Goswami, U. (1990). A special link between rhyming skill and the use of orthographic analogies by beginning readers? *Journal of Child Psychology and Psychiatry, 31,* 301–311.
Goswami, U. (1992). *Analogical reasoning in children.* Sussex, England: Lawrence Erlbaum Associates, Inc.
Holyoak, K. J., & Thagard, P. R. (1989). A computational model of analgical problem solving. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 242–266). New York: Cambridge University Press.
Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development, 54,* 84–90.
Hughes, M. (1986). *Children and number.* Oxford, England: Basil Blackwell.
Inhelder, B., & Piaget, J. (1958) *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.
Kedar-Cabelli, S. (1985, August). *Purpose-directed analogy.* Paper presented at the meeting of the Cognitive Science Society, Irvine, CA.
Lunzer, E. A. (1965). Problems of formal reasoning in test situations. *Monographs of the Society for Research in Child Development, 30*(2, Serial No. 100).
Piaget, J. (1950). *The psychology of intelligence.* London: Routledge & Kegan Paul.
Piaget, J. (1962). *Play, dreams and imitation in childhood.* New York: Norton.
Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438–469). New York: Cambridge University Press.
Soloway, E., Lockhead, J., & Clement, J. (1982). Does computer programming enhance problem-solving ability? Some positive evidence on algebra word problems. In R. J. Seidel, R. E. Anderson, & B. Hunter (Eds.), *Computer literacy.* New York: Academic.
Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological Review, 84,* 353–378.
Vosniadou, S. (1987). Children and metaphors. *Child Development, 58,* 870–885.
Vosniadou, S. (1989). Analogy as a mechanism in knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 413–437). New York: Cambridge University Press.
Winner, E. (1988). *The point of words.* Cambridge, MA: Harvard University Press.

METAPHOR AND SYMBOLIC ACTIVITY *10*(4), 309–331

Children's Memory and Metaphorical Interpretation

Valerie F. Reyna
Department of Educational Psychology
University of Arizona

Barbara Kiernan
Department of Speech and Hearing Services
University of Arizona

Assumptions of constructivism, information-processing, and fuzzy-trace theory were tested in two experiments on memory and comprehension of metaphors with 6- and 9-year-old children. In the first experiment, verbatim memory for metaphors was compared with misrecognition of the gist of metaphorical meaning. In the second experiment, children judged metaphorical meaning, including alternative perceptual and psychological interpretations. We found that (a) children misrecognized the gist of metaphorical interpretations, especially after a delay, much like they do for literal gist (e.g., true inferences); (b) contrary to both constructivism and information-processing theory, misrecognizing metaphorical gist was independent of memory for the metaphor itself; (c) true psychological interpretations were misrecognized and judged acceptable more often than any other type at all ages; and (d) contrary to the idea that literal or perceptual interpretations are suppressed to achieve psychological interpretations, acceptances of alternative interpretations were positively dependent. Results were consistent with those obtained for literal language and for numerical information and are explained by fuzzy-trace theory.

Until recently, theorists assumed that memory and comprehension are closely connected (Bjorklund, 1989; Siegler, 1991). According to information-processing theory, comprehension depends on the ability to encode and sustain information in working memory, at least long enough to extract its meaning (Miller, 1956). The notion that memory and comprehension are

Requests for reprints should be sent to Valerie F. Reyna, Department of Educational Psychology, College of Education, University of Arizona, Tucson, AZ 85721. E-mail: vreyna@ccit.arizona.edu.

related is also central to constructivism, which is the belief that understanding shapes memory. Although the roots of constructivism in cognitive development can be traced to Piaget (e.g., Piaget & Inhelder, 1973), interest was renewed in the 1970s by research on language (e.g., Bransford, Barclay, & Franks, 1972; Liben & Posnansky, 1977). Most recently, the relation between memory and comprehension was the focus of research on fuzzy-trace theory (e.g., Ackerman, 1992; Kreindler & Lumsden, 1994; Reyna, 1995, in press; Swanson, Cooney, & Brock, 1993).

In this article, we explore the relation between memory and comprehension, in particular, between the way children remember and the way they comprehend metaphors. The article proceeds as follows: In the first section, we briefly survey memory and comprehension from each of three developmental perspectives: Piagetian theory, information-processing theory, and fuzzy-trace theory. The remainder of the article is divided into two parts—one on memory and the other on comprehension—corresponding to two experiments. The first experiment concerns children's verbatim memory for metaphors and their tendency to misrecognize the gist of metaphorical meanings. The second experiment examines an implication of the first regarding comprehension, namely, that young children entertain multiple interpretations of metaphors, including abstract psychological interpretations. In the final section, we discuss the relevance of our results to such issues as developmental differences in the interpretation of psychological metaphors (e.g., Winner, Rosenstiel, & Gardner, 1976), the relation between literal and metaphorical meanings (e.g., Glucksberg, 1991), and the indeterminacy of metaphorical meaning (e.g., Kittay, 1987; Reyna, in press).

BACKGROUND

Theorists distinguish metaphors from other kinds of figurative language such as irony and divide them into three types: nominal, predicative, and sentential (Reyna, 1986; Winner, 1988). The three types differ depending on whether a noun phrase, a verb phrase, or a sentence is used metaphorically. For example, the entire sentence is used metaphorically when *"The troops marched on"* refers to children stubbornly refusing to cease some behavior. In the experiments we report in this article, we investigated nominal metaphors, as in *"Juliet is the sun."* It should be noted, however, that developmental differences between nominal and predicative metaphors have been observed, indicating that the former are easier to interpret (Reyna, 1985, 1987).

We also differentiate fresh (or novel) from frozen metaphors. Although the *"leg"* of a table may have been a novel expression at one time, once such metaphorical resemblances become frozen, it is unlikely that they are processed in the same way as fresh metaphors (Reyna, 1986). Of course, metaphors differ in the degree to which they are fresh or frozen. We made an effort, however, to sample metaphors at the fresh end of the continuum, at which active reasoning processes are more likely to be observed.

METAPHOR AND COGNITIVE DEVELOPMENT

It is widely presumed that children's cognitive abilities constrain their interpretation of metaphors. The relation between cognitive development and metaphorical interpretation has been studied from two theoretical perspectives, Piagetian and information processing. Each of these perspectives posits general limitations on cognition, either logical or memorial, that purportedly affect the development of metaphorical interpretation.

The interpretation of novel metaphors can be viewed as a reasoning task, on analogy with classical Piagetian tasks such as class inclusion or probability judgment (e.g., Billow, 1975). For Piaget, cognitive growth was centered primarily on the development of logical reasoning abilities—the attainment of so-called operational thought (e.g., Falmagne, 1975). Thought gradually progressed through a series of stages, from prelogical intuition to formal logic in which it became less influenced by concrete perceptual factors and more influenced by abstract conceptual ones (e.g., Siegler, 1991). Although Piagetian stages have been called into question (e.g., Brainerd, 1978), the idea that "understanding metaphor is primarily a logical-analytic task" (Winner, 1988, p. 10) continues to prevail.

Specifically, current theories retain the Piagetian idea that young children are unable to appreciate abstract, as opposed to perceptual, resemblances (e.g., Keil, 1986; Winner, 1988). This inability to reason abstractly is assumed to hamper the interpretation of certain kinds of metaphors. For example, in interpreting the metaphor *"The prison guard had become a hard rock,"* younger children are expected to be more likely to ascribe physical attributes to the prison guard (e.g., he had hard muscles), whereas older children are more likely to ascribe abstract, psychological attributes (e.g., he was mean and unfeeling; Winner et al., 1976). In this view, which is consonant with Piagetian theory, cognition in general and language comprehension in particular are impelled by the development of reasoning.

In Piagetian theory, memory, too, is subordinated to reasoning (Piaget & Inhelder, 1973). Although memory itself was seen as relatively unimportant, its interaction with reasoning was emphasized. According to the general stance that Piaget advocated—constructivism—reasoning shapes the contents of memory. For example, children who were capable of making inferences from a text should be more likely to recall that text with the inferences (as if they had actually been part of the text) compared with children who

lacked inferential ability (e.g., Brown, Smiley, Day, Townsend, & Lawton, 1977). Hence, children's memories were assumed to reflect their level of logical reasoning (e.g., Paris & Lindauer, 1977; Perner & Mansbridge, 1983; Prawatt & Cancelli, 1976).

In contrast to Piagetian theory, information-processing theory stresses memorial limitations. Such limitations have been used to explain developmental differences in performance on a host of complex cognitive tasks (Bjorklund, 1989; Case, 1985; Daneman & Carpenter, 1980; Siegler, 1991; Trabasso, 1977). The common denominator in these accounts is the claim that memory capacity (or resources) influences the expression of higher cognitive abilities. The assumption that memorial limitations ought to influence metaphorical and other nonliteral interpretations was then incorporated into current models (e.g., Demorest, Silberstein, Gardner, & Winner, 1983; Vosniadou, 1987a; Vosniadou, Ortony, Reynolds, & Wilson, 1984; Vosniadou, Pearson, & Rogers, 1988; Winner et al., 1987). Unfortunately, the evidence on this point is equivocal.

For example, Winner et al. (1987) compared nonliteral interpretation in two conditions that differed in the demands placed on children's memory. They found that although children performed better on one measure of nonliteral comprehension in the condition that minimized memory demands, they performed worse in that condition on the other comprehension measure. Similarly, although children performed better in the low-memory demand condition in comprehending literal falsehoods (i.e., speaker's factual errors), they performed (nonsignificantly) worse in comprehending nonliteral falsehoods (i.e., sarcasm).

Although other studies reported results that are broadly consistent with the memory-limitation hypothesis, research on reasoning (e.g., Reyna, 1992; Reyna & Brainerd, 1990) suggests that such results are subject to two challenges. First, it is crucial to verify that ostensible memory manipulations actually affect memory (see Reyna, 1991, 1995). Rabinowitz, Howe, and Lawrence (1989), for example, also compared high and low "memory load" conditions, but mathematical modeling revealed that processing rather than memory parameters were affected by this manipulation. The implication of this and other research is that inferences about the effect of manipulations on memory can only be made by measuring memory. The second challenge has to do with what such measures of memory reveal. Previous research also indicates that the relation between memory and comprehension cannot be inferred solely by measuring memory. As we discuss later, memory and comprehension must be measured and the relation between them directly assessed.

In summary, theories of cognitive development predict that memory and comprehension of metaphors should be linked. For Piaget, memory should be constructed out of children's understanding of a metaphor's meaning, and

the level of that understanding should change with age. For the information-processing theorist, understanding should depend on the demands placed on a limited memory system, and the ability to handle those demands should also change with age. In the experiments that follow, we present two kinds of data that bear on these hypotheses: (a) false recognition rates for information that was understood but not actually presented, and (b) dependencies between such false recognitions and memory for actually presented information. However, to understand how such data bear on these hypotheses, we must briefly review the literature on memory for literal language.

FALSE RECOGNITION OF UNPRESENTED SENTENCES

Although theorists long maintained that memory is constructive, evidence favoring such claims proved elusive. Bartlett's (1932) early studies are often cited as supportive of constructivism, but his findings were difficult to replicate (e.g., Gauld & Stephenson, 1967). In the 1970s, convincing evidence seemed at hand. In their now classic experiments on constructive memory (e.g., Bransford et al., 1972; Bransford & Franks, 1971), Bransford and colleagues presented participants with literal sentences such as *Three turtles sat on a floating log, and a fish swam beneath it.* After a brief delay, a recognition test was given that contained presented sentences and unpresented sentences consistent with the gist of presented sentences such as *Three turtles sat on a floating log, and a fish swam beneath them,* as well as false sentences. The key result was that participants erroneously claimed that gist-consistent sentences had been presented, sometimes with greater confidence than for sentences that had actually been presented. Paris and Carter (1973) subsequently obtained similar results with children. These false-recognition effects seemed to demonstrate that adults and children remembered what they understood rather than what they experienced.

However, the findings did not go unchallenged (Alba & Hasher, 1983; Fletcher, 1992). For example, Flagg (1976) showed that misrecognition of true inferences, which had been taken to be evidence for constructive memory, could be predicted by a tally model involving memory for verbatim phrases from presented sentences. (For example, given presented sentences such as *The bird is in the cage* and *The cage is under the table,* participants may accept the true inference *The bird is under the table* because it contains phrases that were presented.) Similar false-recognition effects were demonstrated using nonsense syllables, suggesting that understanding had little to do with these phenomena (Reitman & Bower, 1973; Small & Butterworth, 1981). Hasher and Griffin (1978) argued compellingly that there were two parallel, but contradictory, literatures on memory: One claimed that memory was "remarkably accurate—or reproductive" (p. 318) and the other claimed

the opposite (see also Hasher, Attig, & Alba, 1981). This contradiction could not be explained away by differences in materials or other superficial factors.

Therefore, although textbooks routinely characterize memory as constructive, contradictory evidence has been neglected. Hasher and colleagues' work (see also Alba & Hasher, 1983) pointed out these fundamental contradictions, although they were not resolved theoretically. The main issue dividing the two camps was and remains the validity of the false-recognition data. Although there were disputes about the data, observers on both sides accepted the premise that, if the effects were real, memory had to be constructive.

Fuzzy-trace theory introduced a new possibility, namely, that the effects were real but that memory was not constructive. Using fuzzy-trace theory, Reyna and Kiernan (1994) modified instructions and stimuli to address methodological criticisms of earlier sentence recognition experiments. The aim was to obtain unambiguous demonstrations of the classic effects—false recognition of gist—to determine whether they were explained by constructivism. If subjects systematically misrecognized unpresented true sentences, constructivism would have entailed positive dependency between memory judgments for presented sentences and for unpresented sentences consistent with gist. That is, if understanding shapes memory as implied by constructivism, then memory for semantic content must be positively dependent on memory for presented material. Although memory for presented material does not totally determine recognition of semantically consistent sentences, there ought to be some dependency between them.

On the other hand, according to fuzzy-trace theory, two kinds of memory representations of actual experience are deposited—verbatim and gist—that are independent of each other (for a recent review of the theory, see Reyna & Brainerd, 1995). These multiple representations of experience confer cognitive flexibility because reasoners can solve a problem in different ways by using different representations. A central tenet of the theory, however, is that reasoners have a fuzzy-processing preference. That is, reasoning gravitates to the least specified level of representation (i.e., gist) that can be used to accomplish a task (e.g., Reyna & Brainerd, 1991). In most instances, therefore, precisely accurate memory is irrelevant for accurate reasoning. Furthermore, the tendency to rely on gist representations in reasoning increases with development, contrary to traditional cognitive developmental theories (e.g., Brainerd & Reyna, 1993; Reyna & Ellis, 1994). Finally, the tendency to rely on gist representations is modulated by factors affecting the accessibility of gist versus verbatim representations such as instructions, delay, and materials (Reyna, in press; Reyna & Brainerd, 1995; Reyna & Kiernan, 1994).

Under conditions such as those typically used in sentence-recognition experiments (e.g., immediate memory tests), verbatim memory remains accessible and governs recognition of presented sentences. However, systematic

acceptance of new semantically consistent sentences (e.g., true inferences) cannot be based on verbatim memory for the simple reason that these sentences were never presented. Participants sometimes say yes to semantically consistent sentences, however, because they cue memory for the gist of what was presented (e.g., Ackerman, 1992). Thus, under these conditions, recognition of old sentences and misrecognition of new semantically consistent sentences should be independent.

Reyna and Kiernan's (1994) results confirmed the prediction that recognition of old sentences and misrecognition of new semantically consistent sentences were independent. These relations could be manipulated experimentally by varying whether participants were instructed to base judgments on gist or on verbatim memory. Merely instructing participants to base all judgments on gist converted the relation between the same presented sentences and true inferences to positive dependency. Delaying the memory test until verbatim memory was no longer accessible had a similar effect. Both instructions and delay produce positive dependency because participants then use gist memory (rather than verbatim memory) to verify presented sentences and this gist memory is also used to verify true inferences. In addition, younger children, who were less likely to acquire an accurate verbatim representation, were also more likely than were older children to display positive dependency (response biases were not observed for either age group).

Interestingly, both independence and dependency could be elicited within the same task. The verbatim memory test included true paraphrases (i.e., sentences that are synonymous with presented sentences but are worded differently) as well as true inferences. Like true inferences, true paraphrases were misrecognized more often than familiarly worded false sentences. (For example, given presented sentences such as *The orange is bigger than the peach* and *The peach is bigger than the apple*, true paraphrases that contained unfamiliar words such as *The peach is smaller than the orange*, were misrecognized more often than false sentences that contained familiar words such as *The apple is bigger than the peach.*") Although judgments of presented sentences and true inferences were independent, judgments of true paraphrases and of true inferences were dependent. The latter were consistent with the gist of presented sentences, but neither was presented. The first experiment concerns whether results similar to these are obtained for metaphors and their interpretative gist.

EXPERIMENT 1

We applied the techniques used by Reyna and Kiernan (1994) with literal language to study memory for metaphors. In these studies, children received eight vignettes such as:

The woman was shopping in the grocery store.
The woman saw the lost boy near the door.
The woman was an aspirin, kneeling by the lost boy.

Following presentation of each vignette, the children (6- and 9-year-olds) were given a recognition test. They were given the same recognition test again after a delay of 7 days. In the first experiment, children were given verbatim memory tests; they were told to accept only presented sentences. In the second experiment, they were instructed to accept all sentences that were consistent with the meaning of presented sentences. For both experiments, judgments of presented metaphors were compared with those of various metaphorical interpretations and with false sentences.

The main purpose of the verbatim memory test in Experiment 1 was to assess the degree to which children can discriminate between presented metaphors and various potential gists, namely, different metaphorical interpretations. These included sentences incorporating literal synonyms, perceptual interpretations, and psychological interpretations of presented metaphors (e.g., see Winner et al., 1976). These sentences would be classified as true unpresented sentences because unpresented phrases were substituted for presented phrases and because those phrases were semantically consistent with the presented phrases. For example, in the previous vignette, the woman kneeling by the lost boy is described metaphorically as *"an aspirin."* The literal synonym described the woman as *medicine;* the perceptual interpretation referred to her as *round;* and the psychological interpretation indicated that she made the boy feel better. Corresponding false stimuli were not semantically consistent with the presented metaphorical term. These false stimuli were related to the metaphorical term (e.g., *"cough syrup"*) or to the literal synonym (e.g., *nurse*), or they were erroneous perceptual features (e.g., *red cheeks*) or psychological interpretations of the same positive or negative valence as the appropriate psychological interpretation, although they were false (e.g., *proud*), respectively (see Keil, 1986; Nippold, Leonard, & Kail, 1984).

According to most accounts, only the true psychological interpretation represents an apt interpretation of the metaphor. All of the true interpretations, however, are semantically consistent with the metaphor. Therefore, the interesting empirical question regarding verbatim memory tests is whether children misrecognize any of these related true sentences—those related to literal or to metaphorical meanings—as having been presented, as they do for true inferences and other kinds of literal gist.

In addition to examining false-recognition rates for the gist of metaphors, we consider how judgments of the various recognition stimuli are related to one another and whether age, instructions, and delay affect results for these stimuli much like they affect results for other kinds of verbal materials (e.g., Reyna, 1995; Reyna & Kiernan, 1994). Level of verbatim memory can be assessed by comparing acceptance rates for presented sentences with those for semantically consistent, but unpresented, sentences. Such differences (or lack thereof) in acceptance rates between presented and true unpresented sentences (e.g., between *"The woman was an aspirin, kneeling by the lost boy"* and *The woman was medicine, kneeling by the lost boy*) indicate the degree to which children systematically misrecognize the gist of presented sentences.

Method

Participants. A total of 50 children, 25 in a younger group primarily made up of 6-year-olds (M = 6 years, 6 months) and 25 in an older group primarily made up of 9-year-olds (M = 9 years, 7 months), participated in the experiment and one described in Reyna and Kiernan (1994). The children were drawn from an elementary school serving a middle-income residential area, none was identified as learning or language disabled, and there was an equal number of boys and girls.

Design and materials. The design was a 2 × 2 × 2 × 2 × 2 factorial. The first factor, age (6- or 9-years-old), was varied between participants, and the other four factors were varied within participants: Each participant was tested twice: immediately after the story presentation and again after a week's delay. Stories consisted of three sentences: two literal sentences followed by the critical metaphorical sentence. (Consistent with prior usage, we call these connected sets of sentences stories; however, they should be distinguished from full-blown narratives.) The remaining three factors characterized recognition sentences: They were true or false, they were literal sentences or interpretations, and they were closer or further away from the original metaphorical sentence. Closer was operationalized as a closer (as opposed to a more distant) lexical associate for literal sentences and as involving perceptual rather than psychological resemblances for interpretations. In addition to these eight recognition sentences, participants also verified the other two presented sentences.

Procedure. Each child was tested orally and individually. Six different experimenters were used, and they were randomly assigned to children. In addition to securing consent from parents, the child signed an assent form before the session began. After an initial period of familiarization, children were instructed that they would hear some stories and that they should try to remember them. They would then hear some sentences and their task was to say whether they had been in the story. (Sentences were presented and tested orally to guard against such artifacts as differences in reading level.)

Questions relevant to each story were presented immediately after that story, so the child received eight story–test cycles. Children were carefully instructed to say yes only if "it is exactly the same as in the story" and to say no otherwise. They were also told: "Say yes if I say the same words in the same way. Say no if anything is different from the story."

The experimenter then presented a series of examples. Prior to the examples, children were reminded to "try to remember exactly what the story says." Children were given three types of practice recognition sentences: a presented metaphorical sentence, a false sentence, and a true interpretation. The child provided answers for the practice examples, and for each sentence, the experimenter explained the basis for the correct response. Experimenters were trained to repeat the practice examples if children were hesitant or gave incorrect responses, but most children responded quickly and accurately and appeared to have no difficulty understanding the instructions. For the true interpretation, the experimenter acknowledged that the sentence "is true," that this is what happened in the story, but that "the story didn't say that. You should only say yes if I say the same words in the same way." Finally, children were cautioned that each test sentence was different: "Some of these will seem like you heard them before, but you haven't." Stories and recognition sentences were presented in a different random order for each child. Prior to each story, children were reminded to try to remember the exact words.

At the long-term session, children were told that they would be asked about the stories that had been read to them the previous week and that they should answer the questions just like they did then. (Stories were not presented again.) They were also reminded that they "should say yes if I say the same words in the same way, and say no if I say anything different." The order of testing of stories was randomized again, as was the order of recognition sentences, separately for each child. Prior to the recognition sentences for each story, children were again told to "try to remember the exact words."

Results and Discussion

To assess verbatim discrimination for metaphorical sentences, we compared children's acceptance rates for presented metaphors with acceptance rates for true interpretations. We found that children's verbatim memory for metaphors was excellent on the immediate test, although there was some elevation in recognition errors for literal synonyms for both age groups. Note that literal synonyms cannot be rejected by a general rule that eliminates any nonfigurative sentence. Nonetheless, children were generally able to discriminate presented metaphors from highly similar sentences containing literal synonyms, as in *"The woman was medicine, kneeling by the lost boy,"* or the false metaphorical sentence, as in *"The woman was cough syrup, kneeling by the lost boy."*

Over the delay, however, misrecognition of true psychological interpretations increased significantly as did misrecognition of true perceptual interpretations, although the increase was larger for the psychological interpretations. Thus, after the delay, children were more likely to misrecognize a true interpretation of the metaphor as having been presented compared with a false sentence. In this respect, metaphorical interpretations behaved much like true inferences and other literal gist (Reyna & Kiernan, 1994). It is unlikely that the misrecognition of metaphorical gist was due to failure to understand instructions because children had already followed instructions successfully on the previous test, and they were reminded of those instructions prior to the delayed test. Although the level of discrimination between presented sentences and gist sentences was higher in this experiment with metaphors than in earlier experiments with literal sentences, there was a tendency to misrecognize sentences that were semantically related to presented metaphors.

These results were confirmed by a five-factor ANOVA comparing effects of age, delay, and three sentence factors—true–false, literal–interpreted, and close–far. For all significant effects, $F > 4.0$, $p < .04$. All main effects and interactions were significant except Age × Delay, Age × Literal–Interpreted, the three-way interaction of Age × Delay × Literal–Interpreted, the three-way interaction of Age × Delay × True–False, and the five-way interaction containing all factors.

The acceptance rates on the immediate test for each of the eight critical recognition sentences are displayed in Figure 1A. Younger and older children are shown separately. The same comparisons after the delay are displayed in Figure 1B. In addition to the overall patterns noted earlier, it can be seen that younger children showed less verbatim discrimination than did the older children; they accepted the metaphorical target less often and accepted the nearest true distractor more often compared with older children. Further, even on the immediate test, false recognitions of true and false psychological interpretations were elevated for younger children. However, as for older children, false recognitions for true psychological interpretations increased more over the delay than did those for false psychological interpretations. Reyna and Kiernan (1994) obtained similar results with true and false literal sentences, namely that truth, as opposed to surface similarities, determines acceptances more after a delay.

Also consistent with earlier findings for literal sentences, recognition judgments for presented sentences were independent of judgments of semantically related sentences. That is, stochastic dependencies between judgments for each sentence type were evaluated (dependencies were computed between sentences that were related to the same presented metaphor), and children's

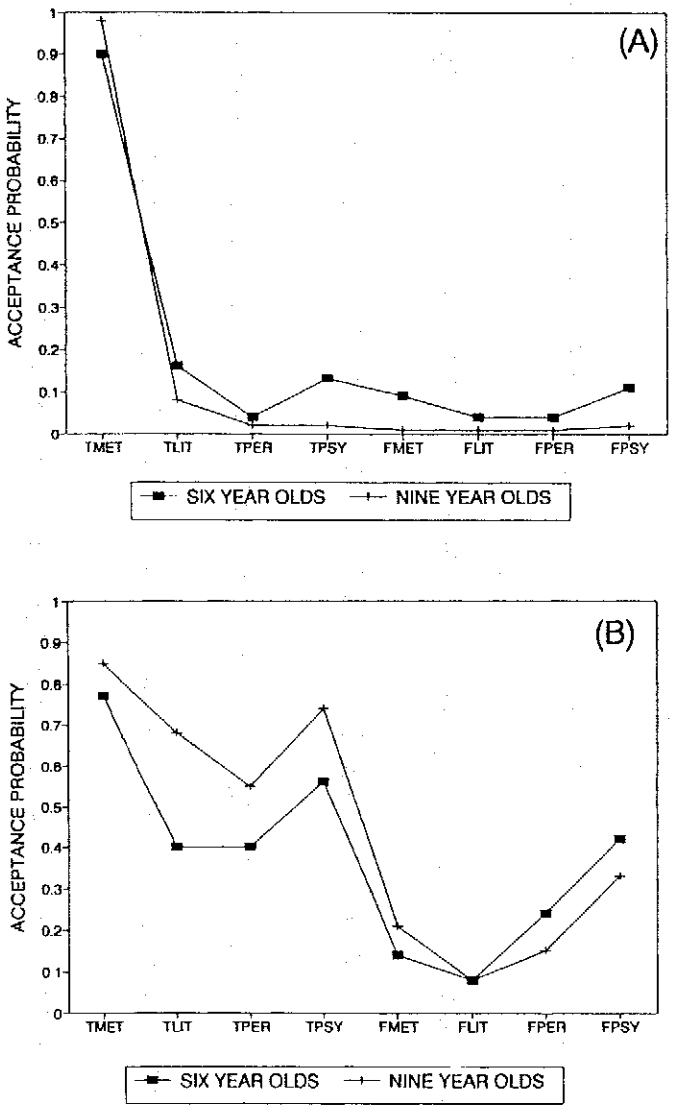

FIGURE 1 Mean proportion of affirmative memory judgments for different sentence types in Experiment 1 for each age level on the immediate test (Panel A) and for each age level on the delayed test (Panel B). TMET = true metaphor (actually presented); TLIT = true sentence with literal synonym; TPER = true perceptual interpretation; TPSY = true psychological interpretation; FMET = false metaphor; FLIT = false sentence with literal associate; FPER = false perceptual interpretation; FPSY = false psychological interpretation (same valence as TPSY).

acceptance of semantically related sentences was found to be independent of their acceptance of presented metaphors. As in Reyna and Kiernan (1994), however, judgments of sentences that expressed the gist of presented sentences were sometimes dependent on one another. For example, judgments of psychological interpretations were positively related to judgments of perceptual interpretations. Overall, positive dependencies were slightly more likely to be observed among younger children than among older ones.

Consistent with fuzzy-trace theory, these results support the conclusion that children retain verbatim representations of sentences even after a week's delay, and that verbatim recognition of presented sentences is independent of gist-based acceptance of semantically related sentences. However, two results regarding acceptance of metaphorical interpretations were surprising. First, true psychological interpretations were accepted more often than were true perceptual interpretations—twice as often for 6-year-olds, for example, after the delay. As we noted earlier, most theories of metaphor development suggest that true perceptual interpretations should be favored, particularly by the younger group. Second, acceptance of these alternative metaphorical interpretations was not mutually exclusive. Indeed, the findings of positive dependency among interpretations imply that children were more likely to accept a true psychological interpretation of a metaphor if they had accepted the true perceptual interpretation of that same metaphor and vice versa. These findings regarding acceptance of interpretations in Experiment 1 are only suggestive, however, because children were instructed to reject true interpretations. As the independence results indicate, contrary to constructivism, memory errors do not necessarily offer a window on comprehension. Therefore, in Experiment 2, we asked children directly about their metaphorical interpretations.

EXPERIMENT 2

The three methods typically used to assess children's metaphorical interpretation are paraphrase (verbal explanation of the metaphor's meaning), forced choice among proposed interpretations, and enactment of interpretations using props (Winner, 1988). As Vosniadou (1987a) pointed out, paraphrase inflates the false-negative error rate. That is, children are usually able to understand a metaphor before they are able to articulate that understanding (Vosniadou et al., 1984). Paraphrase is not simply a more conservative measure, however. As research in many domains of cognitive development has shown, verbal explanation imposes extraneous performance requirements (that are irrelevant to the competence in question), and when the reliability of verbal explanations and other behavioral measures were compared, explanations were decidedly less reliable (e.g., Brainerd, 1973; Reyna & Brainerd,

1990, 1994). Therefore, although paraphrase seems as though it is a direct, if stringent, measure, it is less likely to reflect true metaphorical competence.

Although enactment generally reduces performance requirements (most children can act out interpretations before they can explain them verbally), as Winner (1988) noted, literal props have typically not been provided, which might erroneously reduce the rate of literal interpretations. Even if such props were provided, however, it is difficult to convey subtleties of meaning through enactment. Finally, forced-choice tasks reduce performance requirements, but their sensitivity depends on the alternatives that are provided. Obviously, if alternatives are grossly different, a child can choose the correct response without fully comprehending the meaning of the chosen alternative. Keil (1986) suggested, for example, that children can distinguish the correct polarity of a psychological metaphor—that something positive as opposed to negative is being asserted—without understanding the metaphor's meaning in any specific sense. Nippold et al. (1984) unconfounded polarity from true versus false psychological interpretations (i.e., true and false interpretations for a given metaphor shared the same polarity) and found that children were able to select the correct (true) psychological interpretations. More generally, because children cannot be credited with appreciating distinctions if these distinctions are confounded across response alternatives, the sensitivity of the forced-choice task depends on the subtlety of differences among alternatives.

Therefore, the forced-choice task has the most desirable measurement properties of those tasks that have been used, provided that the right alternatives are included. However, this approach has the major drawback that participants cannot convey that multiple interpretations are acceptable. This is especially troubling given the findings from Experiment 1 that children accepted different metaphorical interpretations and that acceptance of one type did not preclude acceptance of another type (i.e., judgments were not negatively dependent). Of course, forced choice would be justifiable for metaphors that had only one appropriate interpretation. However, such metaphors do not appear to exist. Although it has been argued that so-called psychological metaphors have one appropriate type of interpretation, they actually have at least three possibly appropriate types of interpretations, depending on the context (which may account for differences between metaphors and similes, which lack such ambiguities, see Vosniadou, 1987b).

As Winner (1988) acknowledged, there are many examples of perceptual interpretations of metaphors in adult literature. It would not be difficult to construct a scenario in which *"The prison guard had become a hard rock"* could indeed be interpreted as the prison guard developing hard muscles (Reyna, 1985, 1986). Furthermore, sentences that can be interpreted metaphorically in one context can be interpreted magically in a different context (Reyna, 1985, 1987). In the context of a fairy tale, for example, the sentence could mean that a witch turned the prison guard into a rock. Examples of magical interpretations in adult literature, for instance in science fiction, abound (e.g., the talking cow in *The Hitchhiker's Guide to the Galaxy* series).

Out of context, the sentence *The prison guard had become a hard rock* is simply ambiguous with respect to three possible interpretations: magical, perceptual, or psychological. When forced to choose among these interpretations, children are more likely to interpret sentences magically because of the genres to which they are exposed (e.g., fantasy, fairy tales, and cartoons; see Reyna, 1985). However, as Reyna (1985) showed, even young children are more likely to assign metaphorical than magical interpretations to "metaphors" when the context is disambiguated. (In magical interpretations, words retain aspects of their literal meaning, so these interpretations are often classified as literal; for distinctions between magical and literal interpretations, see Reyna, 1985.) Therefore, when the context does not clearly point to one of the three possible types of interpretations, children must be able to select more than one interpretation. To that end, we adopted a yes–no task in which children could accept (or reject) each of the true alternatives.

Method

Participants. Fifty children, 25 in the younger group of 6-year-olds (M = 6 years, 6 months) and 25 in the older group of 9-year-olds (M = 9 years, 6 months), participated in the second experiment and in one described in Reyna and Kiernan (1994). All of the children were drawn from the same school used in Experiment 1. Again, they were about evenly divided according to gender, and none was identified as learning or language disabled.

Design and materials. The design and materials were the same as those in Experiment 1.

Procedure. The procedure, which included oral and individual administration, was identical to that of Experiment 1 up to the point where children were told that they would hear some stories. Instead of asking children to remember the stories, they were instructed to "try to understand the stories. Think about what the story means." Children were told that they would be asked about what happened in the story and that they should "say yes if that's what the story meant. Say no if the story didn't mean that."

The experimenter then presented a series of examples. Prior to the examples, children were reminded to "think about what the story means." Children were given three example recognition sentences: a presented metaphorical sentence, a false sentence, and a true interpretation. Again, children provided responses to the practice examples, and most responded quickly and accurately. For each sentence, the experimenter also explained

the basis for the correct response. For the true interpretation, the experimenter acknowledged that the story "didn't exactly say that." The experimenter then repeated the story and pointed out "But, all of that means that . . . ," and she repeated the true interpretation. False sentences were used to illustrate that literal lexical association was misleading and that such sentences should be rejected. For example, given a practice vignette involving a policeman patrolling a neighborhood, children were told to reject a false sentence about the policeman turning on his siren, despite the literal association between police cars and sirens, because "That didn't happen in the story." Children were told to say yes whenever sentences "mean the same thing" as what was said in the story. Finally, children were cautioned that "some of these will seem like they mean the same thing, but they don't" and, therefore, that they should "listen carefully." Stories and recognition sentences were presented in a different random order for each child. Prior to each story, children were reminded to think about what the story means.

After the delay, children were told that they would be asked about the stories that had been read to them the previous week and that they should answer the questions just like they did then. They were also reminded that they "should say yes if this is what happened in the story" and they should "say no if it didn't happen this way." The order of testing of stories was randomized again, as was the order of recognition sentences, separately for each child. Children were reminded to answer according to meaning prior to the recognition sentences for each story.

Results and Discussion

Results of this experiment, in which meaning was stressed, bolstered the conclusion that children's misrecognitions in the verbatim task were based on their representations of the gist of the metaphors. When asked to judge meaning, children affirmed true psychological interpretations more often than they did any other type, including true perceptual interpretations. Although the relative popularity of these two types of interpretations was the same for older and for younger children, there was a smaller difference between true and false perceptual and psychological interpretations for the younger children. In other words, both age groups favored psychological and, to a lesser degree, perceptual interpretations, but younger children seemed less able to discriminate true interpretations from false interpretations.

These results were confirmed by an ANOVA with the same five factors specified in Experiment 1. For all significant effects, $F > 4.0, p < .05$. Of the main effects, only true–false was significant. There were 10 two-way and three-way interactions, all of which were included in 2 four-way interactions (Age × Delay × True–False × Literal–Interpreted and Age × Delay × True–False × Close–Far) and in the five-way interaction among all the factors. The results are shown in Figure 2. As can be seen in the figure, acceptance of false interpretations declined with age, but true perceptual interpretations did not.

Unlike in Experiment 1, judgments of presented metaphors and of true interpretations were generally positively dependent. Instructing children to base all judgments on gist converted independence to positive dependency for the same presented metaphors and true interpretations. In addition, as in Experiment 1, judgments of unpresented true interpretations were usually positively dependent. These patterns of dependencies replicate those found with literal language (Reyna & Kiernan, 1994). Thus, when children were asked to judge meaning, there was no evidence that alternative interpretations were suppressed. In other words, children who accepted a psychological interpretation of a metaphor were not less likely to accept a perceptual interpretation or a literal-synonym sentence. On the contrary, they were more likely to accept such sentences (i.e., they were positively dependent).

GENERAL DISCUSSION

The results reported in this article for metaphors are similar to those obtained with literal language (Reyna, 1995; Reyna & Kiernan, 1994) and with numerical information (Reyna & Brainerd, 1993, 1994). At a general level, therefore, the explanation offered for those findings can readily be applied. In the verbatim memory experiment, there was evidence of systematic misrecognition of gist, especially after a delay. Children were more likely to misrecognize a true interpretation of a metaphor as having been presented relative to the misrecognition of false sentences. However, contrary to predictions of information-processing theory and of constructivism, misrecognizing the gist of a metaphor was independent of memory for the metaphor itself.

Such findings can be explained by assuming that presented metaphors cue verbatim memory, but their interpretations cue memory for gist (e.g., Reyna, in press; Reyna & Brainerd, 1995). This explanation is supported by the fact that the independence relation could be manipulated by varying instructions to base judgments on gist, as opposed to verbatim memory, as in Experiment 2. The result that judgments of different true interpretations were positively dependent, regardless of instructions, is also consistent with the claim that these interpretations cue memory for gist.

In the meaning experiment, both age groups accepted the true psychological interpretations of the metaphors more often than they accepted any of the other alternatives. Interestingly, the order of acceptance was the same for false interpretations; false psychological interpretations were affirmed more often than were false perceptual ones. Thus, younger children clearly favored

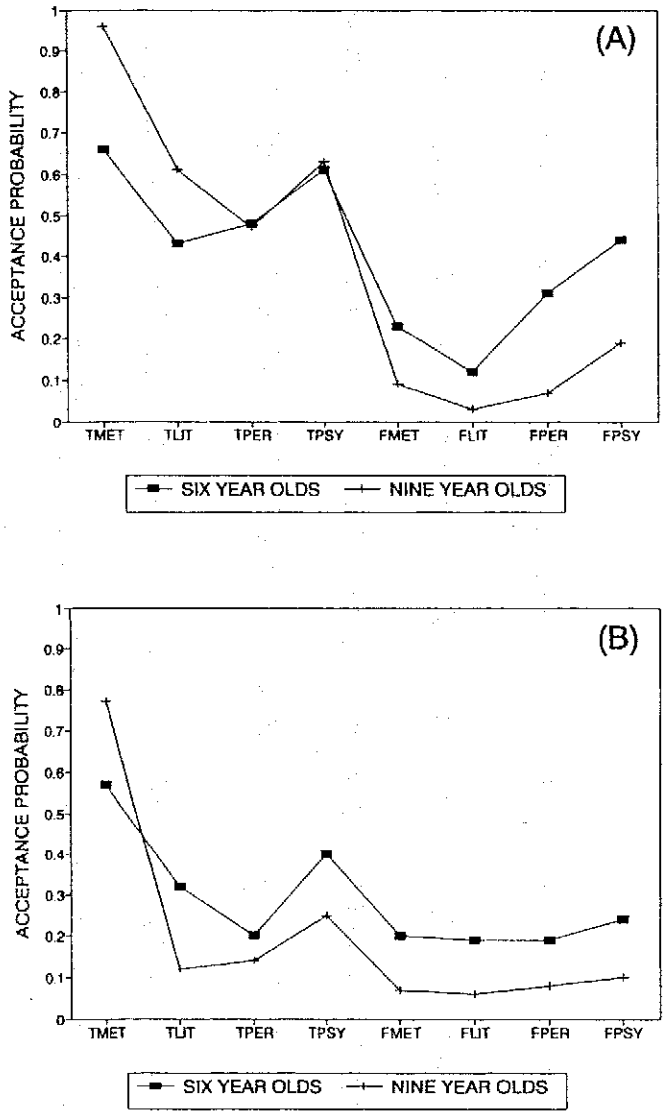

FIGURE 2 Mean proportion of affirmative meaning judgments for different sentence types in Experiment 2 for each age level on the immediate test (Panel A) and for each age level on the delayed test (Panel B). TMET = true metaphor (actually presented); TLIT = true sentence with literal synonym; TPER = true perceptual interpretation; TPSY = true psychological interpretation; FMET = false metaphor; FLIT = false sentence with literal associate; FPER = false perceptual interpretation; FPSY = false psychological interpretation (same valence as TPSY).

the most abstract interpretations even when they failed to ascertain which interpretation was correct. Such results contradict the assumption that younger children would favor perceptual resemblances over psychological ones. Furthermore, there was no evidence that perceptual resemblances were understood at an earlier age. Younger children did not discriminate between true and false interpretations better for perceptual comparisons than they did for psychological comparisons. Also, although true–false discrimination improved markedly with age, it improved almost equally for perceptual and psychological interpretations.

These developmental trends can be put into perspective by considering results for adults. Using the materials and tasks from this study, Lim (1993) found that adults also favored the true psychological interpretations over other types. However, true perceptual interpretations were affirmed at a relatively high rate (40%) roughly comparable to the rate for children (47%) and decidedly higher than the rate for false perceptual interpretations (8%). Also consistent with the results for children, dependencies among the alternative true interpretations were generally positive.

Taken together, the results for children and adults suggest that discrimination between true and false interpretations increases with age, but that the ordering of types of acceptable interpretations remains constant. Psychological interpretations were preferred even among the youngest children, and there was no evidence that this kind of abstract resemblance imposed special cognitive burdens. Such results raise the question of how appropriate interpretations (gist) are achieved.

Positive dependencies indicate that literal features of word meanings, both functional and perceptual, are not suppressed in order to arrive at appropriate metaphorical interpretations. Rather than rejecting these aspects of conventional meaning, interpreters appear to use them as a bridge to metaphorical meaning. For example, consider *"The woman was an aspirin, kneeling by the lost boy."* The literal function of an aspirin, to ease physical discomfort, provides a link to the metaphorical interpretation that the woman eased the boy's psychological discomfort. Positive dependencies would be expected between literal and psychological interpretations (and between perceptual and psychological interpretations) because linked interpretations would share a common gist.

Specifically, Reyna (in press) proposed that children and adults apprehend metaphorical meaning by processing the broad gist of conventional meaning and then reinstantiating that gist so that it coheres with the context. Thus, the conventional property of aspirins—to ease physical discomfort—would be interpreted in a generic sense and then applied to the scenario of a woman kneeling by a lost boy. Response-time data, which show that generic interpretations of novel metaphors are verified more quickly than are specific interpretations, support this view (Reyna, 1981, 1987). Because metaphorical

comparisons traverse different categories (Glucksberg, 1991)—the woman is not really an aspirin—the features of an aspirin must be viewed in a more abstract way in order to apply them to women. Therefore, this account addresses the attribute inequality problem, namely, that "shared" attributes (e.g., easing discomfort) do not mean the same thing when they are applied literally versus metaphorically (Ortony, 1979).

The account also implies that there is a continuum that varies from literal to metaphorical interpretation, depending on how much the interpreter must stretch the meaning of semantic features (Reyna, 1985, in press). Therefore, we assume that literal and metaphorical meaning differ, but by degree (cf. Winner, Levy, Kaplan, & Rosenblatt, 1988). This explains the fact that it is sometimes difficult to judge whether an interpretation is an extended literal sense of a word, common in childhood, or a metaphorical interpretation (Winner, 1988). It also explains how literal and metaphorical interpretation could proceed in parallel so that their effects converge rather than disrupt one another (Glucksberg, 1991; Keysar, 1989). The rationale for this convergence would be that shared gist, as well as similar pragmatic principles, support both literal and metaphorical interpretation. Pragmatic principles would allow interpreters to exploit conventional denotation as well as connotation, relations with other words, and knowledge of the context and the world in order to instantiate meanings (Kittay, 1987; Lehrer, 1970, 1992; Reyna, in press; Vosniadou, 1987a, 1989). The elevated rates of acceptance for false psychological interpretations indicate that children's difficulties primarily are in instantiating the appropriate metaphorical interpretation rather than in an inability to apprehend the broader gist (for analogous difficulties in reasoning, see Reyna, 1995).

In summary, the account of metaphorical interpretation given in this article is similar to that offered by fuzzy-trace theory for verbal and numerical reasoning in general (Reyna, 1995; Reyna & Kiernan, 1994). This makes sense given that the findings for metaphorical interpretation resemble those for verbal and numerical reasoning, especially concerning memory for presented information and its relation to memory for gist. Such similarities indicate that comparisons between reasoning and understanding a metaphor are not wrong. The key to the comparison, however, may be that reasoning is more like understanding poetry than it is like doing logic (Reyna & Brainerd, 1995).

ACKNOWLEDGMENTS

The research reported in this article was partially supported by Grant HO29D90108 from the United States Department of Education to Linda Swisher.

We gratefully acknowledge the assistance of the children, parents, teachers, and staff, and of the principal, Jo Ann Sayre, of St. Cyril's School, Tucson, AZ.

REFERENCES

Ackerman, B. P. (1992). The sources of children's source errors in judging causal inferences. *Journal of Experimental Child Psychology, 54,* 90–119.

Alba, J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203–231.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Billow, R. (1975). A cognitive-developmental study of metaphor comprehension. *Developmental Psychology, 11,* 415–423.

Bjorklund, D. F. (1989). *Children's thinking: Developmental and individual differences.* Pacific Grove, CA: Brooks/Cole.

Brainerd, C. J. (1973). Judgments and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin, 79,* 172–179.

Brainerd, C. J. (1978). The stage question in cognitive-developmental theory. *The Behavioral Brain Sciences, 1,* 173–182.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100,* 42–67.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretative approach. *Cognitive Psychology, 3,* 193–209.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331–380.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development, 48,* 1454–1466.

Case, R. (1985). *Intellectual development.* New York: Academic.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Demorest, A., Silberstein, L., Gardner, H., & Winner, E. (1983). Telling it as it isn't: Children's understanding of figurative language. *British Journal of Developmental Psychology, 1,* 121–134.

Falmagne, R. J. (1975). *Reasoning: Representation and process.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Flagg, P. W. (1976). Semantic integration in sentence memory? *Journal of Verbal Learning and Verbal Behavior, 15,* 491–504.

Fletcher, C. R. (1992). Assessing recognition memory for surface forms in discourse: A methodological note. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 199–203.

Gauld, A., & Stephenson, G. M. (1967). Some experiments related to Bartlett's theory of remembering. *British Journal of Psychology, 58,* 39–49.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusions. *Psychological Science, 2,* 146–152.

Hasher, L., Attig, M. S., & Alba, J. W. (1981). I knew it all along: Or did I? *Journal of Verbal Learning and Verbal Behavior, 20,* 86–96.

Hasher, L., & Griffin, M. (1978). Reconstructive and reproductive processes in memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 318–330.

Keil, F. (1986). Conceptual domains and the acquisition of metaphor. *Cognitive Development, 1,* 73–96.

Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language, 28,* 375–385.

Kittay, E. F. (1987). *Metaphor: Its cognitive force and linguistic structure.* Oxford, England: Clarendon Press.

Kreindler, D., & Lumsden, C. (1994). Extracting a narrative's causal gist: A modeling study based on fuzzy-trace theory. *Journal of Experimental Child Psychology 58,* 227–251.

Lehrer, A. (1970). Indeterminacy in semantic description. *Glossa, 4,* 87–110.

Lehrer, A. (1992). A theory of vocabulary structure: Retrospectives and prospectives. In M. Putz (Ed.), *Thirty years of linguistic evolution.* Philadelphia: Benjamins.

Liben, L. S., & Posnansky, C. J. (1977). Inferences on inference: The effects of age, transitive ability, memory load, and lexical factors. *Child Development, 48,* 1490–1497.

Lim, P. (1993). *Meaning versus verbatim memory in language processing: Deriving inferential, morphological, and metaphorical gist.* Unpublished doctoral dissertation, University of Arizona, Tuscon.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Nippold, M., Leonard, L., & Kail, R. (1984). Syntactic and conceptual factors in children's understanding of metaphors. *Journal of Speech and Hearing Research, 27,* 197–205.

Ortony, A. (1979). Beyond literal similarity. *Psychological Review, 86,* 161–180.

Paris, S. G., & Carter, A. Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology, 9,* 109–113.

Paris, S. G., & Lindauer, B. K. (1977). Constructive aspects of children's comprehension and memory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 35–60). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Perner, J., & Mansbridge, D. G. (1983). Developmental differences in encoding length series. *Child Development, 54,* 710–719.

Piaget, J., & Inhelder, B. (1973). *Memory and intelligence.* New York: Basic Books.

Prawatt, R. S., & Cancelli, A. (1976). Constructive memory in conserving and nonconserving first graders. *Developmental Psychology, 12,* 47–50.

Rabinowitz, F. M., Howe, M. L., & Lawrence, J. A. (1989). Class inclusion and working memory. *Journal of Experimental Child Psychology, 48,* 379–409.

Reitman, J. S., & Bower, G. H. (1973). Storage and later recognition of exemplars of concepts. *Cognitive Psychology, 4,* 194–206.

Reyna, V. F. (1981). The animated word: Modification of meaning by context. *Dissertation Abstracts International, 42,* 3852B.

Reyna, V. F. (1985). Figure and fantasy in children's language. In M. Pressley & C. Brainerd (Eds.), *Cognitive learning and memory in children: Progress in cognitive development research* (pp. 143–179). New York: Springer-Verlag.

Reyna, V. F. (1986). Metaphor and associated phenomena: Specifying the boundaries of psychological inquiry. *Metaphor and Symbolic Activity, 4,* 271–290.

Reyna, V. F. (1987). Understanding verbs: Easy extension, hard comprehension. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 301–315). London: Lawrence Erlbaum Associates, Inc.

Reyna, V. F. (1991). Class inclusion, the conjunction fallacy, and other cognitive illusions. *Developmental Review, 11,* 317–336.

Reyna, V. F. (1992). Reasoning, remembering, and their relationship: Social, cognitive, and developmental issues. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 103–127). New York: Springer-Verlag.

Reyna, V. F. (1995). Interference effects in memory and reasoning: A fuzzy-trace theory analysis. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 29–59). San Diego, CA: Academic.

Reyna, V. F. (in press). Meaning, memory, and the interpretation of metaphors. In J. Mio & A. Katz (Eds.), *Metaphor: Pragmatics and applications.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Reyna, V. F., & Brainerd, C. J. (1990). Fuzzy processing in transitivity development. *Annals of Operations Research, 23,* 37–63.

Reyna, V. F., & Brainerd, C. J. (1991). Fuzzy-trace theory and children's acquisition of mathematical and scientific concepts. *Learning and Individual Differences, 3,* 27–59.

Reyna, V. F., & Brainerd, C. J. (1993). Fuzzy memory and mathematics in the classroom. In R. Logie & G. Davies (Eds.), *Memory in everyday life* (pp. 91–119). Amsterdam: North Holland.

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239–272). New York: Wiley.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7,* 1–75.

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy–trace theory and framing effects in children's risky decision making. *Psychological Science, 5,* 275–279.

Reyna, V. F., & Kiernan, B. (1994). The development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30,* 178–191.

Siegler, R. S. (1991). *Children's thinking.* Englewood Cliffs, NJ: Prentice-Hall.

Small, M. Y., & Butterworth, J. (1981). Semantic integration and the development of memory for logical inferences. *Child Development, 52,* 732–735.

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children's word problem solution. *Journal of Experimental Child Psychology, 55,* 374–395.

Trabasso, T. (1977). The role of memory as a system in making transitive inferences. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 333–366). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Vosniadou, S. (1987a). Children and metaphors. *Child Development, 58,* 870–885.

Vosniadou, S. (1987b). Contextual and linguistic factors in children's comprehension of nonliteral language. *Metaphor and Symbolic Activity, 2,* 1–11.

Vosniadou, S. (1989). Context and the development of metaphor comprehension. *Metaphor and Symbolic Activity, 4,* 159–171.

Vosniadou, S., Ortony, A., Reynolds, R., & Wilson, P. T. (1984). Sources of difficulty in the young child's understanding of metaphorical language. *Child Development, 55,* 1588–1606.

Vosniadou, S., Pearson, P. D., & Rogers, T. (1988). What causes children's failures to detect inconsistencies in text? Representation versus comparison difficulties. *Journal of Educational Psychology, 80,* 27–39.

Winner, E. (1988). *The point of words.* Cambridge, MA: Harvard University Press.

Winner, E., Levy, J., Kaplan, J., & Rosenblatt, E. (1988). Children's understanding of nonliteral language. *Journal of Aesthetic Education, 2,* 51–63.

Winner, E., Rosenstiel, A., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology, 12,* 289–297.

Winner, E., Windmueller, G., Rosenblatt, E., Bosco, L., Best, E., & Gardner, H. (1987). Making sense of literal and nonliteral falsehood. *Metaphor and Symbolic Activity, 2,* 13–32.