Confusions between Memories for Performed and Imagined Actions: A Developmental Comparison

Mary Ann Foley
Skidmore College

Marcia K. Johnson
State University of New York at Stony Brook

FOLEY, MARY ANN, and JOHNSON, MARCIA K. *Confusions between Memories for Performed and Imagined Actions: A Developmental Comparison.* CHILD DEVELOPMENT, 1985, **56**, 1145–1155. Children are often assumed to be more confused than adults are about the origin of self-generated memories (e.g., what they did or thought). The present experiments showed evidence in support of this assumption but only under some circumstances. In Experiment 1, 6- and 9-year-olds were as good as adults in distinguishing what they did from what they saw someone else do. However, children had particular trouble distinguishing what they did from what they imagined doing. Confusion between performed and imagined actions was evident across a range of actions. Clustering data also showed that information about origin is part of the memory for an event; all subjects recalled actions according to who performed what action (Experiment 1). Further, the presence of person categories as a basis for organization reduced clustering based on action class more for children than for adults (Experiment 1 vs. 2). Collectively, these findings indicate that children become sensitive to some distinctions in memories sooner than they do to others.

Many of us believe that children have more difficulty than adults do distinguishing reality and fantasy. Anecdotal reports of children's tendency to confuse their dreams with their waking lives as well as their insistence on the real existence of imaginary companions help perpetuate this notion. This hypothesized developmental trend appears in many theoretical discussions (Broughton, 1978; Brown, Bransford, Ferrara, & Campione, 1983; Flavell, Flavell, & Green, 1983; Foulkes, 1982; Freud, 1922/1953; Morison & Gardner, 1978; Piaget, 1929, 1959; Selman, 1981; Vygotsky, 1962). For example, Werner (1948) wrote that "it is characteristic of primitive mental life (at least until 6 or 7) that it reveals relatively limited differentiation of object and subject, of perception and pure feeling, of idea and action" (p. 59). According to these discussions, the lack of differentiation between object and subject and between idea and action contributes, in part, to children's proclivity to confuse reality and fantasy.

If children have trouble separating ongoing perceptions from thoughts or ongoing ideas from actions, then they should have even more difficulty distinguishing whether a past event was originally real or fabricated. Consistent with this idea, Flavell et al. (1983) recently speculated that children's tendency to tell "whoppers" may be a consequence of their failure to remember that the incidents they are reporting are based on their own fantasies (self-generated events).

Actually, there are many potential sources of confusion involving self-generated memories. Thinking about a past perceptual experience and confusing that thought with the original experience is one example (Johnson, Raye, Hasher, & Chromiak, 1979). Other types of self-generated memories may also be mistakenly attributed to a perceptual origin. For example, both children (Piaget, 1962) and adults (Johnson, Kahan, & Raye, 1984) are sometimes confused about the origins of their

This research contributed to the completion of the Ph.D. requirements at the State University of New York at Stony Brook of the first author. Preparation of the article was supported in part by NSF grant BNS 26145 and NIMH grant MH 37222. We gratefully acknowledge Hugh Foley and three anonymous reviewers for their comments on earlier versions. Also, we thank the children and staffs from the schools in which the studies were conducted, especially Sr. Joan Leavey, Sr. Anne Kelly, and Fr. Vin Rush. Finally, the undergraduates who "played detective" are very much appreciated. Requests for reprints should be sent to Mary Ann Foley, Department of Psychology, Skidmore College, Saratoga Springs, NY 12866-0851.

[*Child Development*, 1985, **56**, 1145–1155. 0009-3920/85/5605-0002$01.00]

dreams. Inferential processes that elaborate ongoing perceptual experiences have memorial consequences (e.g., Bransford & Johnson, 1973), and memories based on these inferences are potentially confusable with perceptual events (Johnson, Bransford, & Solomon, 1973). Also, on occasion, adults mistakenly think that they did something that they only intended to do (e.g., lock the door) (Norman, 1981; Reason & Mycielska, 1982). As these examples show, people are sometimes confused about the origin of memories. However, while it is often assumed that children are uniformly worse than adults in all discriminations involving self-generated memories (Johnson & Foley, 1984), our work suggests that developmental differences in discrimination are limited to specific circumstances (Foley, Johnson, & Raye, 1983; Johnson, Raye, et al., 1979).

For example, Foley et al. (1983) reported that 6-year-olds were as good as adults when discriminating what they said from what they heard. Subjects of all ages found it more difficult to discriminate what they said from what they imagined themselves saying. However, 6-year-olds were at a particular disadvantage. This inferior performance of 6-year-olds was not because this age group is generally confused about memories for imaginations. In a replication, they again had difficulty discriminating what they said from what they imagined saying, but they were able to discriminate what they heard from what they imagined saying.

On the basis of our developmental work, then, it is difficult to argue that children have a general deficit in making decisions involving self-generated memories because they were as good as adults in some conditions. Subjects from all age levels experienced more difficulty on some discrimination problems than on others, and 6-year-olds were worse only in specific circumstances.

The Reality Monitoring model developed by Johnson and Raye (1981) provides a theoretical framework for interpreting these differences between conditions and for interpreting these developmental trends. According to the model, memory representations for perceptions (e.g., what someone said) and for self-generations (e.g., what you said or thought) differ along a number of dimensions. Perceptual memories are generally more detailed and include sensory, temporal, and spatial information. On the other hand, memories for self-generations are thought to include information about the cognitive operations involved in the generation process. These differences in information about sensory features and cognitive operations could serve as the basis for decisions about whether a memory was the result of perception or generation. For example, a memory based on listening to what someone said may include more sensory information than a memory based on what one imagined oneself saying. On the other hand, a memory based on imagining oneself speak includes information associated with cognitive operations mediating the imaginations (Foley & Johnson, 1982). Thus, a memory with a great deal of sensory information and little cognitive operations information would be classified as an earlier perception.

The Reality Monitoring model also predicts that subjects would have a more difficult time differentiating what they said from what they imagined saying than they would differentiating what they heard from what they said (or thought). The former discrimination should be harder because memories for performed and imagined speech are presumably very similar (e.g., each including similar amounts of information about cognitive operations and perhaps information about the activation of motor programs); thus subjects' ability to discriminate is decreased. The model also suggests that one reason for the large age difference in performance when subjects discriminate memories for real and imagined speech is that there may be greater overlap in memories for real and imagined speech in younger children. Perhaps for them, thinking is more like speaking, an idea consistent with Vygotsky's (1962) view about the development of the relationship between language and thought. For example, the similarity would be greater if young children subvocalize more than adults on imagination trials, and there is evidence that young children subvocalize (Garrity, 1977). If speech and subvocalized speech produce more similar memories than do speech and imagined speech that is not subvocalized, then young children should have more trouble separating memories for speech and memories for thought than older subjects do.

Children's difficulty distinguishing real from imagined speech may be only one of many possible instances of a more general phenomenon. MacKay (1981) has suggested that the thought processes involved in imagining actions often include aspects associated with the actual performance of those actions. If imaginal activity is typically similar to real actions, people ought to be confused about what they did and what they imagined doing,

and, in fact, they sometimes are (Anderson, 1985; Foley et al., 1983; Norman, 1981). Given that speech is a type of action, our prior findings raise the question of whether children generally have more difficulty distinguishing real from imagined actions than adults do.

It was to this possibility that the present experiments were directed. Actions such as those in a Simon Says game were used. If children's confusion in the Foley et al. (1983) studies reflects a more general lack of differentiation between memories for performed and imagined actions, then children should still be more confused than adults with the inclusion of this wider range of actions.

In addition to making origin discriminations for actions, subjects were asked to recall the actions. If information about origin is part of the memory for an event, it might well form the basis for clustering items in recall. Thus order of recall was expected to provide an additional index for comparing age groups with respect to the salience of information about origin.

Experiment 1

Method

Overview of the design.—There were three main conditions in Experiment 1. In the Do-Watch condition, subjects performed some actions and watched while another person performed other actions. In the Do-Pretend condition, subjects performed some actions and imagined themselves performing other actions. Finally, in the Watch-Watch condition, subjects watched two people performing actions. This latter condition was included to eliminate alternative interpretations of the performance in the Do-Pretend condition (discussed later). Subjects were selected from three age levels (6 years, 9 years, and adult). Thus, two between-subject variables were combined factorially in a 3 (age) × 3 (condition) design with 12 subjects per cell (N = 108).

Selection of materials.—Six-year-olds, 9-year-olds, and adults who did not participate in the main experiment were used to pretest a set of actions, some of which were used in other memory studies (Foellinger & Trabasso, 1971; Johnson, Perlmutter, & Trabasso, 1979; Lieberman & Altschul, 1971; Minas, 1977). Those actions that were performed with considerable variability either within or between age groups were eliminated.[1]

Thirty-six actions, grouped into six classes, were finally selected (see Table 1). Twenty-four of the actions listed in Table 1 (the first four in each class) were randomly assigned to one of the two types of items within each condition for the activity phase of the experiment. For example, in the Do-Pretend condition, a subject performed 12 actions and imagined performing 12 other actions. Within each type of item, an equal number of actions from each action class was assigned. Thus a subject in the Do-Pretent condition traced two objects and imagined tracing two others. Assignment of actions to item types was counterbalanced so that each action in each class occurred as often as both types of items. The last two items listed in each class in Table 1 were randomly designated as new items for subsequent memory tests. Similar assignments were made in the

TABLE 1

ACTION CLASSES, EXPERIMENT 1

Communicative gestures:	Tracing exercises:	Looking at objects:
Shake your head yes	Trace over this letter *A*	Look up at the ceiling
Shake your head no	Trace over this letter *B*	Look at the door behind you
Wave goodbye	Trace over this square	Look under the table
Make a happy face	Trace over this circle	Look at the floor
Make a sad face	Trace over this letter *C*	Look at the book
Clap your hands	Trace over this triangle	Look at the lightswitch
Touching body parts:	**Extending body parts:**	**Standing exercises (stand up and . . .):**
Touch your nose	Point your toes out in front	Make a motion like an airplane
Touch your shoulder	Point your fingers out in front	Turn all the way around
Touch your knee	Stretch your arms out to your sides	Do a jumping jack
Touch your elbow	Stretch your legs over to the side	Hop once with both feet
Touch your stomach	Lean way over forward	Run in place
Touch your toes	Lean way over backward	Skip with one foot

[1] Pilot data were collected as part of a senior honors thesis conducted by Kathleen McNelis, State University of New York at Stony Brook, 1981.

other two conditions. The order of occurrence of item type was random, except that both types occurred equally often in each quarter of the list; action classes were represented equally across quarters of the activity phase as well.

Procedure.—All subjects were randomly selected and assigned to one of the three major conditions and were tested individually. Each child was met by a female experimenter and accompanied to the experimental room, and if one or more female experimenters were present, the child was introduced. The instructions used in the Foley et al. (1983) studies, inviting children to play a detective game, were adapted for the present experiment. For example, in the Do-Pretend condition, a child was told: "Good detectives are very careful not to give any clues about what they are thinking. So, sometimes I will ask you to do something and other times I will ask you to pretend or imagine yourself going through the motions involved in doing something else. When you imagine yourself doing something, be careful not to give me any clues or hints about what you are pretending to do. If you sit quietly in the chair, and rest your arms in your lap, it will help you not to give me any clues." When performing and imagining actions, subjects completed actions at their own pace (children and adults performed at the same rates). There were 5 sec between the completion of one action and the initiation of the next. The instructions were modified appropriately for the other conditions and for the adult control groups, all of which were told that they were part of a developmental communication study.

Surprise memory tests.—Following the activity phase and a brief retention interval (3 min), the subject was surprised with two memory tests. First, the subject recalled the actions in any order (verbally or by doing). Then, subjects were given a discrimination test composed of all 24 actions from the activity phase plus 12 new ones, all of which were ordered randomly. For example, in the Do-Pretend condition, the experimenter read a test item, and the subject decided if it was an action that the subject performed, one the subject imagined performing, or a new action.

Subjects.—Children from two surburban schools on Long Island were represented proportionally across age levels and conditions. Their mean ages were 6-7 and 9-6. The socioeconomic and cultural backgrounds of these children (middle class) were quite similar. Subjects in the adult control groups were undergraduate volunteers at the State University of New York at Stony Brook. Males and females were represented proportionally across age groups and conditions.

Results and Discussion[2]

This section includes a discussion of four major dependent variables: recognition (hits and false positives), discrimination, recall, and clustering. In each case, the first step was to calculate an appropriate overall analysis of variance (ANOVA). For multiple comparisons subsequent to the ANOVAs, Scheffé's test was used with the .05 level of significance. When significant interactions were found, simple main effects were conducted following the logic and procedures recommended by Kirk (1968, pp. 176–181) for holding the experimentwise error rate at the .05 level of significance.

Recognition.—To measure subjects' memory for what happened during the activity phase, responses on the discrimination test were first scored for simple recognition (hits), without regard for identification of origin, and for the number of new items incorrectly called old (false positives). A 3 (age) × 3 (condition) × 6 (action class) ANOVA on hits showed a main effect for age, $F(2,99) = 4.13$, $MS_e = .55$, $p < .01$, a main effect for condition, $F(2,99) = 9.94$, $p < .001$, and an interaction between age and condition, $F(4,99) = 3.46$, $p < .01$. Tests for simple main effects showed that recognition of old items was equally good for all ages in the Do-Watch condition ($M = 22.9$; max = 24) and the Watch-Watch condition ($M = 21.0$). In contrast, in the Do-Pretend condition, recognition of old items was poorer for both 6-year-olds ($M = 19.8$) and 9-year-olds ($M = 21$) as compared with adults ($M = 22$), $F(1,99) = 181.78$, and $F(1,99) = 51.34$, respectively.

A 3 (age) × 3 (condition) ANOVA on the false positives showed that children made more errors than adults did: the means were 1.11, .78, and .47, respectively, for the three age groups, $F(2,99) = 5.86$, $p < .01$. False positives also varied with condition: the means were 1.03, .94, and .39 for the Do-Pretend, Watch-Watch, and Do-Watch conditions, respectively, $F(2,99) = 4.46$, $p < .01$.

[2] Comparisons using the error term from the overall analysis are reported here, but the results were the same using separate error terms computed only from the conditions included in the subsequent comparisons (Keppel, 1984).

Thus there were some differences in recognition performance in the three conditions, but, in general, memory for occurrence was quite good (on the average, 91%). However, the important point is that these recognition scores did not correlate with discrimination performance, as the next analysis reported demonstrates.

Discrimination performance.—Subjects' responses on the discrimination test were scored by computing a proportion. For example, in the Do-Watch condition, the number of actions correctly identified as those performed by the subject (e.g., 11) plus the number identified as those performed by the other person (e.g., 10) was divided by the total number of actions correctly recognized as old (e.g., 23). In each condition, these proportions (e.g., [11 + 10]/23 = .91) were averaged across subjects.

There was no systematic relationship between hit rate (as reported in the last section) and discrimination performance. For example, for the 6-year-olds in the Do-Pretend condition, the correlation was .23, a nonsignificant value. This absence of a correlation is representative of other ages and conditions. Thus the discrimination data reported next are not simply another measure of memory for occurrence.

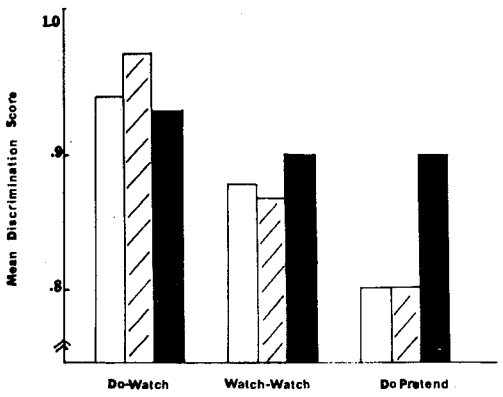

FIG. 1.—Mean discrimination score for the Do-Watch, Watch-Watch, and Do-Pretend conditions. The open bars show the performance of the 6-year-olds, the hatched bars that of the 9-year-olds, and the closed bars that of the adults.

The discrimination scores are shown in Figure 1. A 3 (age) × 3 (condition) ANOVA showed a main effect for age, $F(2,99) = 3.94$, $MS_e = .007$, $p < .02$, a main effect for condition, $F(2,99) = 15.54$, $p < .001$, and an interaction between age and condition, $F(4,99) = 3.98$, $p < .01$.[3]

Tests for simple main effects showed that, in the Do-Watch condition, discrimination performance was quite good and comparable for all age groups ($M = .95$). Although a ceiling effect might have masked developmental differences, the absence of an age trend is consistent with that reported in earlier studies with speech in which performance was not at ceiling (Foley et al., 1983). Also, there were no differences among the age groups in the Watch-Watch condition ($M = .89$), where performance was significantly lower than in the Do-Watch condition, $F(2,99) = 12.07$, $p < .05$. In contrast, discrimination performance varied with age in the Do-Pretend condition, $F(2,99) = 8.17$, $p < .01$; the performance of the 6- and 9-year-olds was comparable ($M = .80$), but both groups performed significantly worse than adults ($M = .92$), $F(1,99) = 12.43$, $p < .05$, and $F(1,99) = 8.97$, $p < .05$, respectively. In addition, for the children, discrimination performance was worse in the Do-Pretend condition than in the Watch-Watch condition, $F(1,99) = 11.20$.

It is important to notice that children were not at a disadvantage as compared to adults in the Watch-Watch condition. This finding shows that the children's inferior performance in the Do-Pretend condition is not due to their having difficulty with any discrimination involving events from the same general class of experience (in the Watch-Watch condition, perceptual experience, and in the Do-Pretend condition, self-generated experience).

In subsequent analyses, we examined the discrimination data in the Do-Pretend condition to see whether subjects were more

[3] The discrimination test actually followed one free recall trial, and the test included both actions recalled and those not recalled. To assess possible effects of recall on the ability to discriminate actions, discrimination scores were computed separately for actions recalled and those not recalled. This within-subject variable was then included in an ANOVA with the other two between-subject variables (age and condition). Discrimination scores for actions recalled ($M = .90$) were comparable to those discrimination scores for actions not recalled ($M = .89$). The overall effects reported for age, condition, and the age × condition interaction were again significant, and there were no interactions between these and the within-subject variable (items recalled vs. items not recalled). Therefore, the results reported for Figure 1 were not contaminated by the prior recall of some items.

likely to make discrimination errors on Do or on Pretend items. That is, were subjects more likely to attribute an action to imagination or an imagination to action? In the six action classes, the only asymmetry involved tracing: subjects were more likely to say that they actually traced an object they only imagined tracing than vice versa.

Recall.[4]—A 3 (age) × 3 (condition) × 6 (action class) ANOVA was conducted on the recall scores. The data in Figure 2 are collapsed across condition because there was neither a main effect nor any interaction between condition and the other variables. The means were 10.00, 10.25, and 11.00 for the Do-Pretend, Watch-Watch, and Do-Watch conditions, respectively. There was a main effect for age, $F(2,99) = 74.84$, $MS_e = 1.18$, $p < .001$. Scheffé's test showed that 6-year-olds recalled fewer actions ($M = 6.72$) than did 9-year-olds ($M = 10.25$) and adults ($M = 14.8$), $F(1,99) = 25.46$, and $F(1,99) = 148.92$, respectively. Also, recall was lower for 9-year-olds than for adults, $F(1,99) = 43.57$.

Recall was greater for some action classes than for others, $F(5,495) = 9.42$, $MS_e = 1.23$, $p < .001$, and this effect was comparable at each age level and condition. Subsequent comparisons using Scheffé's test were made among the means for the action classes; briefly, recall of gestures ($M = .77$, out of 4) < extensions ($M = 1.40$) = touching ($M = 1.45$) = looking ($M = 1.91$) < standing ($M = 2.60$) = tracing ($M = 2.32$).

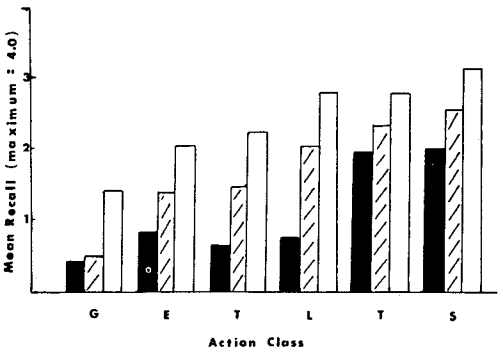

FIG. 2.—Mean recall separated by action class (and collapsed across conditions) for each age, Experiment 1. The closed bars refer to the mean for 6-year-olds, the hatched bars to the mean for the 9-year-olds, and the open bars to the mean for adults. The labels along the *x*-axis refer to action classes: *G*, Gestures; *E*, Extensions; *T*, Touching; *L*, Looking; *T*, Tracing; and *S*, Standing.

Although the action classes obviously varied in terms of the amount of effort required to produce them, they differed in a number of other ways as well (e.g., in time required to complete them, novelty, and complexity). Thus, on the basis of this one study, it is not possible to determine why action class affected recall. However, it is interesting to note that the effects of action class were different for the three dependent measures reported thus far. Recognition was not affected by action class. Discrimination performance was, but only in some conditions (Do-Pretend and Watch-Watch). (The effects of action class on discrimination performance do not require any qualifications of the results reported above.) Recall, on the other hand, was affected by action class in all conditions and in the same manner. This aspect of our findings is further evidence for the dissociation between discrimination and other measures of memory (Johnson, 1985).

Our major interest in action class in the present studies was twofold. First, it was used to assess the generality of the confusion reported in the Foley et al. (1983) studies. Second, as the next analyses show, we were interested in its potential effect on the order in which actions were recalled.

Recall order.—If information about origin (self vs. other) is part of the memory for an event, and our work suggests that it is (Foley et al., 1983; Johnson, Raye, et al., 1979), then this information might well form the basis for clustering items in recall. Our discrimination data also suggest that the cues for differentiating the self from others are available sooner than those for differentiating among different types of self-generations (e.g., performed vs. imagined actions). This led us to expect that children would be less likely than adults to cluster items by Do and Pretend categories.

Clustering scores were based on the number of times an action performed or imagined was followed in recall by an action of the same type. Intrusions were very few and therefore ignored in counting repetitions. These repetitions were then used to compute the Ratio of Repetition (RR) (Bousfield & Bousfield, 1966); the means are shown in Table 2. Clustering exceeded chance level in all conditions; the lowest obtained, $t(11) = 4.44$, $p < .01$, exceeded the critical value (2.20).

[4] The number of subjects who recalled by performing actions rather than by naming the actions was negligible; therefore mode of recall was not included as a variable in the analyses.

TABLE 2

MEAN PROPORTION CLUSTERING RATIO OF REPETITION (RR) BY ITEM TYPE FOR EACH AGE GROUP, EXPERIMENT 1

	AGE OF SUBJECTS		
CONDITION	6 Years	9 Years	Adult
Do Watch	.65	.65	.62
Watch Watch	.44	.33	.48
Do Pretend	.50	.40	.52

In a 3 (age) × 3 (condition) ANOVA on the RR scores, there were no age differences in clustering. However, there were large differences among conditions, $F(2,99) = 9.96$, $MS_e = .05$, $p < .001$ (see Table 2). Clustering was higher in the Do-Watch condition ($M = .64$), as predicted, than in Watch-Watch ($M = .43$) or Do-Pretend ($M = .47$) conditions, respectively, $F(1,99) = 60.20$, and $F(1,99) = 39.33$, respectively. Clustering levels in these latter two conditions were equivalent.

There is another way in which subjects might have clustered their recall, and that is by action class. New RR scores were computed to assess organization in recall by the six action classes in Table 1. These scores are shown in the top half of Table 3.

Single-sample t tests indicated clustering exceeded chance (.22) in all but two cases: the 6-year-olds in the Do-Watch ($M = .16$) and the Watch-Watch ($M = .26$) conditions. A 3 (age) × 3 (condition) ANOVA showed that clustering increased with age, $F(2,99) = 5.81$, $MS_e = .03$, $p < .004$; 6-year-olds clustered less ($M = .26$) than did 9-year-olds ($M = .36$) and adults ($M = .39$). Clustering also varied with condition, $F(2,99) = 3.71$, $p < .03$; clustering was less in the Do-Watch condition ($M = .28$) than in the Watch-Watch ($M = .33$) and the Do-Pretend ($M = .34$) conditions.[5]

To summarize, the results of Experiment 1 clearly indicate that the findings reported by Foley et al. (1983) generalize to a wide range of actions. Children and adults did not differ in their ability to disciminate what they did from what someone else did or in their ability to discriminate between what two other people did. However, the children did have trouble distinguishing what they themselves did from what they only imagined doing.

TABLE 3

MEAN PROPORTION CLUSTERING (RR) BY ACTION CLASS

	AGE OF SUBJECTS		
CONDITION	6 Years	9 Years	Adult
Experiment 1:			
Do Watch	.16	.34	.34
Watch Watch ..	.26	.36	.38
Do Pretend ...	.36	.34	.34
Experiment 2:			
Watch All	.52	.51	.47
Do All	.48	.48	.47

For the most part, the clustering data presented a picture consistent with the discrimination scores. The organization of recall by person categories was as great for children as for adults; thus both discrimination scores and clustering indicate that the distinction between self and other or between two other people is as salient for 6-year-olds as it is for adults. Because adults' discrimination performance in the Do-Pretend condition was superior to that of the children's, we thought that the adults might be more likely to use the categories Do and Pretend to cluster recall. However, this was not the case; clustering by Do and Pretend did not vary as a function of age. Therefore, it appears that discrimination tests may be better than clustering to index the availability of some discriminative information in memory, that is, that associated with different types of self-generations.

Experiment 2

In the Do-Watch and Watch-Watch conditions, 9-year-olds and adults used both category schemes embedded in the materials (person category and action class) to group their action recall. In contrast, 6-year-olds clustered significantly only by person categories. However, the 6-year-olds did cluster by action class in the Do-Pretend condition, indicating that, under some circumstances, action class is a functional recall category for 6-year-olds. That 6-year-olds used action class to cluster actions only in the Do-Pretend condition suggests that this category scheme was

[5] We computed a second measure of clustering (Adjusted Ratio of Clustering; Roenker, Thompson, & Brown, 1971) because this measure, unlike the RR measure, is not confounded with recall (Murphy, 1979). With this more conservative estimate, essentially the same pattern was obtained. The only difference was that clustering by item type was not above chance in either the Watch-Watch or the Do-Pretend conditions for any age group.

overwhelmed by the more salient person categories in the other conditions (self vs. other or two other people). Experiment 2 examines this possibility.

In a Watch-All condition, subjects watched another person perform all 24 actions used in the activity phase of Experiment 1. The impact of person categories on the degree of clustering by action class was assessed by comparing clustering by action class in the Watch-Watch condition of Experiment 1 to that in the Watch-All condition of Experiment 2. In addition, in a Do-All condition, subjects performed all actions. Clustering by action class was compared in the Do-Pretend condition of Experiment 1 and the Do-All condition of Experiment 2 in order to determine whether the presence of two types of self-generated items influences the use of action class to organize recall.

Method and Procedures

The materials and procedures were the same as those in Experiment 1, and, in fact, these two conditions were run in randomized replications with the conditions in Experiment 1 so that comparisons of recall and clustering could be made between the two studies. Subjects were from the same populations described in Experiment 1, with 12 per cell ($N = 72$).

Results and Discussion

Mean recall is shown in Table 4. A 3 (age) × 2 (condition) × 6 (action class) ANOVA was conducted. Recall varied with action class, $F(5,330) = 15.59$, $p < .001$, but the data in Table 4 collapse across this variable because the effect was comparable to that reported for Experiment 1 (see Fig. 2). Recall varied with age, $F(2,66) = 42.78$, $p < .001$. Scheffé's tests showed that there were no differences between the two younger groups; both recalled fewer actions than adults.

Recall was generally higher for subjects who performed actions ($M = 13.88$) than for those who watched another person perform actions ($M = 11.92$), $F(1,66) = 11.67$, $p < .001$. Furthermore, for this between-subject comparison, the "generation effect" (Slamecka & Graf, 1978) increased with age, as shown in Table 4, $F(2,66) = 3.06$, $p < .05$. Simple main effects showed that adults who performed actions recalled more than adults who did not, $F(1,66) = 10.95$, $p < .05$, and that 9-year-olds who performed actions recalled more than 9-year-olds who did not, $F(1,99) = 4.01$. This difference was not significant for 6-year-olds.

TABLE 4

MEAN RECALL COLLAPSED ACROSS ACTION CLASS, EXPERIMENT 2

	AGE OF SUBJECTS		
CONDITION	6 Years	9 Years	Adult
Do All	10.33	12.75	18.42
Watch All	10.00	11.00	14.75

A 3 (age) × 2 (Do vs. Watch) × 6 (action class) ANOVA was performed on the recall data for subjects in the Do-Watch condition of Experiment 1; this provides a within-subject comparison similar to the between-subject comparison just reported. For this within-subject comparison, recall of actions performed ($M = 6.18$) was greater than recall of actions observed ($M = 5.10$), $F(1,33) = 11.06$, $MS_e = .32$, $p < .002$. There was not an interaction between age and type of item as there was in Experiment 2. However, subsequent comparisons for each age group showed that 6-year-olds did not recall significantly more of the actions they performed than of those they saw another perform (Do-Watch), whereas the 9-year-olds and adults did. The general tendency for the generation effect to increase with age is consistent with that previously reported (Foley et al., 1983; see also Johnson, Perlmutter, & Trabasso, 1979, Experiment 1).

Recall order.—Clustering by action class is shown by the RR scores in the bottom half of Table 3. Clustering exceeded chance in all cases (all obtained $t[11]$ values exceeded the critical value of 2.22, and the lowest obtained t value was 6.25). Unlike Experiment 1, there were no significant differences in clustering by action class among the age groups or between the conditions.

Subjects in the Watch-Watch and Watch-All conditions of Experiments 1 and 2, respectively, were similar in that they observed others performing actions. A 2 (Experiment) × 3 (age) ANOVA was calculated to compare the tendency to cluster by action class in these two conditions. Clustering by action class was less in the Watch-Watch condition ($M = .33$) than in the Watch-All condition ($M = .50$), $F(1,66) = 14.22$, $MS_e = .26$, $p < .001$, and this difference was greater the younger the subjects, $F(2,66) = 4.26$ (see Table 4). Thus, the 6-year-olds in the Watch-Watch condition only used person categories for clustering. Either they were less likely to notice the action classes when person catego-

ries were available, or they may have had trouble using two categorical schemes at once. In either event, the results show that, when both types of categories are represented in the materials, person categories are more salient than action classes for the younger subjects.

The clustering by action class in the Do-Pretend and Do-All conditions of Experiments 1 and 2, respectively, were also compared. Here, the 2 (Experiment) × 3 (age) ANOVA showed no significant effects, indicating that cues based on two types of self-generations (Do and Pretend) did not significantly affect the tendency to use action class. This lack of competition from the categories Do and Pretend is another indication (along with lower discrimination scores) of the lack of salience of these categories.

General Discussion

We have compared children's confusion about memories with that of adults, and our work indicates that how confused children are about the origin of memories depends on the nature of the discrimination called for. Children as young as 6 were at no disadvantage in discriminating between perceptual and self-generated memories (what subjects saw another do vs. what subjects did). The Reality Monitoring model proposes that differences in the representations of perceptual and self-generated memories (e.g., greater detail vs. more information about cognitive operations and motor programs, respectively) facilitate decisions about the origin of memories. Evidently, children, as well as adults, can draw on these differences in the memory representations of internal and external events to make decisions about the origin of memories. The fact that all groups ordered recall according to who did what in the Do-Watch condition and that the magnitude of clustering was the same for all age groups is independent support for the notion that self and other are well-differentiated classes of experience for young children. Furthermore, these categories appear to be especially salient for 6-year-olds because the children used them to cluster their recall to the exclusion of another potentially functional category (action class).

The present findings, in combination with earlier developmental work (Foley et al., 1983; Johnson, Raye, et al., 1979), show that young children distinguish between memories for perceptions and self-generations across a wide range of situations involving self-generated memories (speech, pictorial materials, and simple actions). That preschoolers can take into account the characteristics of the person with whom they are speaking (Flavell, Flavell, Green, & Wilcox, 1981; Schmidt & Paris, 1984) is further support for the idea that young children differentiate themselves from others (see also Bretherton & Beeghly, 1982; Johnson & Wellman, 1982).

Children as young as 6 were also as able as older subjects to separate memories based on two different perceptual sources (what two other people did). Within the Reality Monitoring framework, distinguishing between which of two people performed particular actions involves specific information—such cues as perceptual information associated with how each person looked and where they were located when they performed the actions. The fact that children performed as well as adults in the Watch-Watch condition further indicates (along with the children's good performance in the Do-Watch condition) that children can draw on differences in qualities of memories to make attributions about origin.

However, children sometimes are more confused than adults are about the origin of memories. In Experiment 1, they had difficulty distinguishing what they did from what they imagined doing. Thus the findings reported previously (Foley et al., 1983) are not limited to speech but extend to a range of actions in general.

In the Do-Pretend condition (as in the Watch-Watch condition), distinguishing performed and imagined actions involves specific information, such as cues associated with different aspects of self-generated memories (e.g., their initiation, production, and consequences). Apparently, since discrimination was better in the Watch-Watch condition than in the Do-Pretend condition, specific perceptual cues associated with the actions of different people are available sooner than specific cues associated with self-generations.

Interesting questions are raised by these studies. In contrast to our previous findings with verbal materials, 9-year-olds as well as 6-year-olds were at a disadvantage relative to adults in discriminating performed and imagined actions. This could simply reflect fluctuations in samples from an age group in transition. On the other hand, the fact that 9-year-olds did not have any particular trouble separating acts from imaginations with verbal materials, while they did have trouble with the actions investigated here, may point

to something more interesting about changes in the nature of memories or of judgment processes with age that are specific to types of materials. For example, suppose that there is a general tendency, with age, for imagination to become more schematic. As suggested in the introduction, young children may subvocalize as they imagine words. Subvocalization presumably drops out, and people imagine themselves saying words via an abbreviated, or more abstractly represented, imaginal event. One interpretation of the performance of the 9-year-olds in the Say-Think condition in the Foley et al. (1983) studies is that 9-year-olds generate more abstract, "adult-like" schematic representations of words. In general, children may tend to imaginally represent nonspeech actions schematically later than they do speech actions, perhaps because internal speech is highly practiced by the age of 9.

Another interesting question is whether the difficulty in the Do-Pretend condition indicates that children have trouble separating memories based on any type of self-generations (e.g., what I did vs. what I imagined; what I wrote vs. what I traced). If the degree of confusion varies with the type of self-generated acts being discriminated, then it should be possible to determine the particular aspects of memories children have trouble with (e.g., cues associated with initiation, execution, or consequences following self-generations). Finally, if children's confusion is specifically related to memories for performed and imagined actions, it would also be important to determine if this is because both doing and imagining directly involve the self. If the involvement of the self is critical, then children ought to be less confused about memories for what they did and what they thought of someone else doing.

References

Anderson, R. (1985). Did I do it or did I imagine doing it? *Journal of Experimental Psychology: General*, **113**, 594–615.

Bousfield, A. K., & Bousfield, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, **19**, 935–942.

Bransford, J., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383–438). New York: Academic Press.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, **18**, 906–921.

Broughton, J. (1978). The development of the concepts self, mind, reality and knowledge. *New Directions in Child Development*, **1**, 75–100.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering and understanding. In J. H. Flavell & E. M. Markman (Eds.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol 3. Cognitive development* (pp. 77–166). New York: Wiley.

Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, **15**, 95–120.

Flavell, J. H., Flavell, E. R., Green, F. L., & Wilcox, S. A. (1981). The development of three spatial perspective-taking rules. *Child Development*, **52**, 356–358.

Foellinger, D. B., & Trabasso, T. (1977). Seeing, hearing, and doing: A developmental study of memory for actions. *Child Development*, **48**, 1482–1489.

Foley, M. A., & Johnson, M. K. (April, 1982). *The consequences for memory of imagining in another person's voice.* Paper presented at the meeting of the Eastern Psychological Association, Baltimore.

Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for speech and memories for thought. *Child Development*, **54**, 51–60.

Foulkes, D. (1982). *Children's dreams.* New York: Wiley.

Freud, S. (1953). *Beyond the pleasure principle.* In R. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. **18**). London: Hogarth. (Original work published 1922).

Garrity, L. I. (1977). Electromyography: A review of the current status of subvocal speech research. *Memory and Cognition*, **5**, 615–622.

Johnson, C. N., & Wellman, H. W. (1982). Children's developing conception of the mind and brain. *Child Development*, **53**, 222–234.

Johnson, L. R., Perlmutter, M., & Trabasso, T. (1979). The leg bone is connected to the knee bone: Children's representation of body parts in memory, drawing, and language. *Child Development*, **50**, 1192–1202.

Johnson, M. K. (1985). The origin of memories. In P. C. Kendall (Ed.), *Advances in cognitive-behaviorial research therapy* (Vol. **4**, pp. 1–27). New York: Academic Press.

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology*, **98**, 203–205.

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, **40**, 33–50.

Johnson, M. K., Kahan, T. L., & Raye, C. L. (1984). Dreams and reality monitoring. *Journal of Experimental Psychology: General*, **113**, 329–344.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67–85.

Johnson, M. K., Raye, C. L., Hasher, L., & Chromiak, W. (1979). Are there developmental differences in reality monitoring? *Journal of Experimental Child Psychology*, **27**, 120–128.

Keppel, G. (1984). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Kirk, R. E. (1968). *Experimental design procedures for the behavioral sciences*. Belmont, CA: Brooks Cole.

Lieberman, Z., & Altschul, U. (1971). Memory for a list of commands: Imagining, seeing and doing. *Perceptual and Motor Skills*, **31**, 530.

MacKay, D. G. (1981). The problem of rehearsal or mental practice. *Journal of Motor Behavior*, **13**, 274–285.

Minas, S. (1977). Memory coding for movement. *Perceptual and Motor Skills*, **45**, 787–790.

Morison, P., & Gardner, H. (1978). Dragons and dinosaurs: The child's capacity to differentiate fantasy from reality. *Child Development*, **49**, 642–648.

Murphy, M. D. (1979). Measurement of category clustering in free recall. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 51–53). New York: Academic Press.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, **88**, 1–15.

Piaget, J. (1929). *The child's conception of the world*. New York: Harcourt, Brace.

Piaget, J. (1959). *Judgment and reasoning in the child*. Totowa, NJ: Littlefield, Adams.

Piaget, J. (1962). *Plays, dreams and imitation in childhood*. New York: Norton.

Reason, J., & Mycielska, K. (1982). *Absent-minded?* Englewood Cliffs, NJ: Prentice-Hall.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures of clustering and free recall. *Psychonomic Bulletin*, **76**, 45–48.

Schmidt, C. R., & Paris, S. G. (1984). The development of verbal communication skills in children. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 2–47). New York: Academic Press.

Selman, R. L. (1981). What children understand of intrapsychic processes: The child as a budding personality theorist. In E. K. Shapiro & E. Weber (Eds.), *Cognitive and affective growth: A developmental interaction* (pp. 187–215). Hillsdale, NJ: LEA.

Slamecka, N. J., & Graf, P. (1978). The generation effect: The delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, **1**, 592–604.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Werner, H. (1948). *Comparative psychology of mental development*. New York: International Universities Press.