The Relation of Moral Judgment Development and Educational Experience to Recall of Moral Narratives and Expository Texts

DARCIA NARVAEZ
University of Notre Dame

TRACY GLEASON
Wellesley College

ABSTRACT. Moral text processing was used as an ecologically valid method for assessing implicit and explicit moral understanding and development. The authors tested undergraduates, seminarians, and graduate students in political science and philosophy for recall of moral narratives and moral expository texts. Multivariate analyses of covariance using educational experience as an independent variable, age and moral judgment score as covariates, and recall of embedded moral arguments as a dependent variable revealed a relation between education and level of moral arguments recalled. Lower stage moral reasoning was best recalled by undergraduates, whereas higher stage reasoning was best recalled by graduate students, with seminarians intermediate for both types of text. Moral judgment score was related to recall of the highest level moral arguments even when age and educational experience were controlled. Moral judgment development appeared to be particularly helpful in recall of expository compared with narrative texts.

Keywords: expertise, moral development, moral judgment

MOST KNOWLEDGE CANNOT BE EXPLAINED because it is implicit rather than explicit (Hogarth, 2001; Keil & Wilson, 1999). This fact is apparent in moral judgment research. Historically, researchers have studied moral reasoning at the two extremes of understanding. Production tasks such as Kohlberg's Moral Judgment Interview (MJI; Colby et al., 1987) measure knowledge at the high end of understanding, assessing what respondents can express on their own. The MJI requires that participants demonstrate explicit conceptual understanding as well as skills of articulation and persuasive discourse. These capacities are found almost exclusively in highly educated and articulate persons with specialized knowledge (Kohlberg, 1984). In contrast, recognition measures of moral judgment, such as the Defining Issues Test (DIT; Rest, 1979, 1986; Rest,

Narvaez, Bebeau, & Thoma, 1999), assess knowledge at the low end of understanding, providing maximum scaffolding for the respondent. Recognition tasks tap into implicit knowledge and emerging conceptual structures as well as more established understanding (Reber, 1993). Although both types of measures assess basic understanding, an ideal measure would illustrate the full range of moral reasoning available in a person's moral schemas. In recent years, moral text processing has emerged as an alternative method of study in the examination of moral development (Narvaez, 1998, 1999; Narvaez, Endicott, & Thoma, 2001; Narvaez, Lapsley, Hagele & Lasky, 2006).

Moral text processing, such as the recall of narratives, has the potential to measure the full range of understanding, from implicit to explicit. In reading, background knowledge is required to form a mental model of the text (Singer, 1994), and readers compensate for unspecified or vague information in a text with their knowledge of the world (e.g., Bransford & Johnson, 1972). In addition, readers must repackage story events when they summarize, recall, or answer questions about the text on the basis of the mental representation they have constructed by integrating text material with background knowledge (Kintsch, 1988). This process of integrating and repackaging has components of implicit knowledge and explicit, declarative knowledge. Like the DIT, text recall provides a framework for structuring information about a moral issue, and like the MJI, it requires some production on the part of the respondent as recollection is organized. However, it also demands some inference generation on the basis of background knowledge.

Moral text processing provides an avenue for examining moral reasoning in a more ecologically valid manner than do the MJI and DIT because reading is a task that resembles everyday discourse processing. For example, in conversations and in reading or hearing news reports, individuals are often confronted with partial information and partial arguments about rationales and courses of action. They filter this information and, when necessary, apply background knowledge to fill the gaps. Individuals apply schemas, including moral schemas, when making meaning out of events (Narvaez, 2002; Narvaez & Bock, 2002). Consequently, the processing of a text on moral topics is likely to differ as a function of a person's moral reasoning ability and experience with particular moral problems. Individual differences in the processing of moral texts should thus be apparent on the basis of both development and experience in the moral domain.

Experience contributes to differences in domain expertise between experts and novices. Domain expertise generally refers to a specific, studied domain (Alexander, 1992) for which expertise may take something like 10,000 hours of study (Simon & Chase, 1973). *Expertise* refers to a set of conceptual associations, action skills, and conditional knowledge (Abernathy & Hamm, 1995; Sternberg,

Address correspondence to Darcia Narvaez, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556, USA; dnarvaez@nd.edu (e-mail).

1998, 1999). Domains that have been examined by psychologists include well-structured domains such as baseball and chess (e.g., Chase & Simon, 1973; Chi & Koeske, 1983) and ill-structured domains such as medical diagnosis (Johnson, Hassebrock, Duran, & Moller, 1982). Well-structured domains have components that are completely specified in terms of information, possible actions, and outcomes (e.g., baseball), whereas ill-structured domains are characterized by uncertainty regarding how to characterize the initial starting conditions or problem, the goodness or feasibility of operations that can be used, and the goals to be achieved and their inherent goodness (e.g., problems in creative arts; Newell & Simon, 1972). Reasoning about moral issues is regarded as an ill-structured domain (King & Kitchener, 1994) because although experts may agree on a general interest (e.g., to select an ethical decision in a particular circumstance), they will likely not agree on what the problem is (starting conditions), what tools to use to reach a decision (operations used), what a satisfactory answer may be (evaluative functions), or how to determine whether the problem has been solved (goals achieved). Indeed, hospital bioethicists are reported to have these difficulties (Toulmin, 1981).

Reading tasks are particularly appropriate for investigating an ill-structured domain such as moral reasoning because of the relation between understanding a text and domain knowledge. Text comprehension involves not only the nature of the text and reading abilities of the reader but also reader familiarity with the text topic (e.g., Chiesi, Spilich, & Voss, 1979; Fincher-Kiefer, Post, Greene, & Voss, 1988; Spilich, Vesonder, Chiesi, & Voss, 1979; see reviews by Alexander, 1992; Roller, 1990). Text comprehension is greatly influenced by the congruity between reader background and specific text content (e.g., Ohlhausen & Roller, 1988), which is facilitated by a greater amount of knowledge considered analogous to subject matter knowledge (Alexander, Pate, & Kulikowich, 1989; Hayes & Tierney, 1982; Kulikowich & Alexander, 1990; Walker, 1987) and expertise in the subject of the text (Meutsch, 1989). Of course, text comprehension is also related to the educational background of the reader (e.g., Anderson, Reynolds, Schallert, & Goetz, 1977; Birkmire, 1985), but beyond education, expertise differences in a particular domain have been directly related to differences in the ability to make inferences and construct-relevant schematic and conceptual models of text events (Singer, Harkness, & Stewart, 1997; Spilich et al.).

Researchers have explained the processes that underlie the relation between text comprehension and domain knowledge as a function of mental schemas and have tested these processes through examination of readers' recall of domain-specific texts (e.g., Arbuckle, Vanderleck, Harsany, & Lapidus, 1990; Schneider, Körkel, & Weinert, 1989). The findings of these studies suggest that high-knowledge readers achieve a deeper level of understanding than do low-knowledge readers, enabling them to construct an appropriate mental model that allows them to correctly elaborate on the text. In contrast, low-knowledge readers form inadequate mental models of the text, which leads to erroneous elaborations and inferences during recall

(Moravcsik & Kintsch, 1993). When texts are inconsistent with readers' activated knowledge structures, readers will understand poorly (Bransford & Johnson, 1972), misrecall (Steffensen, Joag-Dev, & Anderson, 1979), and even distort memory to fit with their schematic structures (Bartlett, 1932; Reynolds, Taylor, Steffensen, Shirey, & Anderson, 1982). Similarly, recall of a text is superior when the reading topic is familiar (Crafton, 1983; Taylor, 1979) or when it conforms to background knowledge (Chiesi et al., 1979; Spilich et al., 1979).

The use of text recall to test application of domain knowledge in moral reasoning has precedence in the literature. Narvaez (1998) found that readers of moral narratives remembered correctly but also distorted narratives in their recall according to their level of moral judgment development. Although all participants reconstructed moral arguments that were not in the texts during recall, those with higher scores in moral judgment both recalled and invented significantly more high-level moral arguments than did those with lower levels of moral judgment development. Consequently, in the present study, we hypothesized that participants with greater and more focused experience with moral reasoning and rational moral problem solving would recall more of the high-stage moral arguments in moral texts than would participants with less educational experience.

Researchers have mapped developmental differences in moral judgment among students from elementary school through higher education, including undergraduate and graduate students. Age and education are the strongest predictors of moral judgment development (Rest, 1979, 1986; Rest et al., 1999). In fact, researchers have used graduate students in philosophy or political science as proxies for experts in moral reasoning and have used seminarians as middle-level experts (Rest, 1979, 1986). Yet, whether developmental differences alone account for the distinctions between undergraduate and graduate students is unclear because of other differences between the groups, such as age, maturity, and academic experience. We examined influences on moral text recall that included not only moral reasoning but also age and educational experience to see how these factors interacted in explaining recall. We expected group differences due to educational experience, age, and moral reasoning expertise.

The Current Study

We tested whether moral judgment, age, and level of educational experience with moral reasoning and rational, moral problem solving would be reflected in participants' recall for moral stories, including both the events of the story and the moral arguments contained within them. We used moral judgment score as a proxy for moral reasoning expertise and education group or level as a proxy for experience with the tools of moral reasoning. Age was representative of maturation and correlative life experience.

Across stories, we expected that participants with higher moral reasoning expertise would exhibit better recall for the highest stage moral arguments than

would those with low moral reasoning expertise, regardless of age and educational experience. Specifically, compared with low-reasoning participants, we expected high-reasoning participants to better recall the postconventional arguments in stories. We hypothesized that the advantage of moral judgment development would remain even after controlling for age and educational experience.

We used both narrative and expository texts. *Narrative texts* are stories with a beginning, middle, and end that discuss characters, their internal states, goals, and actions and reactions to outcomes that ensue. *Expository texts* are informational texts that convey facts or procedures. We used recall of narrative texts because it is a common form of processing world events (Bruner, 1986; Vitz, 1990), especially moral events (Tappan, 1997), and because narratives are familiar to all participants. However, in daily life, many moral issues are presented and considered in expository rather than narrative form. Consequently, we sought to determine whether recall of expository texts, compared with narrative texts, would vary according to moral judgment development, age, or educational experience.

Although reading processes are similar for narrative and expository texts (Goldman & Varma, 1995; van den Broek, Rohleder, & Narvaez, 1994, 1996), there are systematic differences in how people respond to each type of text (e.g., Einstein, McDaniel, Owen, & Coté, 1990; McDaniel, Einstein, Dunay, & Cobb, 1986; Zwaan, 1994). For example, readers generally do not use comprehension-enhancing strategies with expository texts as easily or as automatically as they do with narratives (Narvaez, van den Broek, & Ruiz, 1999). Consequently, we expected diminished recall for expository texts compared with narrative texts for all readers. However, we expected that recall of the moral expository texts would vary as a function of moral reasoning expertise because previous findings suggest that individuals with more extensive prior knowledge relevant to the text demonstrate more inference generation, hypothesis construction, and other active processes when reading (Haas & Flower, 1988). We also expected that educational experience would correlate with recall of expository texts because these texts require more comprehension effort and background knowledge than do narratives.

We included three groups of participants that varied on the three factors (age, educational experience, moral judgment development). Undergraduate students were the youngest participants and had the least experience in moral reasoning (no philosophy or political science majors were included in the sample). They tended to have lower moral judgment scores than did participants with bachelor's degrees or graduate training. Lutheran seminarians represented a higher educational experience group, were older, and were expected to have intermediate-level moral judgment. Although in their training the seminarians focus on moral questions to a greater degree than do undergraduates, they also learn a particular moral point of view (the Lutheran tradition) rather than learning the complex problem-solving tools in moral reasoning, which includes learning the strengths and weakness of various approaches to solving moral problems. In contrast, graduate students in

philosophy or political science learn the complexities of moral reasoning, and they comprised our third group. They were similar in age to the seminarians, represented a relatively high level of educational experience and the greatest amount of moral reasoning training, and therefore were expected to have the highest moral judgment scores. Admittedly, age and educational experience were confounded to some degree between the undergraduates and the other two groups. We interpreted the results in light of this potential confound.

Method

Participants

Participants were 37 undergraduates from lower division psychology classes in a Midwestern-U.S., public, urban university (M age = 22.19 years, SD = 13.82 years; median = 20 years; 11 men; all White), 34 seminary students from a mainstream Lutheran seminary (M age = 29.97 years, SD = 14.72 years; median = 28 years; 17 men; all White), and 16 graduate students in philosophy or political science at a large Midwestern-U.S. public, urban university (M age = 29.06 years, SD = 10.71 years; median = 27 years; 10 men; all White). The undergraduates received course credit for participating, and the seminary and graduate students were paid for their participation.

Materials

Narratives. The first portion of the study involved reading and recalling five texts: two moral narratives, two moral expository texts, and one nonmoral narrative. The moral narratives were written by the researchers, used in prior research (Narvaez, 1998), and concerned everyday situations (being asked by a friend to do something illegal; receiving undeserved money) in which the protagonist considered action options while trying to make a moral decision. The moral dilemma situations included situational detail along with embedded moral arguments at different levels of Kohlberg's (1984) moral judgment stages (Stages 2–5). The arguments were based on Rest's (1979) conceptualization of Kohlberg's stages. These narratives used partially drawn arguments, an approach successfully used by the DIT (Rest, 1993; Rest, Thoma, & Edwards, 1997) and in previous moral text comprehension research (Narvaez, 1998). The partially drawn moral arguments necessitate reader application of background knowledge to fill in missing information. To do so correctly, the individual must have the conceptual structures required, otherwise the fill-in may be incorrect or may not occur. We presumed that existing moral reasoning schemas would be invoked under these partial-view conditions in the texts. The moral narratives were (a) "Sara's Evening at Home," in which a woman is invited by her best friend to trespass in protest of the production of an inhumane weapon (see Appendix A for a copy of this story

with arguments) and (b) "Penny and the Mail," in which a poor woman with three children and a hardworking husband receives an overpayment from the insurance company that would not be discovered if she kept it. The nonmoral story was "Tom's Week," a story developed by the researchers and used in prior research that described the protagonist's grueling week.

Expository texts. We selected two expository texts from the local newspaper that presented multiple moral arguments at different moral stages. One was a newspaper editorial about euthanasia called "Life and Death" (see Appendix B for the text with arguments). The other was a report called "Sentence Closes Case" about a man turning himself in for the murder of his neighbors 20 years after the fact. This text was a slightly revised newspaper report about the case.

The Defining Issues Test. The second set of materials was the DIT (Rest, 1979, 1986; Rest et al., 1999), a standardized, objective, paper-and-pencil measure of justice-based moral judgment that presents six moral dilemmas. After reading each dilemma, the participant rates the importance of a list of concerns one may have in that situation and ranks the four of most concern. The postconventional or *P score* is the most widely used index on the DIT (Rest, 1993). It is a weighted sum of the postconventional judgment preferred by the participant (i.e., Stages 5 and 6 in the Kohlbergian scheme). The maximum score is 95. Items on the DIT are actual statements from respondents during test development. Examples from a dilemma called "Heinz and the Drug" include "What values are going to be the basis for governing how people act towards each other?" (Stage 5, Postconventional); "Whether a community's laws are going to be upheld" (Stage 4, Maintaining Norms); and "Isn't it only natural for a loving husband to care so much for his wife that he'd steal?" (Stage 3, Personal Interest).

Test–retest reliability for the DIT ranges from .70 to .80 for the P score. Internal consistency as measured by Cronbach's alpha had the same range, .70–.80, in various studies (Rest, 1993). The DIT has been validated according to six criteria: (a) correlations with moral comprehension (e.g., Rest, 1979), (b) differentiating more and less skilled groups (Narvaez, 1998), (c) longitudinal trends (e.g., Rest, 1986), (d) sensitivity to intervention (Rest & Narvaez, 1994), (e) correlations with political attitudes (e.g., Narvaez, Getz, Rest, & Thoma, 1999), and (f) correlations with behavior (e.g., Thoma, 1994; see Rest et al., 1997, or Rest et al., 1999, for a summary of validation studies). The DIT was scored electronically by the Center for the Study of Ethical Development, University of Minnesota.

Scoring

Story events. The stories were each parsed into clauses that constitute events in the broad sense, using rules similar to those proposed by Warren, Nicholas, and Trabasso (1979). We obtained scores for gist recall (getting the general idea of an event) of critical and noncritical events. *Critical events* were those that were

causally connected to three or more other events in the story. *Noncritical events* were those with fewer than three connections to other events in the story. Causal connection was determined according to criteria used by Trabasso, Secco, and van den Broek (1984): by being causally necessary in the circumstances. For example, in the events listed below from "Sara's Evening at Home," Sara's invitation to Cindy to "come in" is causally dependent in the text on Cindy being at the door, hence a causal link is drawn from the item listed here as item 2 to item 5, and from item 3 to item 5.

1. As she was beginning her second bowl of popcorn,
2. the doorbell rang.
3. It was Cindy.
4. "Hey, buddy! How's it going?"
5. "Come on in!"

An entire network was constructed in this way for each story.

Moral arguments. Moral arguments of different stages were in each moral text. Moral argument recall, like nonmoral event recall, was scored using a gist criterion. In other words, a paraphrase of the major components in a moral argument was sufficient for credit. A sample excerpt from "Sara's Evening at Home" that includes a Stage 5 argument fragment is, "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 . . .'"

An example of a participant's response that received credit for the above argument is, "Sara didn't think it was a moral outrage and so it wasn't right to protest." During recall, participants brought up arguments from all stages that were not written in the stories, including Stages 1 and 2. After determining the types of arguments brought up by participants, we included them in the scoring system and scored all protocols for these categories. Table 1 lists the number of

TABLE 1. Number of Moral Arguments Written Into Each Story at Each Moral Stage, by Story Type

Story type	Moral stage					Total
	1	2	3	4	5	
Narrative	3	3	3	4	4	17
Penny	1	1	1	1	2	6
Sara	2	2	2	3	2	11
Expository	1	7	8	10	4	30
Sentence	0	5	3	4	0	12
Life	1	2	5	6	4	18
Total	2	9	11	13	8	47

stage arguments that were scored for each story. The percentages of moral events in stories are listed in Table 2. In the analyses, we standardized participant scores because stories had different numbers of events to recall.

Reliability. Each story had a particular number of events that could be recalled (Tom: 125; Sentence: 95; Sara: 127; Life: 51; Penny: 112). We scored each protocol for either recalling or not recalling each event in each story. To verify that recall scores were reliably scored, we performed two reliability estimates for each story: general (critical and noncritical) event recall and moral argument recall. We evaluated both types of reliability across stories in a random sample of 10 protocols scored by two judges. For the nonmoral story, "Tom's Week," κ = .93 for event recall. For "Penny and the Mail," κ = .85 for event recall and κ = .94 for moral argument recall. For "Sara's Evening at Home," κ = .83 for event recall and κ = .92 for moral argument recall. For "Life and Death," κ = .86 for event recall and κ = .81 for moral argument recall. For "Sentence Closes Case," κ = .83 for event recall and κ = .85 for moral argument recall. All disagreements were settled by discussion and protocols were then rescored.

Internal consistency. We calculated internal consistency using Cronbach's alpha for the recall of each story. Cronbach's alphas for the narratives were .91 for Tom, .84 for Sara, and .88 for Penny. Cronbach's alphas for recall of the expository texts were .78 for Life and .87 for Sentence.

Variables

Dependent variables were recall of general (moral and nonmoral) events in stories and recall of moral arguments for each Stage (1–5). Factors were educational experience and gender. Age, moral judgment (P) score, and nonmoral recall were covariates. Nonmoral story recall was a measure of verbal production and was used in the analyses as a covariate to control for output quantity.

TABLE 2. Number of Total Events and Number of Moral Events in Each Story

Story type	Number of events	Number of moral events
Narrative		
Penny	112	23
Sara	127	48
Expository		
Sentence	95	49
Life	51	23
Total	385	143

Procedure

Participants were tested alone or in groups of 2 to 15. First, the participants were told to "read each of the following stories for understanding." All participants read the stories in the same order ("Tom's Week," "Sentence Closes Case," "Sara's Evening at Home," "Life and Death," and "Penny and the Mail"). When they finished reading all stories, participants exchanged the stories for a set of questions that related to the stories in the same order in which they were read. The instructions were to "complete the following tasks and questions about each of the stories": "Describe the major events of the story" (evoking text-based memory as mental structures allow) and "What were the protagonist's considerations in making a decision?" (evoking the moral schemas used to process the arguments in the text). After performing the story task, participants completed the DIT. Participants were given unlimited time to complete the tasks, and most finished both tasks in less than 90 min. The order of materials was fixed. We placed the DIT at the end so as to not prime the story task for moral reasoning. The neutral story was placed first to acquaint participants with the task. The order of the rest of the stories was determined randomly except for alternating narrative and expository texts. In prior studies of moral texts, order effects have not been found (e.g., Narvaez, 1998).
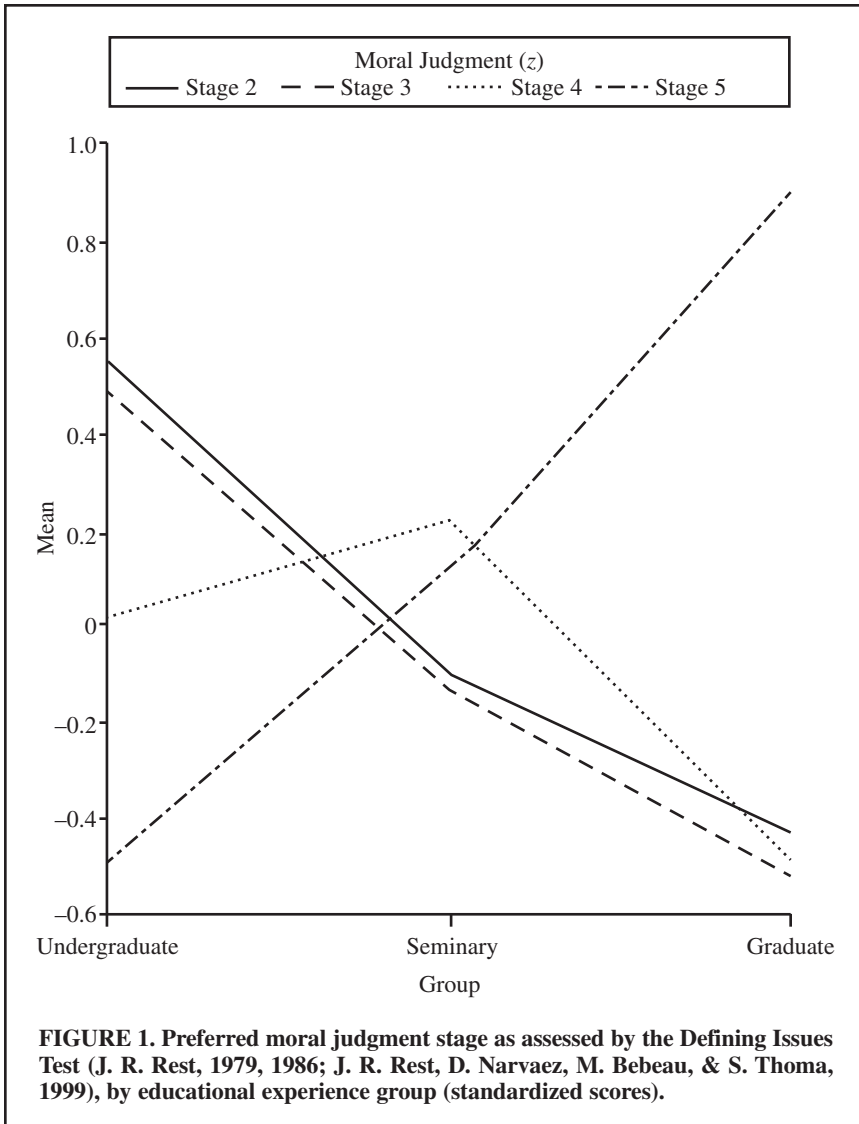
Results

Moral Judgment Scores and General Recall of Texts

We set alpha at .05 for all analyses, and all tests were two-tailed. DIT P scores indicated moral judgment differences between the education groups. Undergraduates had an average score of 43.82 (SD = 13.82), which is at the high end of the norm for college students (Rest, 1993). Seminarians had an average score of 51.61 (SD = 14.75), which is in the range for graduate students (Rest, 1993). Graduate students had an average score of 61.99 (SD = 10.71), slightly lower than the normed score of 65 for graduate students in philosophy and political science (Rest, 1993). A one-way analysis of variance (ANOVA) indicated that the difference among the groups was significant, $F(2, 82) = 9.61$, $p < .0001$. Post hoc tests using Tukey's honestly significant difference (HSD) revealed significant differences among all three groups ($p < .05$) . In general, these scores demonstrate that undergraduates preferred Stage 2 arguments, seminarians Stage 4, and graduate students Stage 5 (see Figure 1).

As is normally the case, we found considerable individual differences within groups. Consequently, we used moral judgment score as a covariate in the analyses. In this way, the effects of education on recall could be weighed separately from moral judgment development.

General recall means for each story are listed in Table 3. We conducted a repeated-measures multivariate analysis of variance (MANOVA) using educational

FIGURE 1. Preferred moral judgment stage as assessed by the Defining Issues Test (J. R. Rest, 1979, 1986; J. R. Rest, D. Narvaez, M. Bebeau, & S. Thoma, 1999), by educational experience group (standardized scores).

experience and gender as independent variables and the three text recall variables as dependent variables (nonmoral, expository, narrative). The multivariate effect was significant for education, Wilks's $\Lambda = 0.66$, $F(6, 104) = 3.99$, $p < .001$, $\eta^2 = .19$, but not for gender, Wilks's $\Lambda = 1.00$, $F(3, 52) = .07$, $p > .95$, nor for the interaction, Wilks's $\Lambda = 0.96$, $F(6, 104) = .34$, $p > .90$. There was a significant univariate effect for expository text, $F(2, 60) = 10.69$, $p < .001$, $\eta^2 = .28$, but not for the nonmoral story ($p > .65$) or the narratives ($p > .08$).

TABLE 3. Mean Proportions and Standard Deviations of Events Recalled, by Text and by Group

| | Educational level | | | | | |
| | Undergraduate students[a] | | Seminarians[b] | | Graduate students[c] | |
Story type	*M*	*SD*	*M*	*SD*	*M*	*SD*
Nonmoral narrative						
Tom	.42	.14	.40	.16	.43	.14
Moral narrative texts overall	.34	.08	.33	.10	.39	.07
Penny	.34	.09	.32	.09	.38	.10
Sara	.34	.09	.37	.12	.40	.07
Moral expository texts overall	.23$^•$	.08	.29$^•$	.10	.39$^•$	.06
Sentence	.33$^+$	.10	.37	.12	.41$^+$	.08
Life	.16†	.08	.23†	.10	.32†	.07

Note. Entries that share superscripts differ at the *p* < .05 level.
[a]*n* = 37. [b]*n* = 34. [c]*n* = 16.

To test the influence of the three key variables (age, educational experience, and moral judgment score), we performed multivariate analyses of covariance (MANCOVAs) with two sets of dependent variables: (a) recall of each stage across texts (5 variables, one for each stage of Stages 1–5) and (b) recall of each stage by type of text (narrative or expository, across stories in the same category; 10 variables). For both analyses, the independent variable was educational experience, and the covariates were moral judgment score, recall of the nonmoral story (as a measure of production), and age. We did not find gender differences initially, so we left this variable out of the final analyses. See Table 4 for mean recall (*z* score averages) of moral judgment stages by group.

Recall of Moral Stages Across Texts

Effects of educational experience. First, when testing moral stage recall, we found a significant multivariate main effect for educational experience, Wilks's Λ= 0.63, $F(10, 75) = 3.90$, $p < .001$, $\eta^2 = .21$. Several univariate tests also were significant. Educational experience was significant for Stage 1, $F(1, 79) = 6.17$, $p < .003$, $\eta^2 = .13$; Stage 4, $F(1, 79) = 8.18$, $p < .001$, $\eta^2 = .17$; and Stage 5, $F(1, 79) = 5.36$, $p < .007$, $\eta^2 = .12$. The means for these scores show a clear novice-to-expert trend in that the lowest stage was best recalled by the low-educational-experience group, whereas the highest stage was best recalled by the high-educational-experience group. Educational experience mattered for all scores except Stages 2 and 3: Undergraduates recalled Stage 1 arguments better

TABLE 4. Recall of Moral Reasoning Stages (*z* Score Average), by Group

Moral reasoning stage	Undergraduate students		Seminarians		Graduate students	
	M	*SD*	*M*	*SD*	*M*	*SD*
1	0.43	1.02	–0.24	0.93	–0.49	0.74
2	–0.06	1.01	0.18	0.95	–0.01	1.00
3	–0.25	0.93	0.18	0.98	0.32	1.13
4	–0.41	0.75	0.28	1.06	0.45	1.10
5	–0.45	0.95	0.02	0.95	0.93	0.71

than did the other groups but were worse at recalling Stage 4 and 5 arguments. Follow-up paired contrasts revealed significant differences between undergraduate and graduate students for Stage 1 ($p = .002$), Stage 4 ($p = .006$), and Stage 5 ($p = .002$), and between seminarians and graduate students for Stage 5 ($p = .02$).

Effects of covariates. No multivariate or univariate effects emerged for age ($p > .11$). However, for P score, a multivariate main effect emerged, Wilks's $\Lambda = 0.76$, $F(5, 75) = 4.74$, $p < .001$, $\eta^2 = .24$, as did univariate effects for Stage 4 recall, $F(1, 79) = 3.99$, $p < .049$, $\eta^2 = .05$, and Stage 5 recall, $F(1, 79) = 21.21$, $p < .001$, $\eta^2 = .21$. Thus, participants with higher P scores did significantly better on Stage 4 and Stage 5 moral recall, showing the effect of moral reasoning expertise on recall of the highest stages.

A multivariate main effect also emerged for nonmoral story recall, Wilks's $\Lambda = 0.80$, $F(5, 75) = 3.54$, $p < .006$, $\eta^2 = .19$. There were also significant univariate effects for the nonmoral story recall for Stage 3, $F(1, 79) = 11.62$, $p < .001$, $\eta^2 = .13$, and Stage 4, $F(1, 79) = 12.67$, $p < .001$, $\eta^2 = .14$, but not for Stage 5 ($p > .10$), suggesting that the number of events recalled was related to recall of the middle stages but not to recall of the lowest or highest stages. This finding may be the result of there being more Stage 3 and Stage 4 arguments to recall.

Recall of Moral Stages by Text Type

Table 5 displays mean recall (*z* score averages) of moral argument stages by story type and group. Stage 1 was not included because too few participants recalled the Stage 1 expository argument.

Effects of educational experience. When testing moral stage recall within type of story, we found a significant multivariate effect for educational experience, Wilks's $\Lambda = 0.69$, $F(16, 146) = 1.82$, $p < .03$, $\eta^2 = .17$, and several univariate effects. Educational experience was related to Stage 4 expository recall, $F(1, 79) = 98.54$, $p < .001$, $\eta^2 = .18$, and Stage 5 narrative recall, $F(1, 79) = 4.02$, $p < .02$,

TABLE 5. Recall of Moral Reasoning, by Stage (z Score Average), Story Type, and Group

Moral reasoning stage	Expository texts						Narrative texts					
	Undergraduate students		Seminarians		Graduate students		Undergraduate students		Seminarians		Graduate students	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
2	−0.19	1.00	0.18	1.00	0.19	0.99	0.25	0.80	0.04	1.11	−0.40	0.88
3	−0.27	0.82	0.19	1.03	0.38	1.20	−0.07	1.10	0.06	0.90	0.02	1.07
4	0.46	0.72	0.39	0.96	0.37	1.23	−0.14	0.92	−0.04	1.10	0.40	0.96
5	−0.39	0.86	0.14	0.99	0.59	1.04	−0.35	0.79	−0.08	1.01	0.89	0.74

TABLE 6. Significant Group Differences in Moral Stage Recall, by Story Type and Factor

	Story type	
Factor	Expository	Narrative
Educational experience	Stage 4, 5	Stage 5
Moral judgment	Stage 2,[*] 4, 5	Stage 5
Nonmoral	Stage 3, 4	Stage 3, 4
Age		Stage 2,[*] 4

[*]Scores were highest for the lower education group.

$\eta^2 = .09$, and was only marginally significant for Stage 5 expository recall, $F(1, 79) = 2.92$, $p = .06$. In these cases, the graduate students performed significantly better than did the other groups.

Using an ANOVA, we found that several post hoc comparisons among groups (with Bonferroni correction) were significant. For Stage 4 arguments in expository texts, we found significant differences between undergraduates and seminarians ($p = .0001$) and between undergraduate and graduate students ($p = .015$). Graduate students were significantly better at recalling Stage 5 arguments in the narratives than were both undergraduates ($p = .0001$) and seminarians ($p = .001$) but were significantly better than only the undergraduates for recall of Stage 5 arguments in expository texts ($p = .002$).

Effects of covariates. A significant multivariate effect emerged for P score, Wilks's $\Lambda = 0.73$, $F (8, 72) = 3.34$, $p < .003$, $\eta^2 = .27$. P score univariate effects were also significant for Stage 5 recall, both expository, $F(1, 79) = 7.07$, $p < .009$, $\eta^2 = .08$, and narrative, $F(1, 79) = 15.75$, $p < .001$, $\eta^2 = .17$. Higher P scores were associated with better recall of Stage 5 moral arguments.

A significant multivariate effect also emerged for nonmoral story recall, Wilks's $\Lambda = 0.78$, $F(8, 72) = 2.68$, $p < .01$, $\eta^2 = .23$. In addition, nonmoral story recall was related to Stage 3 narrative recall, $F(1, 79) = 4.90$, $p < .03$, $\eta^2 = .06$, Stage 3 expository recall, $F(1, 79) = 8.91$, $p < .004$, $\eta^2 = .10$, Stage 4 narrative recall, $F(1, 79) = 7.79$, $p < .007$, $\eta^2 = .09$, and Stage 4 expository recall, $F(1, 79) = 8.70$, $p < .004$, $\eta^2 = .10$. Again, these results may be the result of the preponderance of Stage 3 and 4 arguments.

Last, when testing for moral recall by story type, a significant multivariate main effect emerged for age, Wilks's $\Lambda = 0.81$, $F(8, 72) = 2.13$, $p < .04$, $\eta^2 = .19$. Age mattered for Stage 4 narrative recall, $F(1, 79) = 7.73$, $p < .007$, $\eta^2 = .09$; older participants recalled Stage 4 arguments better than younger participants did. In addition, younger participants surpassed both of the other groups for Stage 2

narrative recall, $F(1, 79) = 4.32$, $p < .04$, $\eta^2 = .05$. Results for moral recall by story type are summarized in Table 6.

Discussion

We tested recall of moral narratives and moral expository texts. The texts presented issues from everyday life that average people encounter. This approach differs from studies in which experts and novices were compared on problems that were easy for the experts but new and difficult for the novices (Hmelo-Silver, Nagarajan, & Day, 2002). Much like political discourse on editorial pages and television news, these texts interwove moral reasons with everyday detail and were on topics with passing familiarity to typical adults.

We tested three groups who differed in amount of educational experience in the moral domain: undergraduates, Lutheran seminarians, and graduate students in philosophy and political science. We examined educational experience, moral judgment scores, and age for their relations to recall of moral arguments at different stages of reasoning, and all three variables had an influence. For example, age appeared to follow a developmental trend in the narrative stories; younger students recalled more Stage 2 and older students more Stage 4 moral arguments. Similarly, both P score and educational experience mattered for moral argument recall. Undergraduates recalled Stage 1 arguments better (no expository Stage 1 was recalled, so Stage 1 recall is of narratives only), whereas the more educated students recalled higher stage arguments better. The specialized experience of the seminarians was evident in their high recall of the Stage 4 and 5 arguments in the expository texts. Their performance on these texts at the higher stages was equivalent to that of the graduate students. However, that their experience with moral problem solving was not as sophisticated as that of the graduate students was demonstrated in their poorer performance in recalling Stage 5 moral arguments overall and in the narrative texts. In combination with the relation between P score and recall of Stage 5 (both overall and by text type), these findings suggest that moral reasoning experience added significantly, beyond educational experience, to recall of Stage 5. In other words, the moral reasoning expertise of advanced students contributed to virtually all higher stage recall, but moral judgment scores still contributed beyond educational experience to recall of the most sophisticated arguments.

Recall of Moral Texts as a Function of Education, Age, and Moral Judgment

Our first hypothesis was that participants with higher moral reasoning expertise, regardless of age and educational experience, would exhibit better recall for the highest stage moral arguments than would those with low moral reasoning expertise. In particular, we expected the advantage of moral judgment development to remain, even when controlling for educational experience and

age. Because both of these latter variables also had a relation to recall of moral texts, we consider them first, followed by a discussion of the value added by moral reasoning expertise.

Educational experience influenced moral recall. As in research of experience differences in other domains, the groups responded distinctively to text events (Johnson & Mervis, 1997; Myles-Worsley, Johnston, & Simons, 1988). The higher-educational-experience group did not have superior performance for the nonmoral text (see Table 3), suggesting that educational experience per se was not a factor for performance on the recall task generally. As shown in Table 4, compared with participants with less educational experience, participants with more education in moral reasoning demonstrated superior performance in response only to high-level moral arguments. Kohlberg's (1984) theory of sequential, hierarchical stages in moral judgment was thus supported. The higher-educational-experience group consistently recalled more complex (postconventional) arguments, demonstrating its distinctive competence with such reasoning, whereas the lower-educational-experience group performed better with moral argument recall at Stage 1. Since we used age as a covariate, these results appear to reflect moral reasoning experience and not only age-related developmental change.

Despite the familiarity of the types of problems in the narratives and equal overall recall levels among groups, the higher education group performed significantly better for high-stage recall compared with the lower education groups. Although the narratives we used contained fragments of higher-moral-stage concepts, the lower education groups were less likely to recall them, suggesting that they did not have the resources to apprehend the intentionally fragmented moral arguments. These results corroborate findings of expert–novice differences in other studies. For example, in research with baseball novices and experts, Voss, Fincher-Kiefer, Greene, and Post (1986) reported that novices seemed to be limited by what was explicitly mentioned about baseball in a text, whereas the experts drew on background knowledge when interpreting the text. Nevertheless, the youngest participants, who were largely in the lower-educational-experience group, did have background knowledge that they employed to identify the fragments of arguments from Stages 1 and 2. In particular, the young, lower-educational-experience participants successfully recalled Stage 1 arguments overall and Stage 2 arguments in the narratives better than their older, more educated counterparts did. Ostensibly, the familiar moral schemas associated with the lower-stage arguments were active, helping these participants recall them. Moravcsik and Kintsch (1993) found that correctness of recall depended on "the availability of appropriate domain knowledge" (p. 361). Clearly, the young, lower-educational-experience participants seemed to have some moral knowledge, but it may have been insufficient or inappropriate for recalling the higher stage arguments. As Neisser (1967) pointed out, "Information can be picked up only if there is a developmental format ready to accept it.

Information that does not fit such a format goes unused. Perception is inherently selection" (p. 55). The asymmetrical nature of the results provides support for the argument that moral texts activated background knowledge as readers processed them. The fact that younger, undergraduate readers were equally if not more adept than the other groups at generating low-level moral arguments but significantly less adept at identifying higher level arguments suggests that both development and experience with moral reasoning improve comprehension of moral texts that contain similar reasoning.

Although we did not test whether the effects we found were the result of encoding or retrieval differences, theoretically, schemas filtered and selected text events generally. Not only did higher-educational-experience participants recall the higher levels of moral arguments better than did the lower-educational-experience participants in the moral texts, but those with more education also remembered the general story events from the expository moral texts better than did those with less education (see Table 3). This finding suggests that those with more experience in moral reasoning may have a wider *causal field* of explanation (a wider set of accessible explanations to choose from) for the moral events in the story (Einhorn & Hogarth, 1986; Ericsson & Kintsch, 1995; Mackie, 1980) and consequently have more resources to infer relations between all events in the text. In these stories, moral arguments were intertwined with general content, so their cohesion may have been interdependent. As a result, the stories may have made more sense or activated a more stable mental representation of the story events in individuals with more formal educational experience.

Age influenced moral recall in a developmentally expected fashion. Age appeared to influence performance on the tasks separately from educational experience or moral reasoning expertise. Age was related to recall of Stage 2 for younger participants and Stage 4 for older participants in the narrative stories. Stage 2 arguments are more salient for emerging adults than for adults, presumably because those arguments still ring true. Moral judgment scores on the DIT (see Figure 1) also showed that for undergraduates, who were on average the youngest participants, Stage 2 was one of the most preferred stages. Moral development is thus reflected as a function of age in both recall of moral texts and moral judgment scores.

Moral judgment contributes to moral recall beyond educational experience and age. Despite the evidence that training or focused study played a role in what was recalled, moral judgment scores predicted recall of Stage 4 and 5 arguments beyond educational experience alone. The standard deviations for P scores were large within each education group, meaning that some individuals' P scores overlapped with those of individuals in the other education groups. Even after taking educational experience into account, individuals with high moral judgment scores recalled more high-stage arguments, on average, than did individuals with lower moral judgment scores. Higher moral judgment

scores contributed to Stage 4 and 5 general recall because those participants had the moral schemas to comprehend them, even though the education group to which they belonged may have had lower P-score averages. This result is not surprising. Higher scores in moral judgment are promoted by enriched social environments, not only by education (Rest, 1986). Participants with higher scores who did not fall into the graduate student group had developed above-average scores for their educational-experience group because of factors beyond their level of formal education.

Moral Recall and Text Type

Our second hypothesis was that recall of expository texts would be more difficult than recall of narrative texts in general but that recall of expository texts, more so than narratives, would be a function of moral reasoning expertise. We also expected educational experience to correlate with recall of expository texts because of the enhanced comprehension requirements of expository compared with narrative writing. These hypotheses were largely supported.

The higher difficulty in recall of the expository texts compared with the narrative texts appeared largely for the undergraduates and less for the other groups (see Table 3); however, we did not test this finding directly. Education beyond the undergraduate level appeared to aid the comprehension of expository material for seminarians and graduate students. In particular, graduate students outperformed the other groups on the editorial "Life and Death." This text contained a high proportion of Stage 4 and Stage 5 arguments (see Table 1) and may have been especially challenging for the other two groups. These group differences support the hypothesis that domain familiarity would facilitate recall of expository texts. Moreover, although educational experience predicted scores for Stage 5 arguments in narrative texts alone, moral judgment score predicted scores for Stage 5 arguments in both types of texts, suggesting that moral expertise made a particular difference in the expository texts.

Using Texts to Study Moral Development

Moral text processing holds promise as a means to test moral thinking. Using moral text recall as a means to test moral thinking overcomes the limitations of requiring participants to explain their understanding, as in an interview, and the limitations of using only recognition to measure understanding. Instead, text recall allows for a more veridical measure of what people do with the discourse they process day to day. This method is ripe for further exploration, particularly in an era when discourse is comprehended so differently according to ideology. For example, participants process the same information with a positive or negative bias depending on whether it refers to a political candidate they like or one they dislike (Westen, Blagov, Harenski, Kilts, & Hamann, 2006).

Future Research

Vitz (1990) proposed that there are differences in processing of narratives and expository texts based on Bruner's (1986) dual view of human thought. *Propositional thinking* is logicoscientific, paradigmatic, and formal interpretation, which possesses public procedures for verification and is context independent. This type of thought is well represented by typical measures of moral judgment such as the MJI and DIT. In contrast, *narrative thinking* is concrete, interpersonal, and subjective. It is characterized by description that aims at verisimilitude and requires understanding human intention. It is appreciative of particularities, such as time and place, character, and action, and its power derives from contextual sensitivity. This type of thinking has not been measured systematically in the moral domain despite the fact that one of the earliest human cognitive faculties to emerge is the ability to create and comprehend narratives (Fivush, 1997; Neisser, 1967).

In a sense, we explored the two types of thought in our comparisons of narrative and expository stimuli. Our data suggest that comprehension of moral expository texts is facilitated by moral judgment development and training. However, Bruner's (1986) narrative thinking may not have been presented sufficiently in our narratives. Our narratives were condensed episodes with a high amount of embedded moral reasoning. A more typical narrative that taps narrative thought would focus less intensely on reasoning and instead highlight other aspects of morality, such as moral goals, moral identity, moral emotion, moral sensitivity, and moral action, or their opposites.

AUTHOR NOTES

Darcia Narvaez is an associate professor of psychology at the University of Notre Dame. In her research, she tries to bridge the gaps in approaches to character education and moral psychology. Integrating cognitive science, expertise, and classical notions of virtue cultivation, she has developed the Integrative Ethical Education model. **Tracy Gleason** is an associate professor of psychology at Wellesley College. Most of her research concentrates on exploring and describing the relationships that some preschool-aged children have with imaginary companions. She is also interested in working on policy issues such as passage of the United Nations Convention on the Rights of the Child.

REFERENCES

Abernathy, C. M., & Hamm, R. M. (1995). *Surgical intuition*. Philadephia: Hanley & Belfus.

Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychologist, 27,* 33–51.

Alexander, P. A., Pate, P. E., & Kulikowich, J. M. (1989). Domain-specific and strategic knowledge: Effects of training on students of differing ages or competence levels. *Learning and Individual Differences, 1,* 283–325.

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

Arbuckle, T. Y., Vanderleck, V. F., Harsany, M., & Lapidus, S. (1990). Adult age dif-
ferences in memory in relation to availability and accessibility of knowledge-based
schemas. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,*
305–315.

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

Birkmire, D. P. (1985). Text processing: The influence of text structure, background
knowledge, and purpose. *Reading Research Quarterly, 20,* 314–326.

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

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University
Press.

Chase, W., & Simon, H. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Chi, M. T .H., & Koeske, R. (1983). Network representation of a child's dinosaur knowl-
edge. *Developmental Psychology, 19,* 29–39.

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., Speicher, B., Hewer, A., Candee, D., Gibbs, J. et al. (1987). *The
measurement of moral judgment* (Vol. 1 & 2). New York: 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.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin,
99,* 3–19.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Coté, N. C. (1990). Encoding and recall
of texts: The importance of material appropriate processing. *Journal of Memory and
Language, 29,* 566–581.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological
Review, 12,* 211–245.

Fincher-Kiefer, R., Post, T. A., Greene, T. R., & Voss, J. V. (1988). On the role of prior
knowledge and task demands in the processing of text. *Journal of Memory and Lan-
guage, 27,* 416–428.

Fivush, R. (1997). Event memory in early childhood. In N. Cowan (Ed.), *Development of
memory in childhood* (pp. 139–161). Hove, England: Psychology Press/Erlbaum.

Goldman, S., & Varma, S. (1995). CAPping the Construction-Integration Model of dis-
course. In C. A. Weaver III, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehen-
sion: Essays in honor of Walter Kintsch* (pp. 337–358). Hillsdale, NJ: Erlbaum.

Haas, C., & Flower, L. (1988). Rhetorical reading strategies and the construction of mean-
ing. *College Composition and Communication, 39,* 167–183.

Hayes, D. A., & Tierney, R. J. (1982). Developing readers' knowledge through analogy.
Reading Research Quarterly, 17, 256–280.

Hmelo-Silver, C. E., Nagarajan, A., & Day, R. S. (2002). "It's harder than we thought it
would be": A comparative case study of expert–novice experimentation strategies. *Sci-
ence Education, 86,* 219–243.

Hogarth, R. M. (2001). *Educating intuition.* Chicago: University of Chicago Press.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic
level of categorization. *Journal of Experimental Psychology: General, 126,* 248–277.

Johnson, P. E., Hassebrock, F., Duran, A. S., & Moller, J. H. (1982). Multimethod study of
clinical judgment. *Organizational Behavior and Human Performance, 30,* 201–230.

Keil, F. C., & Wilson, R. A. (1999). Explaining explanations. In F. C. Keil & R. A. Wilson
(Eds.), *Explanation and cognition* (pp. 1–18). Cambridge, MA: MIT Press.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco: Jossey-Bass.

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

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

Kulikowich, J. M., & Alexander, P. A. (1990). The effects of gender, ability, and grade on analogy performance. *Contemporary Educational Psychology, 15,* 364–377.

Mackie, J. L. (1980). *The cement of the universe*. Oxford, England: Clarendon.

McDaniel, M. A., Einstein, G. O., Dunay, P. K., & Cobb, R. E. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory & Language, 25,* 645–656.

Meutsch, D. (1989). How to do thoughts with words II: Degrees of explicitness in think-aloud during the comprehension of literary and expository texts with different types of readers. *Poetics, 18,* 45–71.

Moravcsik, J. E., & Kintsch, W. (1993). Writing quality, reading skills, and domain knowledge as factors in text comprehension. *Canadian Journal of Experimental Psychology, 47,* 360–374.

Myles-Worsley, M., Johnston, W., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14,* 553–557.

Narvaez, D. (1998). The effects of moral schemas on the reconstruction of moral narratives in 8th grade and college students. *Journal of Educational Psychology, 90,* 13–24.

Narvaez, D. (1999). Using discourse processing methods to study moral thinking. *Educational Psychology Review, 11,* 377–394.

Narvaez, D. (2002). Individual differences that influence reading comprehension. In M. Pressley & C. C. Block (Eds.), *Reading comprehension instruction* (pp. 158–175). New York: Guilford.

Narvaez, D., & Bock, T. (2002). Moral schemas and tacit judgement or how the Defining Issues Test is supported by cognitive science. *Journal of Moral Education, 31,* 297–314.

Narvaez, D., Endicott, L., & Thoma, S. J. (2001, August). *Expertise and the speed of moral information processing.* Paper presented at the 109th annual meeting of the American Psychological Association, San Francisco.

Narvaez, D., Getz, I., Rest, J. R., & Thoma, S. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology, 35,* 478–488.

Narvaez, D., Lapsley, D., Hagele, S., & Lasky, B. (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality, 40,* 966–985.

Narvaez, D., van den Broek, P., & Ruiz, A. (1999). Reading purpose, type of text and their influence on think-alouds and comprehension measures. *Journal of Educational Psychology, 91,* 488–496.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlhausen, M. M., & Roller, C. M. (1988). The operation of text structure and content schemata in isolation and interaction. *Reading Research Quarterly, 23,* 70–85.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Rest, J. R. (1979). *Developing 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. (1994). *Moral development in the professions: Psychology and applied ethics*. Hillsdale, NJ: Erlbaum.

Rest, J. R., Narvaez, D., Bebeau, M., & Thoma, S. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach.* Mahwah, NJ: Erlbaum.

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

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

Roller, C. M. (1990). The interaction between knowledge and structure variables in the processing of expository prose. *Reading Research Quarterly, 25,* 79–87.

Schneider, W., Körkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology, 81,* 306–312.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61,* 394–403.

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

Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing interences in expository text comprehension. *Discourse Processes, 24,* 199–228.

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.

Sternberg, R. (1998). Abilities are forms of developing expertise, *Educational Researcher, 3,* 22–35.

Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology, 24,* 359–375.

Tappan, M. (1997). Analyzing stories of moral experience: Narrative, voice, and the dialogical self. *Journal of Narrative and Life History, 7,* 379–386.

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

Thoma, S. J. (1994). Moral judgments and moral action. In J. R. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 199–211). Hillsdale, NJ: Erlbaum.

Toulmin, S. (1981). The tyranny of principles. *Hastings Center Report, 11,* 31–39.

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* (pp. 83–111). Hillsdale, NJ: Erlbaum.

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., Rohleder, L., & Narvaez, D. (1996). Causal inferences in the comprehension of literary texts. In R. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to discourse: Perspectives from the third IGEL conference* (pp. 179–200). Norwood, NJ: Ablex.

Vitz, P. C. (1990). The use of stories in moral development: New psychological reasons for an old education method. *American Psychologist, 45,* 709–720.

Voss, J. F., Fincher-Kiefer, R. H., Greene, T. R., & Post, T. A. (1986). Individual differences in performance: The contrastive approach to knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 297–334). Hillsdale, NJ: Erlbaum.

Walker, C. H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction, 4,* 25–42.

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,* (Vol. 2, pp. 23–52). Hillsdale, NJ: Erlbaum.

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 U.S. presidential election. *Journal of Cognitive Neuroscience, 18,* 1947–1958.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 920–933.

APPENDIX A
Sara's Evening at Home

Sara drove into her garage stall and sighed. Thank goodness it was Friday. It had been a heck of a day at work—constant interruptions and impossible deadlines. But she wouldn't have to think about it for two days.

She started the bathtub water. A good, hot bubble bath could do wonders. As she began to realize her good fortune of having an unplanned evening, her eyes sparkled with possibilities. She could walk down to the lake and watch the stars later. She could pull out that new romance novel she'd been dying to read. She could finish knitting the baby booties for her expectant sister-in-law, or she could write to her friends who lived out of town.

Or I could call some friends to come over. Jed? No, he was out skiing this weekend . . . Cindy? She was probably busy with her new friends . . . Nona was at a family party, Ted was with his mother in the hospital, Heather was playing at the symphony . . . Everybody was busy. It's okay. I like to be alone sometimes. After a light supper and reading, Sara ended up in front of the television with her knitting. As she was beginning her second bowl of popcorn, the doorbell rang. It was Cindy.

"Hey, buddy! How's it going?"

"Come on in. Have some popcorn. Let me tell you about work today!"

"Actually I can't. I'm working—you know—my volunteering. We are going to be sponsoring a demonstration tomorrow at the weapons plant. I came to remind you and to invite you along. It starts at 8 and goes all day, or as long as we can last before being arrested for trespassing. I've got some chains you can use to lock yourself to the fence." At Sara's silence, she paused. "After our months of debate, aren't you convinced yet that this is important to do?" She sat down and dug into the popcorn.

Cindy was Sara's best friend. They had been friends since she could remember. Cindy had joined the peace movement (actively!) about a year ago and had been pestering Sara to join in a protest ever since. Sara was still hesitant. Cindy was getting impatient with her and had been spending more time with her fellow protesters than with Sara. Sara was worried about losing their friendship if she didn't go along with Cindy's request. (STAGE 3)

Cindy assumed that every reasonable person would share her view. "That company makes napalm. They design that stuff to stick to human flesh while it's burning! You can't get it off while its burning! I've seen pictures of children who are horribly maimed and disfigured by it! It's grotesque! (STAGE 3) We can't allow anyone to use napalm for any reason!"

Sara still wasn't sure that Cindy's protests were right. "But Cindy, it's against the law to trespass. What if everyone took the law into their own hands when they didn't like something? Like, what if I don't like it that McDonalds uses lard in their french fries and go out and chain myself to the golden arches? Think of the chaos there could be . . . 20 people over here chained to an arch, 31 people over there chained to Big Boy's leg . . . 49 people over there marching on Dairy Queen . . . <u>what a mess if everyone started breaking the law</u>!" (STAGE 4)

Cindy was ready for that response. "Listen! This munitions company is making millions of dollars from OUR tax money for the purpose of burning and maiming innocent little children. The military bureaucrats are making decisions to use MY tax dollars to do things that I absolutely abhor! The decision makers can't be trusted, and we can't stand by and let them act for us! <u>They are supposed to represent us. They get their power from us! They have power only as long as we give them the power</u>. (STAGE 5) I can't just stand by and let them make immoral decisions! And I am willing to go to jail to stop them!"

Sara still wavered. "I agree that each of us has to decide on issues of fairness. I agree that it is right to break the law sometimes, when doing so calls attention to some moral outrage. But I can't believe that our government is out to napalm innocent children, although it may have happened in some cases. We are a planet of wars. That means that defense is necessary, military might is necessary, munitions plants are necessary. I can't see that we have to take drastic measures yet in this case."

Cindy spoke quietly now. "Sara, it IS intentional. Napalm is designed for skin. It is a particularly inhuman form of combat. It should not exist. I will chain myself to the fence until my government stops using MY money to have it produced. I won't give up! . . . Won't you come to support me? I want you to be there. You are like family to me."

As Cindy slipped out the door, Sara stood in silence. She thought for a very long time.

APPENDIX B
Life and Death

<u>When patients suffer, doctors try to soothe</u> (STAGE 3: Doctors should try to soothe patients). <u>But sometimes no amount of medicine can ease a patient's agony, and the cry for relief becomes a plea for death</u> (STAGE 3: The right to die may be used because of untreated treatable suffering). What should doctors do then? <u>According to the Ninth U.S. Circuit Court of Appeals, they should feel free to help a terminally ill patient die</u> (STAGE 4: Those who interpret the law sanction euthanasia). <u>Delivered in San Francisco last week, that declaration is sure to stoke the debate over whether life's liberty stops at death's door</u> (STAGE 5: The principle of liberty extends to the right to die).

<u>It's no mystery why some bedridden and pain-wracked victims of terminal illness yearn for an early death</u> (STAGE 3: Patients should not have to suffer). <u>Last week's court decision, which struck down Washington State's ban on doctor-assisted suicide</u> (STAGE 1: Suicide is punishable), is the first appellate answer to their call. Sure to be reviewed by the Supreme Court, the 8–3 ruling was rooted not so much in conscience as in the Constitution. <u>That document, wrote Judge Stephen Reinhardt, guarantees the right to control "the time and manner of one's death"</u> (STAGE 5: Individuals have a right to dignity [humans have special privilege]; STAGE 2: People should have freedom or a right to do whatever they want).

<u>This is bound to be startling news to people who believe that human beings have no business summoning a dallying Angel of Death</u> (STAGE 4: The laws of nature should determine death, not humans). <u>It's sure to dismay those convinced that life is a gift that must be endured even when it can no longer be enjoyed</u> (STAGE 4: Sanctity of life: Life is to be endured even when it is difficult). <u>But as the opinion points out, courts are obliged</u>

to steer clear of such deep moral questions: "Those who believe strongly that death must come without physician assistance," wrote Reinhardt, "are free to follow that creed, be they doctors or patients (STAGE 3: Suicide is wrong). They are not free, however, to force their views, their religious convictions, or their philosophies on all the other members of a democratic society, and to compel those whose values differ with theirs to die painful, protracted, and agonizing deaths" (STAGE 5: We should uphold the principle of tolerance for multiple perspectives; STAGE 3: People have a right to dignity).

The court also questioned the wisdom of characterizing medically assisted death as "suicide." That term is best applied to the tragic, premature deaths caused primarily by untreated depression and other mental illnesses—needless deaths that society must strive to prevent. The appellate court sees the choices of the incurably ill in a different light: "A competent, terminally ill adult, having lived nearly the full measure of his life, has a strong liberty interest in choosing a dignified and humane death rather than being reduced at the end of his existence to a childlike state of helplessness, diapered, sedated, incompetent" (STAGE 5: Individuals have a right to dignity—humans have special privilege).

This decision definitely has a dark side well worth considering. Critics worry that the ruling foreshadows a time when indigent, suffering patients will enjoy no right to medical care but an absolute entitlement to a medically assisted death (STAGE 5: There may be an imbalance of rights with a supportive law—no right to health care but a right to die). That scenario prompts fears that some ailing patients may opt for death merely because Demerol is beyond reach (STAGE 2: Sets a precedent for death as a easy way out when things get hard). And some onlookers speculate that this right to die might evolve into a duty to die (STAGE 4: Slippery slope: May move us towards institutionalized forced death). Dwindling patients could feel pressured to hurry things up just to relieve loved ones of financial and emotional burdens (STAGE 4: Not having a law against it would result in an abuse of the right to die by families).

Of course, any right Americans enjoy can be abused (STAGE 4: Laws against euthanasia protect order). But the possibility of abuse—largely unrealized in the Netherlands, which permits doctor-assisted death—is no reason to foreclose thoughtful talk about the rights of the sick. This court ruling could help Americans seize the last freedom of life: choosing how to die.