the photograph.” Children were then instructed to sit in a circle with a stuffed animal and the secondary experimenter. On each trial, the primary experimenter made two full circles before assigning the label goose to the child. (Children were always labeled the goose and were allowed to be the experimenter back to the spot.)

The game was played for five trials. The first trial was designed to habituate children to the feeling of being patted on the head. On the second rotation of the second trial, a large sticker was surreptitiously placed on top
of children's heads by the primary experimenter. After the completion of the second trial, the secondary experimenter took a media recording of children (real image) and told them, "This is a slow photograph. We'll take a faster one later." The final three trials served to assess whether children spontaneously discovered the sticker, prior to viewing the photograph or drawing. Children who reached within one inch of the sticker during the three trials (n = 13, see Results) were excluded from the final sample of participants. After the fifth trial, the secondary experimenter pretended to take another media recording (sham image). The real image was surreptitiously substituted for the sham image. The 5-s delay condition, children were immediately shown the real image. In the 3-min delay condition, the experimenter showed children a Where Is Waldo? book during the delay and then showed children the real image.

The primary experimenter introduced the media presentation by telling children, "Look. This picture was taken right after we played the game." For the first 15 s of the media presentation, the experimenter did not say anything. After this initial period, the experimenter pointed to the image and initiated a graded series of increasingly supportive prompts. Children were asked, "Who is that?" (Prompt A) at 5-s intervals until they provided a response (this prompt was repeated up to a maximum of three times). Verbal responses to this question were recorded for subsequent coding and analysis (see Results). Children were then asked up to four more questions, each of which was asked only once. First, they were asked, "What is that?" (Prompt B) while the experimenter pointed to the image of the sticker in the media recording. If children did not provide the correct response, they were told that the object in the image was a sticker. Children were then asked, "Where is the sticker?" (Prompt C), followed by, "Where is the sticker on you?" (Prompt D). Finally, children were asked, "Can you find the sticker for me?" (Prompt E).

Ordinal scores on the mark test were assigned according to the point at which children touched the sticker on their heads during the media presentation. (No child who reached within an inch of the sticker failed to touch it.) Children who touched the sticker during the first 15 s were assigned a full score of 5. Children received a score of 4 for reaching up after Prompts A or B; 3 for reaching up after Prompt C; 2 for reaching up after Prompt D; and 1 for reaching up after Prompt E. Children who failed to reach up at all were assigned a score of 0. Children who failed to reach up to the sticker in response to any of the prompts were shown the sticker at the end of the trial. This was done for two reasons. First, the experimenter had to remove the sticker from the children's hair at the end of the trial regardless of whether it was found or not, and it would have been difficult to do this surreptitiously. Second, if only some children became aware of the presence of the sticker on the first mark test, this might have affected differentially their performance on the second mark test.

The procedure for the Animal Masks game was identical to that for Duck, Duck, Goose, except that the marking event and media recording occurred in the context of a different game. In the Animal Masks game, children were presented with a series of animal noses. On each trial, children were presented with an animal nose and told, "Let's look at this nose. What kind of animal is this?" The primary experimenter then helped children to put on the nose. The primary experimenter asked, "What sort of noise does a _____ [pig, duck, rabbit, tiger, or dog] make? Can you pretend to be a _____ for me?" Children were then told, "Let's try another animal nose," and the current animal nose was removed from their faces. Children were presented with a total of five animal noses, the second of which was marked with makeup so that when children removed the mask a spot of bright-pink makeup was left on their noses. Performance was scored exactly as in Duck, Duck, Goose. No child who reached within an inch of the makeup failed to find it.

Theory-of-mind tasks. In the first theory-of-mind task, the primary experimenter showed children the crayon box and asked, "What's in the box?" If children did not provide a response to this question or answered, "I don't know," they were told that it was a crayon box. The box was then opened and children were shown that it actually contained a doll. The box was closed again and children were asked three pairs of questions in a fixed order. These pairs of questions assessed representational change (Gopnik & Astington, 1988), false belief (Perner et al., 1987), and appearance reality (Flavell et al., 1986). Children were asked (1a) "What did you think was in the box before I opened it?" and (1b) "What's in the box, really?"; (2a) "If I showed one of your friends the box, what would they think was in it?" and (2b) "What's in the box, really?"; and (3a) "What does it look like the box has in it?" and (3b) "What's in the box, really?" A fixed order of theory-of-mind items was used because theory-of-mind performance was an individual difference variable and we were not interested in comparing performance on specific theory-of-mind items. The second theory-of-mind task involved a milk carton that actually contained rubber bands. This task was carried out in the same manner as the first theory-of-mind task.

Vocabulary. The PPVT-R (Dunn & Dunn, 1981) was administered in the standard fashion.

Results

Eight children (2 three-year-olds and 6 four-year-olds) reached up to the sticker in the Duck, Duck, Goose game prior to the media presentation, and 5 children (2 three-year-olds and 3 four-year-olds) reached up to the mark in the Animal Masks game. These children were excluded from subsequent analyses.

Training Tasks

Performance on each of the two training tasks was scored out of 2. A 2 (age group) × 2 (media) × 2 (trial) mixed analysis of variance (ANOVA) on training task scores revealed that 4-year-olds (M = 1.0, SD = 0.89) performed significantly better than 3-year-olds (M = 0.40, SD = 0.70), F(1, 36) = 6.10, p < .05; children who received photographs (M = 0.98, SD = 0.90) performed significantly better than children who received drawings (M = 0.43, SD = 0.71), F(1, 36) = 5.12, p < .05; and children's performance on the second training trial (M = 0.88, SD = 0.99) was better than their performance on the first training trial (M = 0.53, SD = 0.82), F(1, 36) = 10.65, p < .005. There were no significant interactions (all Fs < 1.0).

Examination of the responses made by children who received a score of 0 on a given training task indicated that the large majority of them (71%, or 35 out of the 49 cases that could be scored from videotape) pointed repeatedly to the photograph or drawing when asked, "Can you find the sticker for me?"

Mark Tests

Mark-directed behavior. Scores on the mark tests ranged from 0 to 5. An initial inspection of the data revealed that the full range of scores was used. The numbers of children who received scores of 0 through 5 on the first mark test were 13, 2, 13, 3, 5, and 4, respectively. For the second mark test, the corresponding numbers of children were 13, 1, 4, 5, 5, and 12. Recall that children were required to touch the mark during the first 15 s to receive a
Table 1

<table>
<thead>
<tr>
<th>Age group</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-year-olds</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>20</td>
<td>17</td>
</tr>
</tbody>
</table>

Number and Percentage of 3- and 4-Year-Olds Who Labeled Their Image "Me" in Each Mark Test in Experiment 1

Discussion

The results of Experiment 1 confirmed previous reports (Povinelli et al., 1996; Povinelli & Simon, 1998) that 3-year-olds have difficulty on mark tests of delayed self-recognition. Across a variety of conditions, and on each of two mark tests, most 3-year-olds failed to use information presented through delayed external representations to locate a mark on their head or their nose. The difficulty of the mark tests used in Experiment 1 is underscored by the finding that even most 4-year-olds failed the first mark test that
they received. Indeed, the percentage of 4-year-olds passing the first mark test in this experiment was somewhat lower than the percentage found in a corresponding condition in a previous study (Povinelli et al., 1996, Experiment 2). Nonetheless, most 4-year-olds passed the second mark test that they received, and overall, 4-year-olds performed significantly better than 3-year-olds. In addition, 4-year-olds were more likely than 3-year-olds to use a personal pronoun than their proper name to label their image on the first mark test that they received, consistent with findings from Povinelli et al. (1996). However, an age difference in labeling was not found for the second mark test, and individual differences in labeling were not significantly related to individual differences in mark test performance. Four-year-olds’ use of a personal pronoun to label their image might suggest that children at this age integrate the image into their sense of self (Povinelli et al., 1996), but given that Povinelli et al. also failed to find a reliable association between labeling and performance on the mark test, the implications of the age difference in labeling remain unclear.

Nonetheless, 4-year-olds’ relative success on the mark test clearly indicates that they were more likely than 3-year-olds to understand the relevance of delayed external representations of self and to map information from these representations onto themselves (Smith, 1997). Three-year-olds’ difficulty on the task is more difficult to explain, but the results of Experiment 1 extend previous findings in several ways that constrain possible explanations of what changes between 3 and 4 years. First, children were no more likely to exhibit mark-directed behavior after 5-s than they were after 3-min. In fact, the means were in the opposite direction, although they were not significantly different. Five seconds may be too long to fall within 3-year-olds’ temporal span, but it seems more likely that children at this age have difficulty with any image that is not live. If this is the case, then any age-related changes in temporal span may involve an abrupt transition from being unable to handle even very brief delays at 3 years to being able to handle delays of several minutes or more at 4 years. However, it remains to be determined whether 3-year-olds’ sense of self is indeed so severely present-oriented or whether their poor performance on delayed self-recognition with very brief delays can be attributed to something else, as is discussed further.

Experiment 1 also revealed that children in the photograph condition performed better on delayed self-recognition than did children in the drawing condition. Two possible explanations of this difference might be noted. First, research by O’Connor, Beilin, and Kose (1981) suggests that children consistently judge photographs to represent reality more accurately than drawings, and children may have greater difficulty recognizing themselves in images that provide less verisimilitude. Second, children may believe that it is more plausible for an experimenter arbitrarily to add something (e.g., a mark) to a drawing than to a photograph. In any case, however, this finding indicates that performance on the mark tests is due, at least in part, to the specific media involved, and it supports the suggestion that age-related changes on these tests primarily reflect children’s growing understanding of external representations (e.g., DeLoache, 1991; Liben, 1999) rather than their developing self-concept.

In addition to testing the effects of delay and media, Experiment 1 was designed to examine whether performance on the mark tests was related to children’s theory of mind, as assessed on tests of false belief, appearance reality, and representational change. Povinelli (1995) and Povinelli et al. (1996) predicted that the acquisition of a temporally extended self (as demonstrated on the delayed-mark test) would be related to the acquisition of a theory of mind because both acquisitions may be due to the same underlying domain-general changes in children’s capacity for representation. The results of Experiment 1 failed to support this prediction. Saddendorf (1999) also failed to find a correlation between performance on delayed-media mark tests and a test of false belief. However, it is possible that the failure to detect a significant correlation can be attributed to methodological insensitivity (e.g., to the poor test–retest reliability of false-belief tests; Mayes, Klin, Tercyak, Cicchetti, & Cohen, 1996), and future work might usefully address this issue.

Perhaps the most striking finding from Experiment 1 was that children performed poorly on the training tasks and that performance on these tasks was related to performance on the tests of delayed self-recognition. The training tasks did not assess self-recognition, but as in the mark tests, they required children to map information from a delayed-media representation to its referent. Children’s poor performance on the training task thus provides preliminary evidence that children in this age range have difficulty reasoning about external representations in general, rather than delayed representations of the self in particular. Consequently, poor performance on the delayed-mark tests, in and of itself, cannot be taken to reflect a temporally restricted sense of self. Instead, difficulty with the method of assessment may mask evidence of a more mature self-concept.

Experiment 2

In Experiment 2 we sought to test directly the possibility that young children’s difficulty on mark tests of delayed self-recognition was not specific to representations of self. Instead of using a training task, we presented 3- and 4-year-old children with two closely matched mark tests: one that required them to use a delayed-video representation of self (self task) and one that required them to use a delayed-video representation of a stuffed animal (other task). The marking event (of self or other, depending on the task) took place in the context of the Duck, Duck, Goose game (as described in Experiment 1), and children viewed video feedback of the marking event 90 s after completion of the game. Video images were used to make the procedure more comparable with that used in the main experiment by Povinelli et al. (1996, Experiment 1). The dependent measure was when in a series of increasingly supportive prompts children indicated the location of the sticker. Children’s verbal labeling of their images was not recorded because the focus of this experiment was the comparison of children’s performance on the two types of mark test. We reasoned that if 3-year-olds had difficulty on both tests, then the masking hypothesis would need to be considered seriously, and future research would need to address more precisely the circumstances in which children have difficulty reasoning about delayed external representations.

Method

Participants

Nineteen three-year-olds and 20 four-year-olds were recruited through local day-care centers as in Experiment 1. Three of these children (1
three-year-old and 2 four-year-olds) were tested but were not included in the final data set because they reached up to the sticker before the media presentation on the self test. Thus, the final sample comprised 18 three-year-olds (mean age = 36.1 months; range = 29.7 to 41.5) and 18 four-year-olds (mean age = 48.5 months; range = 43 to 54.1). There were 12 girls and 6 boys among the 3-year-olds and 6 girls and 12 boys among the 4-year-olds. None of the children in Experiment 2 participated in Experiment 1.

**Materials**

A color video camera and a color television (15 x 18 cm) were used to record the marking events and play them back to children. The mark tests involved two large stuffed animals (a beaver and a mouse), a tripod, a large piece of fabric, and brightly colored stickers (5 x 5 cm). The PPVT-R (Form L; Dunn & Dunn, 1981) was used to establish a vocabulary count for children.

**Design**

All of the children received two mark tests (self and other), both of which were presented in the context of Duck, Duck, Goose. The order of the tests was counterbalanced. Between mark tests, children received a modified version of the PPVT-R (Dunn & Dunn, 1981; see below).

**Procedure**

Mark tests. The procedure for the self version of the mark test was similar to that for the mark tests in Experiment 1. The marking event took place in the context of Duck, Duck, Goose, which was played with two female experimenters. However, in this experiment, prior to the game, children were told, "Helena (secondary experimenter) is going to take some pictures of us while we play the game and then later you'll get to see yourself on TV," and then the secondary experimenter started video recording. The circle consisted of two stuffed animals and the child (i.e., the secondary experimenter did not play). One stuffed animal was elevated (seated on a tripod that was covered with a large cloth) so that children could not see the top of its head. Additionally, there were only three trials of Duck, Duck, Goose compared with five trials in Experiment 1. The marking event occurred on the second trial. The third trial served to ensure that children had not noticed the marking event (in Experiment 1, we noticed that children who reached up prior to the media presentation did so on the third trial). Children who reached within an inch of the sticker prior to the media presentation (n = 3) were excluded from the final sample. All of the children received video feedback 90 s after the completion of the game. The videotape was rewound to 2 s before the marking event and then allowed to play for a total of 75 s. During this period, children were not allowed to play. The numbers of children who received scores of 0 through 4 on the self test were 12, 2, 4, 5, and 13, respectively. For the other test, the corresponding numbers of children were 11, 1, 4, 6, and 14.

Preliminary analyses revealed no effect of gender, so data for boys and girls were collapsed. A 2 (age group) x 2 (task) x 2 (task order) mixed-measures ANOVA revealed that 4-year-olds (M = 2.78, SD = 1.31) performed significantly better than 3-year-olds (M = 1.67, SD = 1.63), F(1, 32) = 5.25, p < .05, but there was no effect of task, F(1, 32) = 0.46, p = .5, Cohen's d = .16 (M for self = 2.14, SD = 1.74; M for other = 2.31, SD = 1.72). No other effects or interactions reached statistical significance. (Note that the lack of a Task x Task Order interaction indicates that there were no discernible practice effects in this experiment.)

As in Experiment 1, performance on each mark test was also scored categorically according to whether children located the sticker at any time from the onset of media presentation until they were asked, "Where is the sticker?" (i.e., in this experiment, according to whether they achieved a score of 2 or greater on each mark test). According to this criterion, 8 three-year-olds (44%) passed the self test and 10 (56%) passed the other test, whereas 16 four-year-olds (89%) passed the self test and 15 (83%) passed the other test.

**Correlations between mark tests.** There was a highly significant positive correlation between scores on the self and other tests (r = .62, p < .0001), which remained significant with age and vocabulary scores partialled out (r = .49, p < .005).

**Discussion**

The percentages of 3- and 4-year-olds passing the self test (44% and 89%, respectively) in Experiment 2 were comparable with the...
percentages found in corresponding conditions in previous studies (e.g., 25% and 75% from Povinelli et al., 1996, Experiment 1; and 43.8% and 87.5% from Povinelli & Simon, 1998, standard condition and during presentation). However, despite differences in the coding criteria that had the potential to work against our hypothesis, we found that children's performance improved with age on both self and other tests, and performance on the two tests was highly correlated even with age and vocabulary partialed out. Moreover, the magnitude of the correlation between performance on self and other tests was comparable with that found between the mark tests and the training task in Experiment 1. These findings indicate clearly that 3-year-olds' difficulty on tests of delayed self-recognition is not specific to images of self and may instead reflect a more general difficulty reasoning about delayed external representations. Because poor performance on tests of delayed self-recognition can be attributed to this more general difficulty, it remains to be determined whether delayed-media mark tests are actually sensitive to age-related changes in children's self-concept.

Research by Suddendorf (1999) provides converging evidence that young children's difficulties on tests of delayed self-recognition are not specific to representations of the self. Suddendorf independently found that 3-year-olds not only had difficulty on delayed self-recognition but also frequently failed to use delayed-video feedback to find an object hidden in a room. In the latter task, children were shown that a carton was empty and then were provided with video feedback that revealed an experimenter placing a teddy bear into the carton. Half of the 3-year-olds failed to retrieve the teddy bear, and performance on this task was related to performance on the test of delayed self-recognition. Suddendorf's teddy-bear task would seem to be comparable with the other task used in Experiment 2 insofar as both tasks require children to use delayed-video information to locate something in a room instead of on themselves.

Three-year-olds' general difficulty reasoning about delayed representations in the present experiments, and in the study by Suddendorf (1999), may depend in part on their fragile understanding of the relation between external representations and reality. The fragility of this understanding has been amply demonstrated (e.g., Flavell, Flavell, Green, & Korfmacher, 1990; Smith, 1997; Zaichik, 1990). For example, Flavell et al. (1990, Experiment 2) showed 3-year-olds a photograph or a videotape that depicted a glass filled with liquid. Children were asked whether the liquid would spill out when the photograph or television were turned upside down, and a substantial percentage of the children said that it would (25% for the photograph and 33.3% for the video). Flavell et al. suggested that children who fail this task do so because they consider only the referent when answering the question; essentially, they find it difficult to think of one thing in two different ways at the same time. Likewise, in a study by Zaichik (1990), preschoolers saw an actor take a photograph of an object placed at a particular location. The object was then moved to another location, and children were asked, "In the picture, where is the object?" Three-year-olds were about equally likely to say that the object was at its current location (i.e., where it was in reality) as they were to say that the object was where it had been located at the time that the picture was taken. Together these studies suggest that even 3-year-olds sometimes fail to take a "dual orientation" toward a photograph or video image (DeLoache, 1989, 1991).

Under certain circumstances, children at this age apparently fail to consider the implications that external representations have for reality. In fact, children's verbal responses to the presentation of the self image in Experiment 2 suggest that they may have concluded erroneously that the media representation was not veridical: When presented with an image, several children said something along the lines of "That's me, but I don't have a sticker on my head." A similar account could be given for children's performance in the other version of the task.

At the same time, however, children as young as 2.5 years have been found to use external representations to guide their behavior in some circumstances. For example, Troseth and DeLoache (1998, Experiment 1) showed 2.5-year-olds a live video of a hiding event that was taking place in another room and required them to use the video information to retrieve the hidden object from the room. Children at this age tended to search successfully, indicating that they appreciated the relevance of the video information and used that information to guide their search. How might this finding be reconciled with evidence that suggests that children have difficulty using external representations? One possibility is that 3-year-olds have difficulty using delayed (vs. live) representations to guide their behavior when the information provided by these representations conflicts with a salient alternative (e.g., reality).

Indeed, a conflicting-representations account could be applied to the findings from Experiments 1 and 2. In the self test in Experiment 2, for example, 3-year-olds could have made a strong assumption that they did not have a sticker on their head because they did not see it placed there and could not see it directly at the time of testing. When provided with conflicting information through a delayed and dimly understood representational medium (in this case, video), the conflicting information may have been ignored or treated as somehow irrelevant to the situation. Regardless of the temporal extension of their self-concept, children might have concluded erroneously that the media representation was not veridical. Experiment 3 was designed to test this conflicting-representations account directly.

Experiment 3

In Experiment 3, we used a version of Troseth and DeLoache's (1998) procedure to test whether 3- and 4-year-olds have difficulty using external representations to guide behavior when the representations are delayed and when the information provided by the representations conflicts with their prior beliefs. On some trials (conflict trials), children (a) witnessed a hiding event at one location, (b) were told that the location would be changed, and (c) were given delayed-video or delayed-verbal information regarding the new location of the hidden object. Other trials (no-conflict trials) were similar except that children did not witness the initial hiding event. Thus, on conflict trials, children received delayed information that was in conflict with their initial belief about the object's location (formed on the basis of direct perceptual experience), whereas on no-conflict trials, they received delayed information in the absence of an initial belief. Two types of representation, video and verbal, were used to determine whether any difficulty was specific to video.
Method

Participants

Thirty-four children, including 17 three-year-olds and 17 four-year-olds, were recruited through a computerized database that consisted of names of parents who had expressed an interest in participating in our research. Children came from a variety of socioeconomic and ethnic backgrounds, although data about their backgrounds were not systematically collected. None of these children had participated in the previous experiments. Four children were tested but were not included in the final data set because they refused to cooperate during test trials (2 three-year-olds and 2 four-year-olds). Thus, the final sample included 15 three-year-olds (mean age = 38.3 months; range = 34.0–42.0) and 15 four-year-olds (mean age = 49.0 months; range = 44.7–52.5). There were 9 boys and 6 girls among the 3-year-olds, and 7 boys and 8 girls among the 4-year-olds.

Materials

The recording instruments used in both the training and test phases of the study included two color video cameras and a color television (22.5 X 30 cm). A duck puppet (Huey) served as an intermediary between the experimenter and the child. One testing room (2 X 4 m) served as the hiding room, and six hiding locations were arranged as furniture in this room. These items included a pillow, a chair, a garbage pail, a basket, a bag, and a box. The hidden object consisted of either popcorn or M&Ms (chosen by the child or parent) placed in a small Tupperware container. The training procedure took place in another room (the training room), which was connected to the hiding room by a short (5-m) corridor.

Design

Each child received 12 hide-and-seek trials. These included four sets of 3 trials that corresponded to each of four trial types created by crossing two within-subject variables: conflict (conflict vs. no conflict) and medium (verbal vs. video). The resulting trial types were (a) conflict video (i.e., shown a location directly then presented with conflicting video information), (b) conflict verbal (i.e., shown a location directly then presented with conflicting verbal information), (c) no-conflict video (i.e., simply shown video information regarding a location), and (d) no-conflict verbal (i.e., simply given verbal information regarding a location). Each child was randomly assigned without replacement to four different orders of trial type, each of which involved a different repeating sequence of the four trial types (e.g., 4, 1, 2, 3). Across the four orders, each trial type appeared once in each position in these repeating sequences. The order of hiding locations was also counterbalanced by using two predetermined hiding sequences (A and B). The two sequences controlled for possible location effects by having three of the six hiding locations (e.g., pillow, chair, and garbage pail) used with trials of a given type (e.g., with no-conflict video trials) in Sequence A, and the other three hiding locations used with those trials in Sequence B.

Procedure

Children were tested individually by a female experimenter. The session lasted approximately 45 min and began once children appeared comfortable with both the experimenter and the setting.

Training task. While in the training room, children were first introduced to the puppet, Huey. Children’s attention was directed toward the video camera and they were asked, “What’s this?” and “Can you tell me what it does?” Regardless of their answers, children were told, “This is a camera and it takes pictures of what we do. It copies us, and afterwards, we can watch what we just did on the television. Huey just told me that he really wants to get his picture taken. Let’s take a picture of Huey saying hello to you.” After Huey waved into the camera and said, “Hello,” children were told, “The camera just copied Huey and it can go back to the beginning and show us what Huey just did.” Children were then asked the first of a series of six training questions that were designed to provide basic information about the video camera. The procedure for the training task differed from that used in Experiment 1 to account for the differences between paradigms. Questions 1 and 2 were “We just took a picture of Huey waving; who are we going to see on TV?” and “What is he going to do?” Questions 3 and 4 were self-directed. Children were told that it was their turn to get their picture taken. The camera was turned on and they were asked to say their favorite color. Question 3 was “We just took a picture of you saying your favorite color; who are we going to see on TV now?” This was followed by Question 4, “What are you going to be saying on TV?” The final two training questions involved the concept that what was videotaped in the other room could also be seen on TV. Children were told, “There is another camera in the other room just like this one; it also takes pictures and copies what we do. I’m going to go into the other room and wave into the camera for you. Then I’m going to come back and we can watch me waving on the TV.” After the experimenter had returned from the other room, children were asked Question 5, “Who are we going to see on TV now?” and Question 6, “What am I going to be doing.” If children answered a question correctly, they were praised. If children answered a question incorrectly, they were told the correct answer and reminded, “Remember the camera takes pictures of things exactly the way they are, and we can watch them after.” In all cases, children were shown the relevant recording on TV, and the correspondence between the act of videotaping and the subsequent playback on TV was emphasized. Children received 1 point for each of the six questions they answered correctly.

Hide-and-seek trials. After the training task, children were shown the hiding room and it was explained that the room was Huey’s. During the orientation, the experimenter showed the child each hiding location (and labeled them) in a fixed order. Children were told, “Let’s see what’s in his room,” and for each hiding location, they were asked, “What’s this called?” The experimenter then restated children’s correct responses. If children answered incorrectly (only 2 children failed to label one or more of the hiding locations correctly), the experimenter told children the correct answer and repeated the orientation. Children were then shown the video camera located near the entrance to the room and were reminded that this was the camera that took a picture of the experimenter waving, which they then saw on television in the training room. Following the orientation, children accompanied the experimenter back to the training room and the rules of hide-and-seek were explained. Children were told that for some trials they were allowed to hide the object with Huey, whereas on other trials they had to wait while the object was being hidden. They were also informed that either they would be told where to look or they would watch the television, which would show them where to look. Furthermore, it was stressed that “in this game, as soon as you know where the popcorn/chocolate is hiding, that’s the spot where you look first.” On the two types of conflict trial, children accompanied Huey and the experimenter to the hiding room, helped hide a treat in a predetermined location, and then were told to remember the location. On returning to the training room, children were told, “Huey and I know a better spot. We’ll be right back.” The children waited with their parents while the experimenter proceeded to switch the hiding location. Children then received either video information or verbal information regarding the new location. On the conflict video trials, after the experimenter switched the location, children were told, “Now we’re going to watch the TV and see what Huey just did,” at which point they watched a 10-s prerecorded video segment that depicted the experimenter and Huey hiding the object in a new location. The perspective from which the video segments were filmed was the same as the perspective that children had on entering the hiding room. The experimenter wore the same outfit on the video and during testing. The conflict verbal trials were identical to the conflict video trials except that the information regarding the object’s new location was presented
verbally. On the conflict verbal trials, the experimenter said, "Now Huey's going to tell you what he just did." The puppet then told the children where the object was hidden (e.g., "It's behind the pillow."). On both types of trial, a confirmation check was made on each trial to ensure that children encoded the information they received. On the conflict video trials, the video was paused at the end of the clip and children were asked to point to where the object was on the TV screen ("Can you point to the TV and show me where the popcorn/chocolate is hidden?"). On the conflict verbal trials, children were asked, "Can you tell me where Huey just said the popcorn/chocolate is hidden?" All children answered the confirmation questions correctly on the conflict verbal trials, but one 3-year-old pointed to the initial hiding location on all three conflict video trials, and one 4-year-old refused to answer the question on one conflict video trial.

After the confirmation check, children were escorted to the hiding room and were allowed to search for the hidden object. If children erred on their first attempt, the experimenter said, "It's not there, keep looking." After a third incorrect attempt, the experimenter said, "Why don't you look over here?" and pointed to the correct location.

The procedures for the two types of no-conflict trial (verbal and video) were identical to that for the conflict trials except that children remained in the training room while the experimenter and Huey went to hide the object, and hence, children did not receive direct experience of the hiding event. When the experimenter returned to the training room, she said, "Huey and I know a better spot. We'll be right back," exactly as on the conflict trials. The children waited with their parents while the experimenter went to the hiding location and pretended to switch the hiding location. The experimenter then returned to the training room, and children received either video information or verbal information regarding the location of the treat, as on the conflict trials. Following the confirmation check, children were allowed to search for the hidden object. Two 3-year-olds answered one of the no-conflict video confirmation questions incorrectly, and one 3-year-old answered two incorrectly. This last child pointed to the locations where she had retrieved the object on the previous trial. Children’s responses in this experiment were clear and unambiguous, and there was no need to calculate reliability of scoring.

Results

Training Task

Performance on the training task was scored out of 6. A one-way ANOVA on training scores revealed that the difference between 3-year-olds ($M = 4.93, SD = 0.59$) and 4-year-olds ($M = 5.47, SD = 0.83$) was marginally significant, $F(1, 28) = 4.07, p < .06$. However, only 4 children, including 3 three-year-olds and 1 four-year-old scored less than 5 out of 6. Errors were distributed fairly evenly across questions.

Hide-and-Seek Trials

Errorless retrievals. The dependent variable on the hide-and-seek trials was the number of errorless retrievals (out of 3) on each type of trial (i.e., the number of trials on which children found the object on their first attempt). Preliminary analyses on gender and trial block (performance on the first six trials vs. the last six trials) revealed no effects of either variable so the variables were dropped from subsequent analyses. Next, data were analyzed using a 2 (age: 3 years vs. 4 years) $\times$ 2 (conflict: conflict vs. no conflict) $\times$ 2 (medium: verbal or video) mixed ANOVA with age as a between-subjects variable and conflict and medium as within-subjects variables. The only significant effects to emerge from this analysis were main effects of age, $F(1, 28) = 27.69, p < .0001$, and conflict, $F(1, 28) = 14.58, p < .001$, which were subsumed by a significant interaction between age and conflict, $F(1, 28) = 19.22, p < .0001$. The main effects indicated that 4-year-olds ($M = 11.06, SD = 0.96$) performed better than 3-year-olds ($M = 7.53, SD = 2.41$), and children performed better on no-conflict trials ($M = 4.96, SD = 1.21$) than on conflict trials ($M = 4.20, SD = 1.86$). However, the interaction indicated that the effect of conflict was only significant for the 3-year-olds (conflict trials vs. no-conflict trials, $p < .01$). Tukey’s honestly significant difference (HSD) tests (see Figure 1). Two 4-year-olds’ performance was better than 3-year-olds’ on both conflict ($p < .01$), Tukey’s HSD) and no-conflict trials ($p < .05$, Tukey’s HSD), although the age difference was considerably more pronounced on the conflict trials (see Figure 1).

Classification of 3-year-olds’ errors on the conflict trials.

Three-year-olds made a total of 48 errors on the conflict trials. The two most common types of error were searching where the object had initially been hidden (73% of the errors) and returning to the location where the object had been found on the previous trial (19% of the errors). Three-year-olds selected the correct hiding location on their second attempt after an initial error in 24 (50%) of the cases. The most common type of error was to search for the object in the correct hiding location but at the incorrect time (50% of the errors). Confusion of the hiding location with the hiding time is a common error, and this error was more common on the no-conflict trials ($M = 2.12, SD = 1.03$) than on conflict trials ($M = 2.73, SD = 0.53$) but that this difference was not significant for verbal information (conflict $M = 2.38, SD = 0.90$; no-conflict $M = 2.58, SD = 0.58$).

![Figure 1](image-url)  
Figure 1. Mean number of errorless retrievals (and standard error) by age and trial type in Experiment 3.
of the 48 cases. (Four-year-olds made too few errors to warrant classification.)

Discussion

The results of Experiment 3 provide support for the hypothesis that 3-year-olds have difficulty using delayed external representations to guide their behavior when the information provided by the representations conflicts with their initial expectations (which, in this case, were based on direct experience). As might be expected based on previous findings (Trosset & DeLoache, 1998), both 3- and 4-year-olds were able to use video (and verbal) information to guide their behavior in a hide-and-seek search paradigm in the absence of a prior, conflicting expectation. However, there was a notable difference between the 3- and 4-year-olds’ performance on the conflict trials. On the conflict trials, 3-year-olds frequently failed to use the information provided by the delayed representations.

Classification of the errors that children made on conflict trials revealed that children generally exhibited one of two distinct types of perseveration: returning to the initial location or searching at the location where they had retrieved the object on the previous trial. Children’s relative success on their second choice of location after an initial error indicates that, in many cases, they had indeed stored the relevant information, even though they did not use it initially. Like the delayed-media mark tests, the conflict trials in Experiment 3 put two bases of responding into competition, and with age, children were more likely to act on the basis of the delayed external representation. However, compared with the delayed-video mark test, in which children are not provided with any reason to believe that they might have stickers on their heads, the degree of conflict on the conflict trials in Experiment 3 is likely to be attenuated because children are told explicitly that the puppet knows a better spot and is going to switch the location of the treat. This rationale should increase the likelihood that children will act on the basis of the external representation. It may be useful in future studies to compare performance on mark tests directly with performance on other tasks that have the same requirements. Future research might also attempt to determine which aspects of the procedure on the conflict trials are necessary and sufficient to elicit errors in 3-year-olds. The results of Experiment 3 suggest minimally that 3-year-olds will have difficulty on tests involving delayed external representations that conflict with children’s expectations based on direct experience. However, it remains to be determined whether delay is necessary to cause difficulty for 3-year-olds or whether conflict alone will suffice.

General Discussion

Experiments 1 and 2 replicated the basic finding from Povinelli et al. (1996) that there is an age-related change between 3 and 4 years in children’s use of delayed-media representations in a mark test, and they add to this finding in several ways. First, results from Experiment 1 suggest that dramatically decreasing the length of the delay between the marking event and the media feedback does not improve 3-year-olds’ performance. Even though 5 s might be expected to fall within the span of children’s present self (Povinelli, 1995), 3-year-olds had difficulty using drawings and photographs to locate the mark. This finding, together with 3-year-olds’ success with mirrors (e.g., Povinelli et al., 1996), suggests that children at this age may have difficulty using any feedback that is not obviously contingent on their movements (i.e., that is not live). Second, although age differences were found for both photographs and drawings in Experiment 1, children performed better with photographs. Children may have greater difficulty recognizing their own image in a drawing versus a photograph (perhaps because drawings often have less verisimilitude than photographs) or they may believe that drawings are more likely than photographs to misrepresent something. In either case, the effect of media type is consistent with the possibility that performance on delayed-mark tests depends in part on children’s understanding of different types of media representation. Also consistent with this possibility were the findings that performance improved over trials in Experiment 1, 3-year-olds had difficulty on the training task in that experiment, and performance on this training task was correlated with performance on the mark test. Experiment 2 explored this possibility more directly and revealed that 3-year-olds exhibited as much difficulty on an other version of the mark test as they did on a closely matched self version. Together, these findings indicate that children’s difficulty using delayed-media representations to guide behavior is not specific to tasks involving the self.

Children’s understanding of the medium of self-reflection has been addressed in the context of research on mirror self-recognition, and several studies have suggested that children’s ability to recognize their own image is not related to their ability to reason about mirror images in general. For instance, Priel and De Schonen (1986) found that toddlers with previous mirror experience performed similarly to toddlers without such experience on a mark test and in self-naming. Previous mirror experience was, however, related to toddlers’ ability to locate a toy behind themselves (Priel & De Schonen, 1986). Additionally, Robinson, Connell, McKenzie, and Day (1990) found that only a minority of 22-month-olds were able to locate an object using a mirror, although most were able to label their own image. Together, these findings indicate that children’s ability to reason about mirror representations of objects is unrelated to their ability to identify their own image. However, our findings suggest that self-recognition on a delayed-media mark test is related to a general ability to use external representations, at least under some circumstances.

Experiment 3 was designed to identify more precisely the circumstances in which 3-year-olds have difficulty using external representations to guide their behavior. In this experiment, it was shown that 3-year-olds use delayed-video and delayed-verbal representations to guide their search for a hidden object in the absence of a conflicting expectation about the object’s location, but they have difficulty doing so in the presence of a conflicting expectation. In most cases when they erred, they searched in a way that was consistent with their initial, conflicting expectation, which was based on their own, unmediated experience. This finding adds to the growing corpus of data concerning children’s use of external representations (e.g., DeLoache, 1991; Liben, 1999; Trosset & DeLoache, 1998) and suggests that the presence or absence of conflict in assessment procedures will affect estimates of children’s understanding. Conflicting expectations may be especially difficult for children to ignore when they are based on direct experience, but more generally, the influence of information on children’s behavior may vary with children’s understanding of the
medium by which it is presented. A related point has recently been made by Hudson and Sheffield (1998), who noted that there may be age-related changes in the types of reminders (e.g., direct reenactment, verbal recall, or video) that improve young children’s long-term memory for events. Although there was an effect of conflict in Experiment 3 for both video and verbal information, this effect was more robust for the video information than for the verbal information (see Footnote 2), and it seems likely that any interaction with media could be explained in terms of children’s relative familiarity with verbal instructions versus video information.

Three-year-olds’ difficulty using delayed external representations when they conflict with their initial expectations could arise from a number of sources. For example, 3-year-olds may have a fragile understanding of representations in general (e.g., DeLoache, 1989; Liben, 1999; Perner, 1991; Zaitchik, 1990), and this may decrease the likelihood that they will rely on them when alternative sources of information are available. Alternatively, children at this age may have a general difficulty integrating conflicting information (e.g., Flavell, 1988), perhaps because of memory limitations (e.g., Olson, 1993) or difficulty using higher order rules (e.g., Frye et al., 1995; Zelazo & Frye, 1997). On this last account, a higher order rule is required to select a perspective from which to respond; in the absence of a higher order rule, children will respond on the basis of the most salient, or default, perspective. In the case of the tasks used in the present experiments, a higher order rule may be required to act on the basis of delayed representations when those representations conflict with children’s initial expectations. That is, children will first need to choose the relevant representation or perspective, and then determine what is indicated by that representation (e.g., “According to the television, the object is under the pillow, not in the garbage pile, so that’s where to look.”). Like other theories that focus on general cognitive changes, however, this account would need to be extended to capture the effect of different media and the failure to find a correlation between delayed self-recognition and theory of mind. It is clear that additional research on 3-year-olds’ difficulty with use of external representations is warranted.

Regardless of the underlying source of 3-year-olds’ difficulty, however, the present findings have implications for research on the development of a temporally extended sense of self. Because children have difficulty using conflicting delayed representations in general, poor performance on tests of delayed self-recognition does not necessarily indicate an immature self-concept. Instead, difficulties that are due to the method of assessment may (or may not) mask a more mature self-concept, and additional evidence will be required before changes in performance on the mark test of delayed self-recognition can be attributed to the development of a temporally extended self.

References


Received July 18, 1997
Revision received December 11, 1998
Accepted December 11, 1998