Age-related Changes in Confusion between Memories for Thoughts and Memories for Speech

Mary Ann Foley and Marcia K. Johnson
State University of New York at Stony Brook

Carol L. Raye
Bell Laboratories

FOLEY, MARY ANN; JOHNSON, MARCIA K.; and RAYE, CAROL L. *Age-related Changes in Confusion between Memories for Thoughts and Memories for Speech.* CHILD DEVELOPMENT, 1983, **54,** 51–60. The present experiments compared people's abilities to make decisions about the origin of their memories. Experiment 1 demonstrated that 6-year-olds were as good as 17-year-olds in discriminating memories originating from what they said earlier (self-generations) from memories of what another person said earlier (external presentations). However, in both experiments 1 and 2, 6-year-olds were not as good at discriminating what they had said earlier from what they had only thought. The possibility that younger children simply have more difficulty distinguishing between memories originating from the same class, internal or external, was ruled out because 6-year-olds performed as well as 9-year-olds when differentiating between memories from 2 external sources (experiment 2). Nor could their difficulty be attributed to a general problem in distinguishing memories for their thoughts from *any* other class of memories because they were at no disadvantage in discriminating their earlier thoughts (words they imagined themselves saying) from words someone else said (experiment 2). Our findings suggest that some distinctions, self versus other, emerge as cues in memory sooner than other distinctions, thoughts versus actions.

Asking people to remember whether they said something or somebody else said it is an instance of what Johnson and Raye call "reality monitoring" (1981; Note 1). More generally, reality monitoring refers to the process of discriminating between self-generated memories and memories produced by perceptual processing of external events. Few experiments have directly explored whether this ability improves with age, in spite of the theoretical and practical importance of this issue (e.g., the question of the reliability of children's legal testimony). One exception is a study by Johnson, Raye, Hasher, and Chromiak (1979) that found no developmental differences in reality monitoring. Eight-year-olds were as good as adults at differentiating memories derived from imagining pictures and memories derived from seeing pictures. Thus, while it might be expected that this ability would improve with age (Piaget 1929, 1959), there was no evidence to support this notion.

The present set of studies not only tested whether the prior findings would generalize to memories for verbal information, but they also used a paradigm that allowed us to investigate reality monitoring in even younger children, 6-year-olds. In addition we looked at the development of another ability, distinguishing memories of what one said from what one only thought.

According to Piaget, the process of thinking undergoes substantial developmental change

This research was supported by NSF grant BNS7813054A02. Some of the present data were initially presented at the annual meeting of the Eastern Psychological Association, Hartford, Connecticut, 1980. We thank Mark Aronoff and Hugh Foley for their helpful suggestions and comments on earlier versions of this manuscript. In addition, we express our thanks to Charlene Reeder, Kathleen McNelis, Nadine Abbott, and Lisa Gershoff for their generous help in conducting experiments 1 and 2. We especially thank Sister Joan Leavey, D.W., principal, and Father F. Pizzarrelli, administrator, and the teachers and children in the school in which this work was completed. Requests for reprints should be sent to Marcia K. Johnson, Department of Psychology, State University of New York, Stony Brook, New York 11794.

[*Child Development*, 1983, **54,** 51–60. 0009-3920/83/5401-0022$01.00]

and, for children as young as 6, thinking is equivalent to taking (e.g., "thinking is with the mouth" [Piaget 1929, p. 39]). Also, he proposed that it is not until children have reached the stage of formal operations, beginning about 11–12 years of age, that they are able to reflect on their own thought processes (see also Flavell 1977). If young children are less able to differentiate between *ongoing* self-initiated activities such as speech and thought, then one would expect that *memories* originating from these two types of self-initiated activities would be particularly confusable. Thus, from this point of view, children might have more difficulty than adults in differentiating between memories of their own speech and thoughts, as well as more trouble distinguishing what they said from what someone else said.

Another purpose of the present studies was to test a particular prediction of the Johnson and Raye model. In describing the reality monitoring process, Johnson and Raye proposed that the two classes of memories (those originating from internal events and those originating from external events) differ on several general dimensions. These class differences (e.g., level of sensory information, information about cognitive operations) are then used to make reality monitoring decisions, such as distinguishing between what you said and what someone else said. Of course, on the average, any two memories from the same class should be more similar on these general dimensions than any two memories from different classes, and it should therefore be harder to distinguish what one thought earlier from what one said than to distinguish what one heard from what one said.

Experiment 1

Method

Design.—The manipulated variable was the type of discrimination required. In the say-listen condition, two adult experimenters were present in the room with the subject. One experimenter asked the subject to pronounce words out loud and to listen while the other experimenter pronounced different words out loud. In the say-think condition, the experimenter asked the subject to pronounce some words out loud and to simply think of saying other words. In this second condition, then, the subject engaged in two types of self-generated activities—one overt and the other covert. Subjects were selected from each of three age levels—6, 9, and 17 years of age. There were eight subjects in each combination of age and treatment condition ($N = 48$).

Subjects.—The 6- and 9-year-olds were enrolled in a summer recreational program sponsored by a suburban Catholic parish. The mean ages of these children were 6-9 and 9-8. The 17-year-olds were counselors in the same summer program and, because they were members of the community from which the children were drawn, the three age groups were similar with respect to socioeconomic backgrounds. Males and females were represented equivalently across age levels and conditions. All children were given parental consent to participate.

Materials.—Sixty words representing common objects familiar to young children were selected from children's storybooks (e.g., "cookie," "puddle," "straw," "truck," "mother," "flower"). Half of the words were randomly designated as new words (distractors) for a later recognition test. The other 30 (target) words were randomly assigned to one of two types of items within each condition. In the say-listen condition, 15 target words were designated as items the subject would pronounce (S items) and 15 were designated as items the experimenter would pronounce (E items). In the say-think condition 15 target words were designated as those that the subject would express overtly (S items) and the other 15 were designated as those the subject would generate covertly (T items). Across subjects, target items occurred equally often as S or E items (or S or T items). The 30 target items were presented in random order. Both types of items, S or E(T), occurred equally often in each quarter of the acquisition list.

On the recognition test the order of target and new items was determined by randomly assigning items to the 60 positions with the restriction that different types of targets from different portions of the acquisition phase be distributed as similarly as possible throughout the recognition test sequence.

Procedure.—All subjects were tested individually by a female experimenter in one of two unused rooms at the school sponsoring the recreational program. A 6- or 9-year-old child was selected randomly and then randomly assigned to the say-listen or say-think condition. Similarly, counselors were assigned randomly to one of the two conditions as they signed up for appointments, usually scheduled before or after the recreational program each day. The counselors were tested in the same rooms as the children.

The child was met by a female experimenter and accompanied to the experimental

room. The experimenter talked to the child to help the child feel comfortable, and, if a child was in the say-listen condition, the other experimenter was introduced, and the three engaged in conversation. After the experiment was over, the experimenter accompanied the child back to the classroom.

Children were told that they were about to play a detective game. A child in the say-listen condition was told that good detectives listen carefully for clues when other people speak. "Sometimes I will ask you [John] to say some words, and other times I will ask [Kathy] to say other words." The experimenter then cued the child and the experimenter by saying, for example, "John, say crayon." Four practice trials were then presented.

A child in the say-think condition was told that good detectives are careful not to give away clues about what they are thinking. "Sometimes I will ask you to say a word out loud, and other times I will ask you to think of yourself saying a word out loud." When asked to "think" of saying a word, a child was told to imagine themselves actually saying the word to themselves. After four practice trials, a child in this condition was then asked "What did you do when I asked you to think of yourself saying cat?" If the child responded incorrectly (e.g., "I thought of my cat" or "I thought of a cat"), then the instructions were repeated to emphasize that the child should actually think of saying the word. There were very few instances in which these instructions needed to be repeated.[1] No subjects were unable to follow the instructions. Once the acquisition phase began, subjects rarely said a word aloud when they were supposed to think of it, and the frequency of this type of error was the same for all age groups. In addition, children were good detectives in that they did not engage in overt lip movements when asked to think of themselves saying words.

After completion of the acquisition phase, the experimenter(s) engaged in conversation with the child for about 3 min and then proceeded with the test phase. Here, the child was told that "good detectives remember" and that we wanted to see if the child was a good detective. This was the first mention of a memory test, and subjects of all age groups often expressed surprise at this point. During the test, both experimenters remained in the room in the say-listen condition. Children in the say-listen condition were asked to decide whether they said a word, the other experimenter said a word, or the word was a new word. Children in the say-think condition were asked to decide whether each word was one they said out loud, one they imagined themselves saying, or a new word. The experimenter read each test item aloud. All of the children seemed to enjoy the game. This technique also made it easy to encourage children to "keep [the game] a secret, because good detectives keep secrets," in order to discourage discussions about the game among the children in the program. The instructions were written initially for the youngest age group and modified appropriately for the older age groups; for example, references to detectives were deleted for the oldest age group.

Results and Discussion

Discrimination scores.—For each subject in the say-listen condition, the total number of words attributed to the correct source (i.e., the number of words correctly identified as items the subject said, plus the number of words correctly identified as items the experimenter said) was divided by the total number of items correctly identified as old. For each subject in the say-think condition, the number of words correctly identified as "say words," plus the number of words correctly identified as "think words," was divided by the total number of words correctly identified as old.

As the means in table 1 suggest, there was an interaction between age and condition, $F(2,42) = 15.69$, $p < .001$. There were no differences between the say-listen and say-think conditions for the 9-year-olds, $F(1,14) = 1.74$. Discrimination scores were higher for both the

TABLE 1

DISCRIMINATION SCORES, EXPERIMENT 1

	AGE OF SUBJECTS (Years)		
CONDITION	6	9	17
Say-listen.....	.92	.73	.89
Say-think.....	.63	.82	.74

NOTE.—Discrimination scores are total items attributed to the correct category divided by total items correctly recognized as old.

[1] When the instructions needed to be repeated in experiment 1 it was usually because a child younger than 6 was brought to the session by mistake. Though these data were excluded from analysis, it is important to note that these younger children had a very difficult time suppressing overt responses when asked to imagine speaking. Also, they had a hard time with the three-choice recognition test.

6- and 17-year-olds in the say-listen compared with the say-think condition, $F(1,14) = 35.44$, $p < .001$, and $F(1,14) = 17.77$, $p < .001$, respectively. Moreover, this difference was greater for the 6-year-olds than for the 17-year-olds, $F(1,28) = 4.90$, $p < .03$. This developmental difference indicates that the youngest subjects had particular trouble discriminating what they said from what they thought. At the same time, the 6-year-olds were as good as the 17-year-olds in distinguishing what they said from what the other speaker said, $F < 1$, analogous to the finding reported by Johnson et al. (1979), in which 8-year-olds were as good as adults in distinguishing what they thought from what they saw.

Recognition memory.—An analysis of the recognition of old and new items, without regard for the correct identification of the origin of old items, provides an overall index of recognition performance. The total number of errors is given by the number of items that the subject failed to recognize (misses) plus the number of items that the subject mistakenly identified as old (false positives). This particular measure takes into account decision criteria and correlates highly with d' (Underwood 1974). The mean number of errors are summarized in table 2.

There was a significant interaction between age and condition, $F(2,42) = 8.02$, $p < .001$, but no main effects. As can be seen in table 2, recognition performance was about equal in the two conditions for the 6-year-olds, was better in the say-think condition for the 9-year-olds, and, for the 17-year-olds, performance in the say-listen condition was superior to performance in the say-think condition. A comparison of recognition errors categorized by type of item will be discussed after experiment 2 has been presented.

In general, recognition scores did not match the pattern of discrimination scores. Furthermore, in no condition and at no age was there a significant correlation between the discrimination scores and recognition errors. These results are consistent with the suggestion that these two memory tests capture different aspects of memory (Johnson & Raye 1981).

False positives.—Previously we have shown that adults tend to attribute a new item to the other person rather than to themselves when they mistakenly identify a new item as old in situations similar to the say-listen condition in the present study (Johnson, Raye, Foley, & Foley 1981). Table 3 shows the mean number of false positives of each type for each condition for the present experiment. In the say-listen condition, there was a main effect of type of false positive, $F(1,21) = 20.11$, $p < .001$, a main effect of age, $F(2,21) = 5.64$, $p < .01$, and an age × condition interaction, $F(2,21) = 5.62$, $p < .01$. There was a general bias to say a new item was said by the other speaker rather than by oneself, and this bias was particularly pronounced for the 9-year-olds.

The false positives in the say-think condition demonstrate a different bias: young adults exhibited a tendency to misidentify new items as ones they thought rather than said, $F(1,7) = 9.21$, $p < .05$. Children, on the other hand, did not show this tendency in the least.

Experiment 2

From a developmental perspective, the primary finding from experiment 1 was the particularly poor discrimination performance of the 6-year-olds in the say-think condition. At the same time, 6-year-olds were at no disadvantage compared with older subjects in the say-listen condition. These findings suggest that the use of cues for differentiating the self from others may develop sooner than the use of cues for differentiating self-initiated activities such as speech and thought.

We should, however, consider some alternative explanations. First, young children may

TABLE 2

MEAN NUMBER OF RECOGNITION ERRORS (Misses Plus False Positives), EXPERIMENT 1

	AGE OF SUBJECTS (Years)		
CONDITION	6	9	17
Say-listen	10.00	11.75	6.00
Say-think	9.50	6.88	10.38

TABLE 3

MEAN NUMBER OF FALSE POSITIVES, EXPERIMENT 1

	AGE OF SUBJECTS (Years)		
	6	9	17
Say-listen condition:			
New items called S	.63	1.25	.12
New items called E	1.75	4.90	.75
Say-think condition:			
New items called S	.38	.38	.50
New items called T	.38	.38	4.13

simply have a difficult time discriminating between any two memories originating from the same class of experience. Thus, a new condition, listen-listen, was added in which subjects listened while the experimenter asked two adults to pronounce words out loud. Here the two sources of memories are from the same class (external) and, if same-class membership creates the problem for young children, they should do poorly in the listen-listen condition as well as in the say-think condition. In another new condition, think-listen, the subjects thought of themselves saying some words and listened while another person said other words. This condition addressed the possibility that young children might generally have difficulty discriminating memories derived from thoughts from any class of memories rather than from their own speech in particular.

Design and subjects.—As in the first study, the major manipulated variable was the discrimination called for, say-listen, say-think, listen-listen, and think-listen. These four independent conditions were combined factorially with two age levels, 6 and 9 years, with eight children per cell ($N = 64$). The children were enrolled in a local parochial school, and their mean ages were 6-11 and 9-6. Male and female children were represented proportionally across conditions. All children were given parental consent for their participation.

Procedure.—The materials and basic procedures used were identical to those included in experiment 1. Items were counterbalanced across all conditions in a manner identical to that described for experiment 1. Children were randomly assigned to one of the four conditions and tested individually by one of two female experimenters in a reading room at the school. The two female experimenters were represented equivalently across age and condition.

Children in the say-listen and say-think conditions were given the same instructions as the children in experiment 1. Children in the listen-listen condition were told that good detectives listen carefully to what other people say. "Sometimes [John] I will ask [Nancy] to say some words, and other times I will ask [Lisa]." Four practice trials were then given.

Children in the think-listen condition were told that good detectives are careful not to give away any clues about what they are thinking. "Sometimes [Susan], I will ask you to imagine yourself saying some words, and other times I will ask [Nancy] to say some words." After four practice trials, children in this condition were asked what they did when the experimenter asked them to imagine themselves saying words. There were very few instances in which instructions had to be repeated because children did not understand what they were asked to do. Also, no children had to be excluded because of an inability to follow instructions in any of the conditions.

After completion of the acquisition phase, the experimenter(s) engaged in brief conversation with the child and then proceeded with the test phase. Children in the say-listen and say-think conditions were given instructions as in experiment 1. In the listen-listen condition, children decided whether each word was a new word, or who said each word (e.g., Nancy or Lisa). Finally, children in the think-listen condition were asked to decide whether they thought of themselves saying a word, heard the other experimenter say a word, or the item was a new word. All of the experimenters remained in the room during the test phase, where appropriate. Children were permitted to indicate their responses either verbally or by pointing.

Results and Discussion

Discrimination scores.—Table 4 presents discrimination scores for the various conditions. The analyses consisted of several planned comparisons addressing the questions outlined earlier.

First, discrimination performance in the say-listen and say-think conditions was compared. Our findings in experiment 1 were replicated in that the scores were higher in the say-listen than in the say-think condition, $F(1,28) = 20.13$, $p < .001$, and the 6-year-olds had a particularly difficult time in discriminating between items in the say-think condition compared with the say-listen condition, $F(1,14) = 21.66$, $p < .001$. Here, the 9-year-olds showed the expected pattern in that the discrimination scores were higher in the say-

TABLE 4

DISCRIMINATION SCORES, EXPERIMENT 2

CONDITION	AGE OF SUBJECTS (Years)	
	6	9
Say-listen	.90	.85
Say-think	.63	.74
Listen-listen	.79	.71
Think-listen	.81	.82

listen than in the say-think condition, but, as in the first study, this difference between conditions was not significant for the 9-year-olds; hence the prediction from the reality monitoring model that say-listen would be easier than say-think was not supported for 9-year-olds in either experiment but was for 6-year-olds in both experiments and 17-year-olds in experiment 1. Also, a finding consistent with this prediction was reported by Anderson (Note 2) where, in a forced-choice test, subjects could distinguish a drawing they had traced from one they had merely looked at (in the present framework, a discrimination between an internally generated and an external event) better than they could distinguish a drawing they had traced from one they had only imagined tracing (two internally generated events).

If the poor performances of the 6-year-olds in the say-think condition in experiments 1 and 2 were simply because younger children have a particularly hard time discriminating between memories that originated from the same class, then the 6-year-olds should also have particular difficulty discriminating between memories originating from two external sources. Therefore, we compared the listen-listen and say-think conditions next. As is clear from table 4, this was not the case. Performance of the 6- and 9-year-olds did not differ significantly in the listen-listen conditions. However, in contrast to the 9-year-olds, the scores for the 6-year-olds were much lower in the say-think than in the listen-listen condition, $F(1,14) = 5.14$, $p < .05$. This pattern produced the overall significant interaction between age and condition, $F(1,28) = 4.68$, $p < .04$.

When the say-listen and think-listen conditions were compared, there were no differences in discrimination performance. Thus, memory for origin was quite good when subjects' generations were only covert, and, more importantly, 6-year-olds did not appear to find it particularly difficult to separate memories of their thoughts from externally derived memories.

An additional comparison was made between the discrimination scores in the say-listen and listen-listen conditions, because the Johnson and Raye model (1981) predicts a difference between these two conditions. Scores were higher in the say-listen condition, $F(1,28) = 8.39$, $p < .007$, and this difference did not interact with age. Raye and Johnson (1980) reported a similar finding with adults. According to the model, the greater difficulty of the listen-listen condition is because subjects have to rely on *specific* sensory attributes of their experiences such as the speaker's voices or contextual attributes such as where the speakers were seated in the room to help them differentiate among their memories for who said what. In contrast, in the say-listen condition subjects can take advantage of the several general dimensions on which the classes of internally and externally derived memories differ, such as, self-generated memories usually are richer in information about cognitive operations that occurred when the memory traces were established.

Recognition memory.—In an overall analysis there were no significant differences in the recognition errors (misses plus false positives). As in experiment 1, there were no significant correlations between the overall number of recognition errors and the discrimination scores for any age or condition. The recognition errors separated by type will be discussed with comparable data from experiment 1 later.

False positives.—Table 5 shows the mean number of false positives, categorized by type of response for each condition in experiment 2. In the say-listen condition, 6-year-olds were equally likely to attribute new items to themselves or another speaker (as we found in experiment 1), but 9-year-olds were more likely to attribute false positives to the other speaker, producing a significant interaction between age and type of false positive, $F(1,14) = 5.73$, $p < .03$. In the think-listen condition, the bias was in a similar direction but was not significant.

In the listen-listen condition there would be no reason to expect a difference in the number of false positives attributed to the two speakers, and there were no differences for either 6- or 9-year-olds.

TABLE 5

MEAN NUMBER OF FALSE POSITIVES, EXPERIMENT 2

	AGE OF SUBJECTS (Years)	
	6	9
Say-listen condition:		
New items called S.....	.38	.38
New items called E.....	.38	1.12
Think-listen condition:		
New items called T.....	2.12	1.25
New items called E.....	3.37	2.37
Listen-listen condition:		
New items called E_1's...	1.25	2.50
New items called E_2's...	1.12	3.10
Say-think condition:		
New items called S.....	2.00	1.25
New items called T.....	2.62	1.00

In the say-think condition, the number of new items called "thought" items was not greater than the number of new items called "said" items. Thus, in both experiments, if 6- and 9-year-olds misidentified new items as old, there was no bias to report that "I thought it." In contrast, older subjects (experiment 1) demonstrated a pronounced bias to label new items as those they thought rather than ones they said.

In summary, the false positive data from experiment 2 were consistent with those reported from experiment 1 and together suggest that children as young as 9 show an E > S attribution bias when they overtly generate S items but do not yet show a T > S bias. In addition, these data suggest that the developmental pattern of false positives may depend on the decision required. Berch and Evans (1973), for example, found a greater tendency among third graders to misidentify a new item as old compared with kindergarten children and wondered if this effect was specific to the kind of task or type of materials used in their study (i.e., continuous recognition of number pairs). Our data confirm their suspicion that the nature of the developmental differences in the magnitude of false positives is related to the type of decision involved.

Recognition for different types of items.— A secondary though interesting question addressed by these experiments is whether or not recognition memory is dependent upon item types in each of the conditions. While differences across conditions in the overall number of recognition errors were minimal, there is reason to expect errors might vary with item type within each condition. For example, previous work has shown superior memory for self-generated information for adults (Johnson et al. 1981; Slamecka & Graf 1978).

Table 6 shows recognition errors (misses plus false positives divided by two). In both experiments 1 and 2 there were more errors on E items than S items, $F(1,21) = 7.80$, $p < .01$, and $F(1,14) = 25.80$, $p < .001$, respectively. This difference was greater for 9-year-olds than for 6-year-olds in experiment 2, $F(1,14) = 15.35$, $p < .002$, and, when the data for 6- and 9-year-olds from the two experiments were combined, there was also a significant age × item type interaction, $F(1,30) = 8.29$, $p < .007$. The data for the 6-year-olds alone never showed a significant difference in recognition of E and S items. It appears, then, that the benefit of overtly expressing self-generations increases developmentally over the age range of 6–9.

Also, there was an overall main effect for item type in the think-listen condition, $F(1,14) = 13.51$, $p < .002$, with more errors related to E items than to T items. Thus, superior recog-

TABLE 6

RECOGNITION ERRORS SHOWN SEPARATELY FOR EACH TYPE OF ITEM, EXPERIMENTS 1 AND 2

	AGE OF SUBJECTS (Years)		
	6	9	17
Say-listen condition:			
Experiment 1:			
Errors on S-items	4.68	5.19	2.00
Errors on E-items	5.31	6.56	4.06
Experiment 2:			
Errors on S-items	4.06	2.75	...
Errors on E-items	4.80	8.00	...
Think-listen condition:			
Experiment 2:			
Errors on T-items	6.19	3.62	...
Errors on E-items	8.94	8.18	...
Say-think condition:			
Experiment 1:			
Errors on S-items	4.12	4.00	4.81
Errors on T-items	5.38	2.87	5.56
Experiment 2:			
Errors on S-items	7.69	4.50	...
Errors on T-items	8.19	4.25	...
Listen-listen condition:			
Experiment 2:			
Errors on E_1-items	8.19	6.69	...
Errors on E_2-items	6.19	7.44	...

nition memory for self-generations was observed even when items were not expressed overtly. (Johnson et al. [1981] reported similar findings with recall for adult subjects, and we have similar unpublished findings with recognition for adult subjects.)

In the say-think condition of experiments 1 and 2 the main effect of item type was not significant. Thus, subjects were as good at recognizing items simply thought as they were at recognizing items expressed overtly. Similarly, in the listen-listen condition there were no differences in the number of recognition errors related to each speaker.

The superior recognition of T items in the think-listen condition along with the equal recognition of S and T items in the say-think condition support the assumption that the subjects were indeed thinking the words when instructed to do so.

General Discussion

Our findings of no age differences in the ability to discriminate what was said from what was heard (experiments 1 and 2) or what was thought from what was heard (experiment 2) are consistent with those reported by Johnson et al. (1979). Together, these experiments support the idea that cues for differentiating the self from others in memory may be well developed in quite young children. Furthermore, Raye, Johnson, and Taylor (1980) suggested that there are a number of potentially interesting components of memories created from internal or self-generated activities. These components include the residual of central processes leading to the initial activation of a concept and of any further processing (e.g., covert rehearsal), as well as the residual of more peripheral aspects, such as memories of one's own movements in the process of speech production. Logically, any of these components may potentially serve as cues for recognition (Ghatala & Levin 1976; Zechmeister & Gude 1974), recall (Flavell 1977), or origin decisions. For reality monitoring decisions, however, Raye et al. (1980) emphasized the importance of more central aspects. That reality monitoring decisions were as good in our think-listen as in our say-listen condition supports this notion.

However, this very characterization of self-generated as comprised of many types of mental events and overt actions suggests an important class of discrimination problems, for example, differentiating between unrealized ideas and ideas that have been realized or expressed in action (such as vocalization or other activities). Common experiences of this sort of discrimination problem include wondering whether you have said or only thought something, or wondering whether you have mailed a letter or only intended to do so. Such realization judgments, then, may depend on information preserved in memory about motor movements, self-initiated effects on the environment, etc. Our say-think condition required a discrimination between self-generated memories differing in their degree of realization, and here 6-year-olds were at a clear disadvantage relative to older subjects.

There are at least two potential explanations for this age-related pattern. One is that memories originating from speech and thought are more alike for younger children. This would be the case if, for example, younger children were doing more subvocalizing when they imagined themselves saying a word (Garrity 1977). If speech and subvocalized thought produce more similar traces, then young children should have more difficulty discriminating between memories of speech and thought. This explanation suggests that very specific details of the original experience (such as whether subvocalization took place when a concept was activated) are preserved in memory over fairly long intervals.

Alternatively, the developmental differences observed in this study may indicate that younger subjects have less sophisticated reasoning processes. Thus, while memories for speech and thought may differ in their basic characteristics as much for younger subjects as for older subjects, young children may not know how to use cues they have available in memory. This latter interpretation is consistent with the developmental memory literature in its demonstration that there is a discrepancy between young children's acquisition of knowledge and the use of this knowledge for the purpose of monitoring their own memory performance (Brown 1975; Brown & De Loache 1978; Flavell & Wellman 1977; Ornstein 1978). Of course, these two interpretations are not mutually exclusive, and there may be some truth to both of them.

While not conclusive, the false positive data support the second suggestion above, that there are differences across ages in the reasoning applied to the say-think discrimination. These data also demonstrate the potential usefulness of false positives as an index of metamemory assumptions. Intuitions about the way

memory works affect judgments about past events. Children may learn very early that experiences they remember vary in strength. Some memories, for example, are more detailed and vivid than others. Children may also notice very early that what they say or do is more memorable than what they observe others say or do. Such experiences should produce the belief that memories are better for self-initiated activities. How might such a metamemory assumption affect a subject's judgments about items that seem only vaguely familiar? If subjects believe that what they say or do is particularly memorable, and if some event seems only vaguely familiar, then there should be a tendency to attribute the memory to someone else (Johnson et al. 1981). The greater number of false positives attributed to the experimenter, instead of the subjects themselves, is consistent with this presumed belief and is clearly present by 9 years of age.

Similarly, a belief that "thoughts are weaker than deeds" would produce the notable bias in experiment 1 among 17-year-olds to misidentify new items as ones they thought rather than as ones they said (a bias we have subsequently replicated with undergraduates). There was no evidence in the false positive data that such a bias was operating in the judgments of either 6- or 9-year-olds. This indicates that young children may not yet have or apply all the same metamemory rules that older subjects use in say-think discriminations. The false positive data, then, as well as the discrimination scores, indicate that cues that may help one monitor one's memory are not equally used at every period of development and that some distinctions, such as that between self and other, are salient before others, such as that between one's own thoughts and speech. This would be consistent with Piaget's suggestion that speech and thought are undifferentiated for the young child (see also Flavell 1971).

In summary, these experiments found that children as young as 9, as well as adults, showed an advantage in remembering self-generated compared with presented information. This generation effect increased over the age range of 6–9. Also, children as young as 6 failed to show any disadvantage in reality monitoring decisions. However, 6-year-olds did have marked difficulty deciding whether or not their thoughts were realized in audible speech. Thus, our studies suggest there are some grounds for viewing children's memory with suspicion, but not perhaps as much as one might have expected.

Reference Notes

1. Johnson, M. K., & Raye, C. L. A working model of reality monitoring. Annual meeting of the Psychonomic Society, San Antonio, Texas, November 1978.
2. Anderson, R. Did I do it or did I imagine doing it? Annual meeting of the Psychonomic Society, Washington, D.C., November 1977.

References

Berch, D. B., & Evans, R. C. Decision processes in children's recognition memory. *Journal of Experimental Child Psychology,* 1973, **16,** 148–164.

Brown, A. L. The development of memory: knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior.* Vol. **10.** New York: Academic Press, 1975.

Brown, A. L., & De Loache, J. Skills, plans and self-regulation. In R. S. Siegler (Ed.), *Children's thinking: what develops?* Hillsdale, N.J.: Erlbaum, 1978.

Flavell, J. First discussant's comments: what is memory development the development of? *Human Development,* 1971, **14,** 272–278.

Flavell, J. *Cognitive development.* Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Flavell, J., & Wellman, H. M. Metamemory. In R. V. Kail, Jr., & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition.* Hillsdale, N.J.: Erlbaum, 1977.

Garrity, L. I. Electromyography: a review of the current status of subvocal speech research. *Memory & Cognition,* 1977, **5,** 615–622.

Ghatala, E. S., & Levin, J. R. Children's recognition memory processes. In J. R. Levin & V. L. Allen (Eds.), *Cognitive learning in children, theories and strategies.* New York: Academic Press, 1976.

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

Johnson, M. K.; Raye, C. L.; Foley, H. J.; & Foley, M. A. Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology,* 1981, **94,** 37–64.

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

Ornstein, P. (Ed.). *Memory development in children.* Hillsdale, N.J.: Erlbaum, 1978.

Piaget, J. *The child's conception of reality.* London: Routledge & Kegan Paul, 1929.

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

Raye, C. L., & Johnson, M. K. Reality monitoring vs. discrimination between external sources. *Bulletin of the Psychonomic Society,* 1980, **15,** 405–408.

Raye, C. L.; Johnson, M. K.; & Taylor, T. H. Is there something special about memory for internally-generated information? *Memory & Cognition,* 1980, **8,** 141–148.

Slamecka, N. J., & Graf, P. The generation effect: delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, **4,** 592–604.

Underwood, B. J. The role of association in recognition memory. *Journal of Experimental Psychology,* 1974, **102,** 917–939.

Zechmeister, E. B., & Gude, C. Instruction effects in recognition memory. *Bulletin of the Psychonomic Society,* 1974, **3,** 13–15.