Journal of Educational Psychology
1998, Vol. 90, No. 1, 13–24

0022-0663/98/$3.00

The Influence of Moral Schemas on the Reconstruction of Moral Narratives in Eighth Graders and College Students

Darcia Narvaez
University of Minnesota, Twin Cities Campus

Much attention has been focused on the importance of reading moral stories to children (e.g., W. Bennett, 1993). Although research on general discourse comprehension is flourishing, little attention has been given to how moral discourse is understood by individuals; that is, what affects an individual's comprehension of a moral text? Eighth-grade and college students read and recalled four complex moral narratives in which moral arguments at different Kohlbergian stages were embedded. Participants then took the Defining Issues Test (DIT), a measure of moral judgment development. Those with higher reasoning scores on the DIT reconstructed more high-stage moral arguments during recall, including adding high-stage moral reasoning that was not in the original text. Significant age-level differences in cumulative moral judgment concepts were also found. Prior moral knowledge affected the comprehension of complex moral narratives.

In the 1990s, there has been widespread popular interest in reading moral stories to children to develop moral literacy (e.g., Bennett's, 1993, 1995, best-selling books). Underlying this popularity, there seems to be an implicit assumption that individuals (e.g., adult writers and child readers) comprehend moral texts in the same way. However, text comprehension research has demonstrated that readers do not comprehend (nonmoral) texts in the same way because of individual differences in skill and background knowledge (see, e.g., Gernsbacher, 1994). In other words, a comprehender does not necessarily understand what the author intended. In addition, considerable empirical evidence exists for developmental and expert–novice differences in moral judgment (e.g., Rest, 1986) that suggests individuals often view social events differently and, as a result, perform moral comprehension tasks distinctively (e.g., Rest, Thoma, & Edwards, 1997).

Darcia Narvaez, Department of Curriculum and Instruction and Department of Educational Psychology, University of Minnesota.

This article is based on data collected for a doctoral dissertation submitted to the University of Minnesota in partial fulfillment of the requirements for the doctorate degree. The design has been presented at the annual meetings of American Educational Research Association, Association for Moral Education, and Jean Piaget Society. This research was supported by the University of Minnesota's Center for the Study of Ethical Development and Center for Research in Learning, Perception, and Cognition and National Institute of Child Health and Human Development Grant HD-07151.

The following people provided invaluable assistance in completing this work: Muriel Bebeau, Amy Briggs, Jean Evens, James Rest, Kirsten Risden, and Steve Thoma. I gratefully acknowledge the support and guidance of Paul van den Broek on the dissertation.

Correspondence concerning this article should be addressed to Darcia Narvaez, College of Education and Human Development, University of Minnesota, 150 Peik Hall, 159 Pillsbury Drive Southeast, Minneapolis, Minnesota 55455. Electronic mail may be sent to narvaez@turtle.psych.umn.edu.

What factors are involved in understanding moral texts? The main purpose of this study was to begin to examine this question. Methodologies from two research traditions—text comprehension and moral development—were used as well as groups of students who have different levels of moral judgment development. This research opens a new window into the moral mind for moral judgment research by using memory for moral arguments embedded in narratives. For both text comprehension and morality research, it offers an examination of schemas and their effects. As it turns out, comprehending moral narratives is more complicated than it first appears.

Text Comprehension Research

Many factors are involved in comprehending texts. One factor is individual differences; that is, individuals who read the same text often end up with different mental representations of the text. Reading researchers have studied differences in the comprehension of texts along two lines. One branch focuses on basic reading and language abilities, such as vocabulary or memory, and finds that readers with more skills are better at comprehending texts (e.g., Cunningham, Stanovich, & Wilson, 1990; Palmer, MacLeod, Hunt, & Davidson, 1985). The second branch addresses differences in specific knowledge brought to the text by readers and has demonstrated that readers with more text-relevant knowledge are better able to comprehend the text (e.g., Anderson & Pearson, 1984; Spilich, Vesonder, Chiesi, & Voss, 1979). In other words, prior knowledge about the world can affect how a reader reads and remembers text.

In general, as a reader reads and remembers text, he or she attempts to create a coherent mental representation not only by integrating text information but also by elaborating on the text with prior knowledge about the world (van den Broek, 1994) and by building a mental model (overall meaning structure) of the text (McNamara, Miller, & Bransford, 1991; van Dijk & Kintsch, 1983). Prior knowledge often

comes in the form of general knowledge structures. General knowledge structures such as scripts (e.g., Brown, Smiley, Day, Townsend, & Lawton, 1977; Nelson, 1986; Schank & Abelson, 1977) and schemas (e.g., Anderson & Pearson, 1984; Bartlett, 1932; Bobrow & Norman, 1975; Rumelhart, 1980; Rumelhart & Ortony, 1977) have been shown to affect how readers comprehend a text. For example, because of extensive familiarity with sit-down restaurants, a reader likely has a general knowledge "script" of the type and order of events that occur in restaurants (a restaurant script), which could affect the reader's recall of a text about a restaurant visit. When a reader familiar with restaurants reads a text like the following, a restaurant script may be activated: "John ordered from the menu. When he was satisfied, he paid the bill." The reader might add details at recall that were not in the text such as "John sat at a table, received food, ate it, and was given a bill." Such added detail would be evidence for the existence of a restaurant script. A schema functions in a way similar to a script, except that it is less rigidly structured and ordered. For example, if a speaker indicated a prior visit to the beach, the listener would infer that the speaker relaxed on the beach and interacted with the adjacent body of water somehow but, unlike a script, not in any particular manner, order, or duration.

The effects of schemas on text understanding have been documented in situations involving culturally specific texts (e.g., Bartlett, 1932; Harris, Lee, Hensley, & Schoen, 1988), reader orientation at reading or recall (e.g., Anderson & Pichert, 1978; Anderson, Reynolds, Schallert, & Goetz, 1977; Pichert & Anderson, 1977), reader familiarity with text material (e.g., Chiesi, Spilich, & Voss, 1979; Crafton, 1983; Spilich et al., 1979), and reader prior knowledge (e.g., Bartlett, 1932; Reynolds, Taylor, Steffensen, Shirey, & Anderson, 1982; Steffensen, Joag-Dev, & Anderson, 1979). Schema effects are strongest with ambiguous material in which referential specificity is low (it is not clear to what the sentence or phrase refers), local coherence is weak (the phrases and sentences are not very related), and the message is unclear or nonsensical until a theme or title is provided (Bransford & Johnson, 1972; Dooling & Lachman, 1971). In addition, the longer the interval before recall, the more inaccuracies there are and the more likely it is that memory reconstruction is affected by the individual's own perspectives (e.g., schemas) in terms of theme sharpening (embellishment, emphasis, rationalization) and theme leveling (discarding, condensation) of irrelevant material (Bartlett, 1932; Brown et al., 1977; Dooling & Christiaansen, 1977; Dooling & Lachman, 1971; Sulin & Dooling, 1974). In short, schemas and scripts can influence the reader's mental representation of a text and are demonstrated by the characteristics of what a reader recalls or does not recall from the text, including distortions, intrusions, and the elimination of information that does not match the schemas of the reader.

Schema-based or top-down processing contrasts with bottom-up or data-based processing. In top-down processing, a whole knowledge structure (such as a restaurant script or beach schema) is evoked by a word or event in the text. Later events in the text are then interpreted according to the schema. (For example, "When he was satisfied, he paid the bill" is an ambiguous sentence that is interpreted according to the schema activated by the previous sentence, "John ordered from the menu.") There has been a shift away from top-down, schema-based theories as theory and evidence for bottom-up processing have grown (see reviews by Pressley & Afflerbach, 1995; Whitney, Budd, Bramucci, & Crane, 1995). In bottom-up processing, the words and propositions of the text activate prior knowledge in the reader to produce word and conceptual associations. The particular meaning of the word or proposition that remains activated as the reader continues reading depends on the context (on the other words in the text). Therefore, bottom-up processing is highly dependent on the particular characteristics of the text itself.

One characteristic of texts that influences comprehension is cohesion: how related or interconnected the elements in the text are according to how closely they refer to one another or how causally connected they are. For example, the following text is not causally or referentially coherent because there is no causal connection between the events in the first sentence and the second sentence and because the referents "she" and "it" in the second sentence are not related to the first sentence: "Jake carried the drinks to the cooler. She drank it." Of course, we could make it a referentially and causally coherent text by adding a middle sentence: "He handed a carton of milk to Myra." A causal chain of events is then established. If the text continued and discussed how the milk made Myra sick, which made her leave the party early and miss seeing a friend who came late, the sentence "She drank it" would have a high number of causal relations to other events in the text. Text events that have a greater number of causal connections are better recalled, more likely to be included in summaries, and rated as more important (Trabasso, Secco, & van den Broek, 1984; Trabasso & Sperry, 1985; van den Broek, 1988, 1994).

Theories that combine schema-based, top-down processes with bottom-up processes have been proposed to account for data that support both types of processing (e.g., Pressley & Afflerbach, 1995; Whitney et al., 1995). Some reading theorists contend that schemas relevant to the discourse guide the construction of the mental model during reading (e.g., Kintsch, 1988; van Dijk & Kintsch, 1983) and help in the selection of what is relevant or irrelevant to keep in the mental representation (Singer, 1994). It is this latter view that underpins the research reported here. A particularly relevant theory is Pressley and Afflerbach's (1995) "constructively responsive reading," in which readers actively search for meaning, constructing interpretations based on prior knowledge and the reader's processing of the text. The data here offer support for that view.

Moral text understanding involves not only general reading processes such as schema-based processing and bottom-up, text-based processing but moral cognitive processes as well. The texts used in this study contained not only general events but events that represent types of moral thinking.

Moral Schemas

Piaget (1932/1965) and Kohlberg (1969, 1984) studied moral thinking by presenting participants with a moral dilemma, asking what action should be done and justification of the action choice. Kohlberg classified the moral justifications that people produce into one of six categories. The six types of response can be viewed as schemas for various concepts of cooperation—different solutions to the problem of getting along with others (see Rest, Narvaez, Bebeau, & Thoma, 1998, for a thorough discussion). Kohlberg observed that logically simpler concepts of justice develop earlier and the logically more complex concepts develop later. For instance, a Stage 2 concept of cooperation involves a simple, direct, one-time exchange (you do me a favor and I'll do you a favor). A Stage 4 concept of cooperation involves organizing lasting society-wide cooperation not only among friends and familiar intimates but also among strangers. The construct validity of the Kohlbergian sequence of six schemas or stages has been supported in several ways (Rest, Thoma, & Edwards, 1997), more specifically by longitudinal studies (e.g., Colby & Kohlberg, 1987; McNeel, 1994; Rest, 1986), cross-sectional age–education trends (e.g., Kohlberg, 1984; Rest, 1986), high correlations with theoretically similar constructs and low correlations with theoretically dissimilar constructs (e.g., Rest, 1986; Thoma & Rest, 1997; Walker, 1991; Walker, deVries, & Bichard, 1984), and reactions to predicted experimental manipulations (e.g., Rest, Narvaez, Bebeau, & Thoma, 1998; Walker, 1988).

A neo-Kohlbergian reformulation of moral judgment development proposes that stages represent moral schemas that can be characterized as "prior moral knowledge" about different ways to get along with others (see Rest et al., 1998, for a thorough discussion). The relation between prior moral knowledge and moral judgment schemas has been illustrated by moral comprehension studies. These studies measure the capacity of participants to understand moral schemas (e.g., Rest, 1973; Rest, Turiel, & Kohlberg, 1969; Walker et al., 1984) regardless of whether or not the participant actually uses the schema to solve the moral problem. Participants are asked to paraphrase reasoning statements from different stages of moral thinking. For example, the following is a Stage 4 statement from the Moral Comprehension Test:

> If Heinz steals, he is breaking his agreements with other members of society. In most countries, men have agreed not to steal because they see that not stealing is better for them. Heinz himself would have to admit that a law against stealing is a good law to have. And so if Heinz wants to have laws that he and other people think are good to have, he should abide by them. (Rest, 1979, pp. 82–83)

Comprehension studies examine whether the participant can correctly paraphrase the reasoning statement or whether the participant distorts the statement during the response task. Correct paraphrasing of a statement indicates that the participant is capable of reasoning at that level of moral understanding. The major finding in these studies of relevance here is that comprehension of moral schemas is cumulative (i.e., a participant who comprehends Stage 5 also comprehends Stages 4, 3, 2, and 1; a participant who has reached an understanding of Stage 3 only comprehends in addition Stages 2 and 1). Moral comprehension is significantly correlated with scores in moral judgment (range = .32–.67; see Rest, 1979).

Moral judgment schema development can be assessed by the Defining Issues Test (DIT), an objective test derived from Kohlberg's theory (Rest, 1979). The DIT presents moral dilemmas and asks participants to rate and rank justifications that represent different moral stages. The items are very brief fragments of a justification that make sense to a participant who has the schema that undergirds the justification. For instance, one item that represents Stage 4 is "whether a community's laws are going to be upheld." This DIT item presupposes that the item is ranked as important if a Kohlbergian-type schema about "law and order" underlies a person's thinking about the moral dilemma. DIT items can evoke a particular schema if the person has indeed been thinking about a dilemma in terms of that schema. If the schema is important to the participant, then the item representing it will receive a high rating. If the person has not been thinking in terms of the schema or if it is not considered important, the participant will not give that item a high rating. In summary, the DIT presupposes that items evoke certain schemas; how the person ranks the item represents how important the participant regards that schema.

Unlike Kohlberg's Moral Judgment Interview (Colby & Kohlberg, 1987), the DIT does not "stage type" the participant (i.e., categorizing him or her into a Stage 3 or a Stage 4 thinker). Instead, the DIT is based on a "soft-stage" model, which assumes that, with experience, people learn to use a variety of moral judgment schemas whose pattern of use changes with development. In other words, with increased experience, people use more of the higher stages and less of the lower stages (see Rest, 1979). The DIT P score measures the relative importance attributed to postconventional thinking (Stages 5 and 6) when a choice is given among different types of thinking (Stages 2–6 are presented and scored). In other words, the P score of the DIT provides a percent score that indicates the amount of postconventional thinking (in contrast to other kinds of thinking) preferred by the participant. Generally, previous research (Rest, 1986) has indicated percentages in the 20s for junior high students, 30s for senior high students, 40s for college students, and 50s for general graduate students.

The DIT uses a recognition task that assesses the moral schemas preferred by the participant in solving the moral dilemma. In contrast to preference selection, the moral recall task used in this study is similar to a moral comprehension task in being an inventory of moral schema capacity. Similar to measures of other kinds of knowledge, moral judgment can be measured at different levels of capacity, from recognition to generation. A recognition task can be described as measuring the low end of capacity because a correct response requires only familiarity. On the other hand, a verbalization task like the Moral Judgment Interview (Colby & Kohlberg, 1987) can be described as a high-end task in terms of capacity because it necessitates the formulation of a response based on both accessible conceptual

understanding and verbal ability. A recall task (used in this study) is a middle-level task that is more difficult than response selection based on familiarity (recognition) and yet easier than conceptualizing and articulating justifications (verbalization). The recall task requires the participant only to remember (not generate) the concepts that were presented in the stimulus materials. Like the comprehension task, recall of moral arguments can inventory the capacity for moral schemas. Therefore, moral recall, like moral comprehension, should produce a cumulative pattern of moral schema capacity.

Like the DIT, the recall task utilized here uses the "fragment strategy" (embedding moral reasoning arguments at different stages within the narratives) to evoke schemas (i.e., both recall and the DIT evoke schemas with short moral arguments, which do not completely spell out the whole line of reasoning distinctive for a stage). It is presumed that the fragments would elicit understanding from higher level reasoners who have developed the corresponding moral judgment schemas. As with Piaget's object concept, once the concept is present, concept evocation does not require presentation of the entire concept but merely a fragment (e.g., seeing the foot of the doll from under a blanket, the child realizes that the whole doll is under the blanket; Piaget, 1952/1963). On the DIT, if an item does not evoke the postconventional schema, then the participant does not choose that item as important; on the recall task, if a short moral argument in the moral narrative does not evoke the schema, it is less likely to be recalled.

Current Study

The current study differs from earlier studies in text comprehension and moral judgment development in several respects.

1. Type of text: Moral texts can differ in focus, for example, from examining the perspective and feelings of a character to didactic teaching of a moral rule or attitude to advocating a policy or decision, or to some combination of these. Previous research has focused on didactic stories such as fables, Bible stories, or folktales (e.g., Goldman, Reyes, & Varnhagen, 1984; Johnson & Goldman, 1987; Lehr, 1991; Stein & Trabasso, 1982); literary children's stories (Narvaez, Bentley, Gleason, & Samuels, in press); or constructed moral stories for children (Narvaez, Gleason, Mitchell, & Bentley, 1998). In this study, real-life, complex narratives were used with embedded moral reasoning at different stages of moral judgment.

2. Format: One can examine the influence of moral schemas on cognitive behavior directly or indirectly. Earlier studies in moral comprehension (Rest, 1973; Rest et al., 1969; Walker et al., 1984) have examined it directly by presenting moral reasoning in isolation and asking participants to paraphrase or recall it. In this study, moral arguments were presented in a stream of contextual detail. In other words, the participant's attention was not directed to isolated arguments. As in real life, the narratives used here intertwine events and people's rationalizations and interpretations of those events.

3. Type of task: As a new approach to studying the effects of moral schemas, participants were asked to read and recall narratives. Participants were asked not only to recall what actions generally occurred in the narrative but also what the protagonist was thinking about in the narrative. As in real life, the participant had to think over a decision situation while trying to sort out the reasoning and reconstruct what happened.

4. Type of individual difference: This is the first time that the effects of moral judgment development on moral text comprehension have been examined.

Differences in moral judgment development were expected to affect the comprehension of the moral texts in particular ways. In previous research, it has been demonstrated that familiarity with the content of a text improves recall (Chiesi et al., 1979; Crafton, 1983; Spilich et al., 1979; Taylor, 1979). Similarly, it was hypothesized that participants who on the DIT ranked Stages 5 and 6 items as highly important (implying that Stages 5 and 6 schemas were evoked) would also recall more Stage 5 reasons in the narratives (indicating that Stage 5 moral schemas were evoked during the reading recall task). It is the shift to Stage 5 thinking, measured effectively by the DIT, that was expected to be particularly evident among junior high and early college students, providing enough spread in development to detect the effects of moral schemas on moral text understanding. Research with the DIT shows that moral judgment development scores (P scores) increase in high school and especially in college (see McNeel, 1994; Rest & Narvaez, 1991; Rest et al., 1998). In this study, moral judgment development scores will be used as evidence for prior moral knowledge or schemas about how to cooperate with others. Higher moral judgment development scores indicate a facility with more kinds of moral judgment schemas. Higher scores in moral judgment on the DIT (i.e., greater preference for postconventional schemas) should be related to recall of postconventional moral arguments in the narratives because schemas for those arguments should be evoked only in those who have them.

Method

Participants

To ensure a spread in the pattern of moral judgment scores, two age levels participated in the study. Eighth-grade students (the lowest age level appropriate with the DIT)—80 from a private preparatory school and 81 from a public suburban school—represented the lower level. Sixty-two undergraduates enrolled in introductory psychology courses at a public university represented the higher level. No compensation was given to the eighth-grade participants of either school. The university students were volunteers who received extra credit for their participation. Because of incomplete protocols ($n = 12$) or failing the Consistency Check on the Defining Issues Test ($n = 44$), the final sample consisted of 63 private school students (32 females and 31 males), 55 public school students (38 females and 17 males), and 49 university students (23 females and 26 males). (See Rest, 1993, for a discussion of the DIT Consistency Check.)

Materials

Three sets of materials were used: four narratives, a set of questions for each narrative, and the DIT.

Moral narratives. The narratives were written by the researcher. Each narrative concerned a situation in which the protagonist had to consider many options before making a moral decision. The titles and topics of the narratives were (a) "Penelope and the Check": An impoverished woman wonders whether or not to keep an overpayment from an insurance company; (b) "Tom, the Manager": A manager wonders whether or not to fire his nephew, an incompetent worker; (c) "Watching the Game": Young men try to decide whether or not to sneak into a ball game for free; and (d) "Sara and the Demonstration": A woman is invited by her best friend to trespass in protest of the production of an inhumane weapon.

The texts were "moral" texts because they involved moral elements mentioned earlier: conflict over specific instances of getting along with others (moral issues) and considering the moral defensibility of alternative courses of action (moral judgment). Woven into the narratives were moral arguments considered by the story character that represented various moral judgment stages. For example, "Penelope" debates whether or not to keep the mistaken overpayment from the insurance company, considering such issues as the needs of her family (Stage 3) and the implicit agreement she has made with the company and other policyholders about whom and what the insurance money covers (Stages 4 and 5). Each narrative was written with situational detail (e.g., Tom pulls out the seventh gray hair that night) along with the character's moral arguments at different levels of Kohlberg's moral judgment stages. The arguments were based on Rest's conceptualization of Kohlberg's stages. See Appendix A for a sample narrative.

Embedded moral reasoning. The moral stage arguments in the narratives were not full-blown explications of the moral schemas (as appear in Kohlberg's moral judgment scoring guides). Partially drawn reasoning statements were used in the narratives, an approach successfully used by the DIT (Rest, Thoma, & Edwards, 1997). In other words, only fragments of moral reasoning arguments were used. For example, "Tom" mentions that part of doing his job is to fire unproductive workers. The full-blown "duty to your social role" schema with the need for order and predictability was not fully described. This "fragment strategy" approach was taken for two reasons: (a) to keep the narratives short (the narratives would not have been short if a full, complex moral argument had been developed for each argument in the middle of the narrative) and (b) to make the texts less coherent for those who do not have the corresponding moral schema. Across narratives, there were two Stage 1 arguments, three Stage 2 arguments, six Stage 3 arguments, five Stage 4 arguments, and six Stage 5 arguments. The narratives were written naturalistically, resulting in differing amounts of stage reasons in each text (see Rest, Thoma, & Edwards, 1997, for a discussion of the unimportance of counterbalanced moral stories). See Table 1 for the number and type of stage arguments by narrative. The experimenter obtained 100% agreement from an expert judge about the validity of the arguments representing the particular moral stage claimed.

Table 1
Number and Type of Moral Stage Reasoning by Story

Story	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Penelope	1	1	1	1	1
Sara	1	1	2	2	3
Tom	0	0	2	1	2
Game	0	1	1	1	0
Total	2	3	6	5	6

DIT. This is an objective, pencil-and-paper measure of moral judgment development that presents six moral dilemmas. After reading each vignette, the participant rates the importance of a list of concerns one might have in that particular situation and then ranks the four of most concern. The postconventional, or P, score is the most widely used index (Rest, 1993; Rest et al., 1998). It is a weighted sum of items in Stages 5 and 6 preferred by the participant. The score ranges from 0 to 95 and indicates the percentage of postconventional thinking preferred by the participant. Test–retest reliability for the DIT ranges between .70 and .80 for the P score. Internal consistency as measured by Cronbach's alpha has the same range (.70–.80) in various studies examining the full range of development (Rest, 1993). Cronbach's alpha for the sample in this study was .64. Evidence for validity includes studies of longitudinal trends (e.g., McNeel, 1994; Rest, 1986); cross-sectional age–education trends (e.g., Rest, 1986); correlations with theoretically similar and theoretically dissimilar constructs (e.g., Rest, 1986; Thoma & Rest, 1997); and predicted experimental manipulations (e.g., Rest et al., 1998). (For a thorough discussion see Rest, Thoma, & Edwards, 1997.) The P score from the DIT was used as a measure of moral judgment development. (Throughout the rest of this article, it is referred to as "moral judgment score.")

Procedure

Participants were tested in groups. Because of scheduling issues, groups varied in size from 2 to 88. Materials were printed on paper and distributed to individuals randomly.

Participants performed three tasks. First, four multiple-moral-stage narratives were read one after the other by the participants. The instructions were to "read each of the following stories for understanding." The order of the narratives was counterbalanced among participants and randomly assigned. Second, when they had finished reading, participants exchanged the narratives for a set of tasks and questions for each narrative in the same order in which each participant had read them. The instructions were to "complete the following tasks and questions about each of the stories." The tasks and questions were "Describe the major events of the story" and "What were the protagonist's considerations in making a decision?" Participants were given unlimited time to complete the tasks. Most finished the reading and writing tasks in less than 1 hr.

Third, after performing the narrative tasks, university participants took the DIT. The eighth-grade students took the DIT in a separate session 1 to 2 weeks later. Most students completed the DIT within 45 min.

Variables

To study the effects of prior moral knowledge, two kinds of moral reasoning responses were studied in the recall task: moral recall and moral reconstruction. Moral recall is straight, text-based recall of moral arguments at each moral stage. The participant was given credit for a paraphrase of an embedded moral argument in the narrative. Moral reconstruction is a combination of moral recall added to moral construction. Moral construction refers to participants' reasoning responses in the recall that were not in the original narrative, such as a participant attributing to a story character a Stage 3 moral argument that was not in the story. For analyses, the second dependent variable, moral reconstruction, was formed by adding together moral construction responses and moral recall.

(Separate analyses were not performed for constructed responses because there were too few of them.)

Three scores were used as independent variables: general content recall, age level, and moral judgment score. The rationale for the three independent variables is as follows. Because people differ in basic reading abilities, general content recall was entered as a control for reading ability differences. Age level (eighth grade or college) was entered as a second control variable because age-based developmental differences were expected. Finally, the moral judgment score was entered as a special measure that can tap into preexisting moral schemas.

Scoring

Content recall. General content recall was used as a measure of general recall ability, because some people have better memories for text in general (Daneman, 1991). Recall for critical and noncritical events was used as a validity check to make sure these particular narratives were processed in a manner similar to other text studies.

The narratives were each parsed into clauses that constitute events in the broad sense, using rules similar to those proposed by Warren, Nicholas, and Trabasso (1979). A scoring system was devised for the four narratives whereby general content memory was scored using a gist criterion, in which either a paraphrase or exact wording qualified as a correct answer. Scores were obtained for critical event recall and noncritical event recall. Critical events were those that were causally connected to three or more other events in the narrative. Causal connection was determined according to criteria used by Trabasso et al. (1984), that is, by being "causally necessary in the circumstances." For example, here are two events from "Penelope and the Check":

1. Penelope heard a yelp downstairs.
2. Penelope runs downstairs.

In order for Event 2 to take place in the story, Event 1 is necessary in the circumstances in that Penelope would not have run downstairs if she had not heard a cry. A link signifying a causal connection is made from Event 2 to Event 1. The causal network of the entire narrative is built up by connecting events to one another with such causal links.

Moral recall. Moral argument responses were scored in two ways. The first was text-based moral recall. A gist paraphrase of the major components of a moral argument was sufficient for credit. All five moral reasoning stages were scored. (See Appendix B for an example of moral scoring criteria.) Here is a sample excerpt from the narrative, "Sara and the Demonstration," which includes a Stage 5 argument fragment:

Sara still wavered. "I agree that each of us has to decide on what's fair. I agree that it is right to break the law sometimes, when doing so calls attention to some moral outrage."

Here is an example of a participant's response that received credit for this argument:

Sara didn't think it was a moral outrage and so it wasn't right to protest.

Moral construction. In contrast to straight, text-based moral recall, participants also provided reasons that did not appear in the stories but were part of their response to the recall task. These reasons were scored as moral constructions and were scored for all five moral stages using a gist criterion. For example, several participants wrote that Tom, the manager, was worried about getting into trouble if he did not fire his nephew, a Stage 1 argument that was not mentioned as one of Tom's concerns in the story. Constructed moral responses reflect what was activated in the participant's mind beyond text information.

Interjudge reliability. Interjudge agreement was determined both for general, nonmoral, content recall and for moral argument responses. Judges were unaware as to the classification of events as well as individuals' DIT moral judgment scores. Twenty percent of the protocols were scored by another judge. Kappa reliability computed on these protocols was .95. Disagreements on the subset were discussed and resolved. The entire group of protocols was then rescored by the researcher.

DIT scoring. Moral judgment scores were obtained from the DIT. The DITs were scored by the Center for the Study of Ethical Development, University of Minnesota.

Results

Analyses

One general form of regression equation was used with two sets of dependent variables. (Regression was used because the DIT moral judgment score is a continuous variable.) The first set involved the dependent variable moral recall and was first used as a "weak" test for schema effects. Multiple regressions were run for Stages 1 to 4 and Stage 5 alone. The second set of dependent variables, moral reconstruction (moral recall added to moral construction), was used as a "stronger" test for moral schemas; again, regression equations were run for Stages 1 to 4 and for Stage 5 separately. In sum, the analyses consisted of four multiple regressions. The main interest was in the results for Stage 5. Although one can argue that moral schemas are important when moral judgment score accounts for significant variance in text-based moral recall (the "weak" test), the evidence is much more convincing if moral judgment schemas are shown to affect moral reconstruction (the "stronger" test) because participants are then generating moral reasoning from their own schemas (prior knowledge), not only from what was given in the stimulus material.

All statistical analyses were conducted with the alpha level set at .05. There were no order effects.

Main Hypothesis

It was expected that readers with higher scores in moral judgment (higher preference for postconventional moral thinking) would reconstruct more of the Stage 5 arguments in the narratives during the recall task and that this effect would be significant beyond general content recall and age level. Regressions were performed for Stages 1 to 4 and for Stage 5, with the expectation that only Stage 5 would be significantly related to the moral judgment score after accounting for reading comprehension and age level. The results are shown in Table 2.

Moral recall. General content recall was the only factor significant in predicting scores for Stages 1 to 4. However, for Stage 5, age level was also significant ($p < .01$), and moral judgment score was nearly significant ($p = .054$) in explaining the variance.

Moral reconstruction. The same analysis was conducted on Stages 1 to 4 moral reconstruction and Stage 5 reconstruction. (Moral reconstruction is moral recall added to moral construction.) Only content recall was significant in

Table 2
Summary of Regression Analyses for Recall and Reconstruction of Stages 1 to 4 and Stage 5

Stage	Variable	*B*	*SE B*	β	*p*
		Recall			
1–4					
	Content recall	17.85000	2.6000	.5440	<.0001
	Age	−0.16500	0.3220	−.0400	
	Moral judgment	0.00500	0.0100	.0340	
5					
	Content recall	0.00310	0.0008	.2956	<.0020
	Age	0.06870	0.0252	.2147	<.0100
	Moral judgment	0.00150	0.0007	.1446	=.0540
		Reconstruction			
1–4					
	Content recall	25.37000	3.7500	.5480	<.0001
	Age	−.50900	0.4631	−.0870	
	Moral judgment	−.00005	0.0144	−.0002	
5					
	Content recall	.01720	0.0055	.2531	<.0100
	Age	.45880	0.1618	.2240	<.0050
	P score	.01200	0.0050	.1788	<.0200

predicting reconstruction of Stages 1 to 4. However, all three independent variables—general content recall ($p < .002$), age level ($p < .005$), and moral judgment score ($p < .02$)—contributed significantly to explaining the variance for Stage 5 moral reconstruction. To the extent that participants are supplying elements not given in the stories, this provides clear evidence that the Stage 5 moral schemas measured by the DIT matter beyond general recall and age level in the reconstruction of Stage 5 moral arguments.

Age-level differences. To look more closely at moral judgment schema development, age-group differences by stage (Stages 1–5) were examined. Figure 1 shows the increasing disparity in stage recall scores between the eighth graders and college students. For example, Stage 1 recall was virtually identical for the groups (eighth grade: $M = 32\%$, $SD = 32\%$; college: $M = 34\%$, $SD = 36\%$), whereas Stage 5 recall was significantly different (eighth grade: $M = 5\%$, $SD = 11\%$; college: $M = 18\%$, $SD = 18\%$). To test for significant differences between stage recall scores beyond reading comprehension differences, a multivariate analysis of variance (MANOVA) was conducted for the set of recall variables with age as a factor and general content recall as a covariate. There was a significant main effect for age, $F(5, 160) = 2.31$, $p < .047$. The only significant univariate analysis was Stage 5, $F(1, 164) = 10.01$, $MSE = .016$, $p < .002$. In addition, a MANOVA was conducted for the set of reconstruction variables (Stages 1–5) with age as a factor and general content recall as a covariate. The MANOVA indicated a significant main effect for age, $F(5, 160) = 2.80$, $p < .019$. Again, only the Stage 5 univariate analysis was significant, $F(1, 164) = 11.18$, $p < .001$. When reading comprehension was taken into account, college students were better at recalling and reconstructing the Stage 5 arguments, whereas Stages 1 to 4 were recalled and reconstructed equally well by both groups.

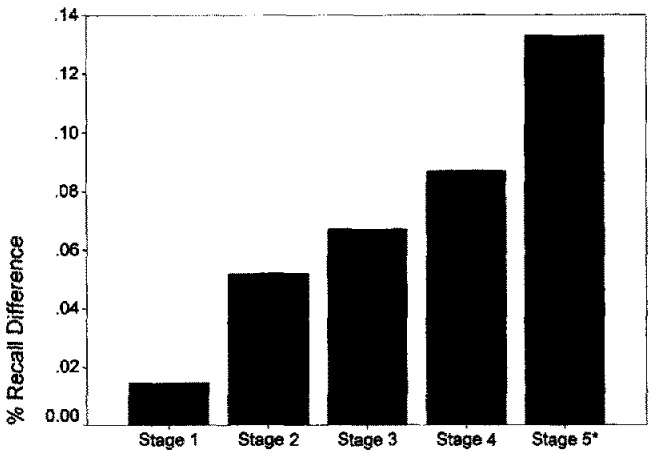

Figure 1. Difference scores between college and eighth-grade students for moral stage recall. The eighth-grade mean was subtracted from the college mean. *$p < .002$.

Validity Checks

Several analyses were conducted to ensure that the participants were representative of their age groups and that the general, nonmoral content in the narratives was recalled in a manner comparable to other studies (not bearing on the special hypotheses of this study). The results are reported in Table 3. First, as expected, older students scored significantly higher ($M = 38.3$) than younger students ($M = 27.2$) on moral judgment, as is typically found in moral judgment research (Rest, 1979, 1986, 1993). The mean for college students was slightly below the college average of 40, and for eighth graders, above the junior high average of 23.2. Scholastic achievement reading scores for the eighth-grade students indicated an above-average sample with a mean average at the 80th percentile (see Narvaez, 1993). Second, the general, nonmoral content in the narratives was processed in a manner comparable to the processing of other narratives used in text comprehension research. Participants recalled a greater amount of critical than noncritical events

Table 3
Percent Score Means (±Standard Deviations) and t Tests by Age for Moral Judgment Score, Content Recall, Critical Content Recall, Noncritical Content Recall

Variable	Eighth grade	College	*t*	*df*
P score	27.2 ± 12.9	38.3 ± 13.2	−5.02*	165
Content recall by story				
Penelope	13.7 ± 6.3	18.5 ± 7.2		
Tom	15.0 ± 6.4	22.5 ± 6.6		
Sara	14.5 ± 6.9	22.0 ± 7.7		
Game	19.9 ± 6.7	27.2 ± 9.5		
Total content recall	14.4 ± 4.6	20.7 ± 5.8	−6.75*	74.12[a]
Critical recall	29.2 ± 8.1	39.8 ± 8.7	−7.52*	165
Noncritical recall	9.3 ± 4.0	14.1 ± 5.9	−5.26*	67.55[a]

[a]Due to unequal variances as tested by Leven's test.
*$p < .0001$.

(32.3 > 10.7), comparing favorably with the findings from other studies in which readers remembered more of the critical events in a story (Fletcher & Bloom, 1988; Trabasso et al., 1984; Trabasso & van den Broek, 1985; van den Broek & Trabasso, 1986) and in which long stories were recalled (van den Broek, Rohleder, & Narvaez, 1994). Third, as in most text comprehension research, there were age differences in general content recall. The older students recalled more general content (20.7 > 14.4) and more critical events (39.8 > 29.2) than the younger participants (Casteel, 1993; van den Broek, 1988, 1989; van den Broek, Lorch, & Thurlow, 1997).

Discussion

This study examined the effects of moral schemas on the reconstruction of moral texts. Using the DIT as a corroborating measure of Stage 5 moral schemas, it examined schema effects on the recall of moral texts. The texts were created using Kohlbergian stage level arguments at different stages embedded in narratives about moral situations. The results demonstrate the interaction between moral judgment development and moral text comprehension in that prior moral knowledge played an important role in the recall of the moral texts.

Prior knowledge effects were evident in the recall and reconstruction of the Stage 5 moral arguments. It is presumed that readers who had Stage 5 conceptual structures had them stimulated by the Stage 5 fragments in the text. In other words, the fragments of moral stage ideas in the narratives evoked prior knowledge, which in turn affected the reconstruction of the text. There is support for the notion of "theme sharpening" by the high moral reasoners, demonstrated by embellishment for Stage 5, and for "theme leveling" by the low moral reasoners, through the discarding of what seemed to be irrelevant (Stage 5) material (e.g., Bartlett, 1932; Brown et al., 1977; Dooling & Christiaansen, 1977). This finding fits with the theory of "constructively responsive reading" in which prior knowledge has an influence on text processing (Pressley & Afflerbach, 1995). As readers read, they develop and refine hypotheses about the text based on what they already know about the world or the topic at hand. In the current study, readers responded constructively to the texts, although it is not clear whether this was done during reading (encoding) or when recalling (retrieval). As found with ambiguous and culturally disparate texts (Bartlett, 1932; Brown et al., 1977), readers may have misunderstood, deleted, added, or ignored information according to specific prior knowledge at the time of reading or during reconstruction.

The data provide converging evidence for the claim that moral judgment schemas exist, that they change with age and education, and that they influence the recall of moral narratives like these. There are three sources of evidence for this claim. First, individuals with higher scores in moral judgment (preference for postconventional reasoning), no matter what their age level, reconstructed Stage 5 reasoning significantly more often than those with lower scores in moral judgment, even after taking into account both age level and general content recall. This outcome points to a schematic effect. Neither general content recall nor age level were sufficient in explaining Stage 5 reconstruction.

Second, the college students recalled significantly more of the moral arguments from Stage 5 but not from Stages 1 to 4, supporting a cumulative, developmentally based moral schema pattern. They also reconstructed more arguments from Stage 5 than the younger students. Univariate MANOVA analyses indicated a gap between the recall performance of the two groups for the highest stage, reconfirming findings from moral comprehension studies that show that higher stages are increasingly difficult for lower stage reasoners. The eighth graders were significantly less facile with the presented higher stage (Stage 5) arguments. Otherwise, there would have been evidence for a developmental pattern only in reading comprehension; college students would have remembered all of the moral stage arguments better.

Third, both age groups of readers reported inferred reasoning that was not present in the stories, a phenomenon that has been found with other types of schemas in other reading research (e.g., Reynolds et al., 1982; Steffensen et al., 1979). Here, as in those studies, readers reported a representation that was constructed under the influence of prior knowledge, resulting in a distortion of the text.

The method of using moral text comprehension provides another window into the moral mind beyond the standard protocol of Piaget, Kohlberg, and other cognitive developmentalists. Moral narrative recall mimics everyday experience of hearing or reading about moral situations, recounting them, and evaluating or offering solutions. The method used in this study allows for an ecologically valid look at the real effects of moral judgment schemas on the comprehension of moral discourse.

Persuasive discourse pervades our lives: from news shows, documentaries, talk shows, political speeches, and policy discussions to lawyer arguments in a jury trial. Persuasive discourse of any kind may be understood differently by different comprehenders in correspondence to their levels of moral judgment development. In addition, much of current political and social discourse contains implicit moral reasoning. When faced with implicit or fragmented moral reasoning, moral schemas may more strongly come into play, as found with other schema effects (e.g., Bransford & Johnson, 1972; Dooling & Lachman, 1971). If a communicator is interested in composing morally persuasive communication, he or she needs to take into account the moral reasoning capacity level of the targeted comprehender.

Those who use moral stories to build moral character should be aware that children may be understanding the stories in ways different from the author's intention or the perspective of the instructor. In fact, explicit educational curricula and instruction concerning moral topics such as social behavior change (e.g., drug use prevention or abuse recovery) may not be properly understood if the moral judgment capacities of the audience are not accommodated. Just as teachers attempt to match the reading level of a text with the student's level of reading skill, moral and social education programs should attempt to match the moral

reasoning level of a text with the student's level of moral reasoning.

In short, the development of "moral literacy" is more complicated than often believed. Merely reading moral stories to children is unlikely to be enough for them to understand the intended message.

References

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 225–291). New York: Longman.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17,* 1–12.

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal, 14*(4), 367–381.

Bartlett, F. C. (1932). *Remembering.* Cambridge, England: Cambridge University Press.

Bennett, W. (1993). *The book of virtues.* New York: Simon & Schuster.

Bennett, W. (1995). *The children's book of virtues.* New York: Simon & Schuster.

Bobrow, D., & Norman, D. (1975). Some principles of memory schemata. In D. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 131–149). New York: Academic Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

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.

Casteel, M. (1993). Effects of inference necessity and reading goal on children's inferential generation. *Developmental Psychology, 29*(2), 346–357.

Chiesi, H., Spilich, G., & Voss, J. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 257–273.

Colby, A., & Kohlberg, L. (1987). *The measure of moral judgment, vol. 1: Theoretical foundations and research validation.* Cambridge, England: Cambridge University Press.

Crafton, L. K. (1983). Learning from reading: What happens when students generate their own background information? *Journal of Reading, 26,* 586–593.

Cunningham, A. E., Stanovich, K. E., & Wilson, M. R. (1990). Cognitive variation in adult college students differing in reading ability. In T. H. Carr & B. A. Levy (Eds.), *Reading and its development: Component skill approaches* (pp. 129–159). New York: Academic Press.

Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research, vol. 2* (pp. 512–538). New York: Longman.

Dooling, J. D., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 428–436.

Dooling, J. D., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology, 88,* 216–222.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27,* 235–244.

Gernsbacher, M. A. (Ed.). (1994). *Handbook of psycholinguistics.* New York: Academic Press.

Goldman, S. R., Reyes, M., & Varnhagen, D. (1984). Understanding fables in first and second languages. *NABE Journal, 8*(2), 35–66.

Harris, R. J., Lee, D. J., Hensley, D. L., & Schoen, L. M. (1988). The effect of cultural script knowledge on memory for stories over time. *Discourse Processes, 11,* 413–431.

Johnson, D. F., & Goldman, S. R. (1987). Children's recognition and use of rules of moral conduct in stories. *American Journal of Psychology, 100,* 205–224.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95,* 163–182.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago: Rand McNally.

Kohlberg, L. (1984). *The psychology of moral development: The nature and validity of moral stages.* New York: Harper & Row.

Lehr, S. (1991). *The child's developing sense of theme.* New York: Teachers College Press.

McNamara, T. P., Miller, D. L., & Bransford, J. D. (1991). Mental models and reading comprehension. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research, vol. 2* (pp. 490–511). New York: Longman.

McNeel, S. P. (1994). College teaching and student moral development. In J. R. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 27–50). Hillsdale, NJ: Erlbaum.

Narvaez, D. (1993). High achieving students and moral judgment. *Journal for the Study of the Gifted, 16*(3), 268–279.

Narvaez, D., Bentley, J., Gleason, T., & Samuels, J. (in press). Moral theme comprehension in third graders, fifth graders, and adults. *Reading Psychology.*

Narvaez, D., Gleason, T., Mitchell, C., & Bentley, J. (1998). *Measuring moral thinking in children using moral theme comprehension.* Manuscript submitted for publication.

Nelson, K. (1986). Event knowledge and cognitive development. In K. Nelson (Ed.), *Event knowledge: Structure and function in development.* Hillsdale, NJ: Erlbaum.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language, 24,* 59–88.

Piaget, J. (1963). *The origins of intelligence in children.* New York: W. W. Norton. (Original work published 1952)

Piaget, J. (1965). *The moral judgment of the child* (M. Gabain, Trans.). New York: Free Press. (Original work published 1932)

Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology, 69*(4), 309–315.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading.* Hillsdale, NJ: Erlbaum.

Rest, J. R. (1973). The hierarchical nature of moral judgment. *Journal of Personality, 41,* 86–109.

Rest, J. R. (1979). *Development in judging moral issues.* Minneapolis: University of Minnesota Press.

Rest, J. R. (1986). *Moral development: Advances in research and theory.* New York: Praeger.

Rest, J. R. (1993). *Guide for the DIT.* Minneapolis, MN: Center for the Study of Ethical Development.

Rest, J. R., & Narvaez, D. (1991). College experience and moral development. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development, vol. 2: Research* (pp. 229–245). Hillsdale, NJ: Erlbaum.

Rest, J. R., Narvaez, D., Bebeau, M., & Thoma, S. (1998). *Moral judgment development in adolescents and adults: A neo-Kohlbergian approach based on the Defining Issues Test.* Manuscript submitted for publication.

Rest, J. R., Thoma, S., & Edwards, L. (1997). Designing and validating a measure of moral judgment: Stage preference and stage consistency approaches. *Journal of Educational Psychology, 89,* 5–28.

Rest, J., Turiel, E., & Kohlberg, L. (1969). Level of moral development as a determinant of preference and comprehension of moral judgment made by others. *Journal of Personality, 37,* 225–252.

Reynolds, R., Taylor, M., Steffensen, M., Shirey, L., & Anderson, R. (1982). Cultural schemata and reading comprehension. *Reading Research Quarterly, 17,* 353–366.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 265–303). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Matague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, and goals.* Hillsdale, NJ: Erlbaum.

Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479–516). New York: Academic Press.

Spilich, G., Vesonder, G., Chiesi, H., & Voss, J. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 275–290.

Steffensen, M., Joag-Dev, C., & Anderson, R. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly, 15,* 10–29.

Stein, N., & Trabasso, T. (1982). Children's understanding of stories: A basis for moral judgment and dilemma resolution. In C. J. Brainerd & M. Pressley (Eds.), *Verbal processes in children, vol. 2* (pp. 161–188). New York: Springer-Verlag.

Sulin, R. A., & Dooling, D. J. (1974). Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology, 103,* 255–262.

Taylor, B. M. (1979). Good and poor readers' recall of familiar and unfamiliar text. *Journal of Reading Behavior, 11,* 375–380.

Thoma, S. J., & Rest, J. R. (1997). *The relationship between decision-making and patterns of consolidation and transition in moral judgment development.* Manuscript submitted for publication.

Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (p. 86). Hillsdale, NJ: Erlbaum.

Trabasso, T., & Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language, 24,* 595–611.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of causal relations in stories. *Discourse Processes, 12,* 1–12.

Van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language, 27,* 1–22.

Van den Broek, P. (1989). Causal reasoning and inference making in judging the importance of story statements. *Child Development, 60,* 286–297.

Van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). New York: Academic Press.

Van den Broek, P., Lorch, E., & Thurlow, R. (1997). Children and adults' memory for television stories: The role of causal factors, story-grammar categories and hierarchical level. *Child Development, 67,* 3010–3028.

Van den Broek, P., Rohleder, L., & Narvaez, D. (1994). Cognitive processes in the comprehension of literary texts. In H. van Oostendorp & R. Zwaan (Eds.), *Naturalistic text comprehension, vol. LIII: Advances in discourse processes* (pp. 229–246). Norwood, NJ: Ablex.

Van den Broek, P., & Trabasso, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes, 9,* 1–15.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Walker, L. (1988). The development of moral reasoning. *Annals of Child Development, 5,* 33–78.

Walker, L. (1991). Sex differences in moral reasoning. In W. Kurtines & J. Gewirtz (Eds.), *Handbook of moral behavior and development, vol. 2: Research* (pp. 333–364). Hillsdale, NJ: Erlbaum.

Walker, L., deVries, B., & Bichard, S. L. (1984). The hierarchical nature of stages of moral development. *Developmental Psychology, 20,* 960–966.

Warren, W. H., Nicholas, D. W., & Trabasso, T. (1979). Event chain and inferences in understanding narratives. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 23–52). Hillsdale, NJ: Erlbaum.

Whitney, P., Budd, D., Bramucci, R. S., & Crane, R. S. (1995). On babies, bath water, and schemata: A reconsideration of top-down processes in comprehension. *Discourse Processes, 20,* 135–166.

Appendix A

Sample Story With Moral Argument Fragments Marked

"Tom, the Manager"

Tom was having another sleepless night. He quietly crawled out of bed for the fourth time, grabbed his robe, and walked to the bathroom. As he stared at the exhausted face in the mirror, he thought about his dilemma for the 100th time that night.

He was the manager of a store that was part of a department store chain with branches all over the country. It was challenging work that he greatly enjoyed. He had given a job to his nephew, Freddie, because Freddie's mother, Tom's sister, had pleaded with him to do so. Freddie had been in a lot of trouble at school and needed a fresh start. Tom liked his older sister and wanted to help her out. He remembered the times when she had helped him: selecting clothes to wear for a special date, advising him on how to study for the tests of notoriously difficult teachers, and attending every wrestling match in which he had competed. These days, their families got together every Sunday for dinner and an afternoon of sports television.

After about 3 weeks, Tom noticed that Freddie was not doing a good job. Freddie seemed unwilling to do anything. He wouldn't stock the shelves, he was reluctant to help customers, he would even arrive late.

Several times, Tom had tried to straighten Freddie out. He gave him special instructions and encouragement about how to do his job. He gave him tips on helping customers, techniques for stocking shelves, shortcuts on doing inventory. But nothing seemed to help. Freddie didn't change. The only thing that Freddie had ever done right was when he had been sent on a special errand to pick up a regional director from the airport. But that was back in the first week Freddie was employed.

Tom pulled out a gray hair. It was the seventh one tonight. *On the one hand, he hated to fire his sister's child. He was afraid that such an action would strain the relationship with his sister* [Stage 3] and hurt Freddie's chances for success even more. In Tom's family, they always took care of their own. And, if he were to fire Freddie, it would be next to impossible to talk to or to ever visit his sister's family.

What had been keeping him tossing and turning sleeplessly every night for the last 2 weeks, however, was his feeling of responsibility to the business as well. *Part of doing his job was to fire unproductive workers. Although he was working for a big company, it still wouldn't be fair for the company to be paying Freddie for doing a poor job.* [Stage 4] *If this sort of thing got out of hand, the whole company could potentially lose scores of customers and loads of money.* [Stage 3]

Tom went to the den and sank into the easy chair. As he pondered the difficulty, it was clear that no matter what he decided, he wouldn't be able to make everyone happy. Tom realized that *his position in the company and his special responsibilities were designed to create the greatest benefit for the most people: the customers, the workers, and the investors. It was a necessary part of his job to rid the company of unproductive employees. It was a policy that he fully supported.* [Stage 5] In fact, Tom would not want to work for a company that didn't have work standards. Furthermore, Tom was working on the assumption that he and everyone else in the company tacitly agreed with this policy.

But Tom realized that Freddie was not a rotten boy. Maybe Freddie was a kid who needed a little more time and a little more support in order to straighten out. Tom had been hoping that a portion of tolerance on his part temporarily might bring about a fundamental change in the boy. *Turning a young person around was an investment that his business, as well as the society at large, should support. It wasn't just that Freddie was his sister's child; Tom would be willing to make a similar investment in any young person he thought he could help change for the better.* [Stage 5]

The cat jumped in his lap and settled in for a snooze. Her purring soothed him. As his body relaxed a bit, Tom continued to consider the options. Toward morning he was able to make a decision with which he could live.

Appendix B

Scoring Criteria for Moral Arguments in "Tom, the Manager"

#T1—Con: Tom's Family Bonds (Stage 3)

Score *hit* if student (S) (a) mentions family closeness or affection, helping Freddie, or family upset as reason for not firing Freddie or (b) argues against this. Example of a hit: "But also felt that if he did so it would put a strain on the relationship with his family and sister." Example of a miss: "He was close to his family."

#T2—Pro: Might Hurt the Company (Stage 3)

Score *hit* if S (a) mentions that damage might be done to the company, such as loss of money and customers, (b) is concerned about the prevention of specific damage, (c) mentions the need for productivity–efficiency, (d) mentions Tom's sense of loyalty to his job. Example of a hit: "He had a responsibility as manager of the store. The company was losing money by paying an employee that couldn't do the job." Example of a miss: "His job is on the line."

#T3—Pro: Against the Rules (Stage 4)

Score *hit* if S mentions (a) that it is Tom's special responsibility (his job) as manager to fire anyone who isn't working properly (company rules), (b) the widespread damage if everyone did a bad job, or (c) the inverse. Example of a hit: "Tom doesn't want to fire him, but it is his job to fire bad employees." Example of a miss:

"Plus the store really needs it and he was putting his job on the line."

#T4—Pro: Social Contract With the Company (Stage 5)

Score *hit* if S mentions that because Tom wants to have this general policy, he in effect has a tacit agreement with the company and the other workers to abide by this agreement. (So it is not just that there is a company rule but rather Tom's support of that policy—his consent that it is a good rule—that makes it binding.) Example of a hit: "He firmly believed in the worker policy of the company and he supported this along with other workers." Example of a miss: "What his job meant."

#T5—Con: General Priority of Long-Term Human Welfare Over Short-Term Institution Needs (Stage 5)

Score *hit* if S (a) clearly distinguishes Tom's family bonds from a general value to put long-term human welfare over short-term institutional needs (S must make it clear that Tom would be acting on a general principle that applied to anyone, not just to Freddie [special relations]), or (b) argues against this. Example of a hit: "He also said that he would have given the second chance to any other youth in the same position."

Received March 2, 1997
Revision received August 7, 1997
Accepted August 7, 1997 ■