MORAL THEME COMPREHENSION IN THIRD
GRADERS, FIFTH GRADERS, AND
COLLEGE STUDENTS

DARCIA NARVAEZ, JENNIFER BENTLEY,
TRACY GLEASON, AND JAY SAMUELS
University of Minnesota

Abstract

Proponents of character education (e.g., Bennett, author of The Book of Virtues) argue that children should hear moral stories in order to develop moral character. But do children necessarily understand the moral message of a narrative? Third and fifth grade students and adults were tested for their ability to apprehend the moral messages in three stories selected from The Book of Virtues and other sources. After listening and reading along with three moral stories, participants were asked to generate the message, select the message from a list of possible messages and select a vignette with the same message from three possible vignettes. Subjects also heard and read a non-moral story and answered comprehension questions about it. Results indicate a developmental difference in moral theme understanding unexplained by reading comprehension scores alone.
Ten years after the first edition of The Golden Notebook, Doris Lessing (1974) wrote about reader reaction to it:

I can get, in one week, three letters about it...One letter is about the sex war, about man's inhumanity to woman, and woman's inhumanity to man...The second is about politics...and here she writes many pages about politics and never mentions any other theme. The third letter...is written by a man or a woman who can see nothing in it but the theme of mental illness. But it is the same book! And naturally these incidents bring up again questions of what people see when they read a book, and why one person sees one pattern and nothing else at all of another pattern, and how odd it is to have, as author, such a clear picture of a book, that is seen so differently by its readers." (p. xi)

The research in this article addresses issues raised by Lessing: what do children perceive to be the theme when they read or listen to stories that have a moral message?

The use of narratives and stories for moral education has been advocated by societies around the world (e.g., Ching, 1989), including the U.S.A., as evident since the popularity of the McGuffey reader in the U.S.A. over 100 years ago. More recently it is supported by theorists (e.g., Tappan & Brown, 1989; Vitz, 1990) and more popularly by character education proponents like Bennett. In fact, Bennett's Book of Virtues (1993) was a runaway bestseller for more than a year only to be replaced on the bestseller list by its sequels. The popularity of these books is due in part to the claims of Bennett and others (Kilpatrick, 1992; Lickona, 1991; Wynne & Ryan, 1993) who argue that, in order to develop moral values and good character,
children should hear and read moral stories. Despite a lack of research supporting recommendations to read stories in order to build values (Leming, 1993), these advocates appear to assume that children comprehend moral stories in the way that adults do and in the way that authors intend. In fact, narrative comprehension research demonstrates that children do not understand narratives in the same way as adults do (e.g., Collins, 1983; Perfetti, 1985; van den Broek, Lorch & Thurlow, 1997). There is evidence that adults themselves may not understand moral narratives in a uniform fashion since there are recall differences for moral reasoning in narrative texts due to differences in moral judgment development (Narvaez, in press).

Research on theme comprehension

Theme comprehension or extraction has been approached in various ways. Lukens (1982) defined a theme as "the idea that holds the story together, such as a comment about society, human nature, or the human condition" (p. 3). Lukens suggested that a theme answers the question: "What does it all mean?" Using a similar definition, several researchers have looked at children's ability to extract a theme. Taylor (1986) studied summary writing in 4th and 5th grade students including summarizing the point of a narrative. Although Taylor found that summarizing the narrative was easy for the children, summarizing its point was not; it was as difficult as summarizing an expository text. In a study about the development and processes of summarization, Johnson (1984) found that summarization was more difficult than recall for elementary school children. Brown, Campione, & Day (1981) tested rules for summarizing texts and found that the most difficult task for novices was adding information, as is
required in constructing a theme, rather than just repeating or modifying existing sentences. Goldman, Reyes & Varnhagen (1984) studied K-6 children's understanding of fables, including their ability to extract a lesson from each fable. Children were generally unable to extract a lesson until fourth grade. Younger children tended to give more concrete, story-specific statements as lessons. Lehr (1988) gave kindergartners, 2nd and 4th graders three stories and asked them to do several things including matching two by theme, summarizing and extracting the lesson. Some stories were realistic and some were folktales (and some were familiar like *The three little pigs*). Previous high exposure to literature improved scores on theme identification but there was a developmental trend for theme identification among low-exposure students. She found that it was easier to generate and identify themes from realistic fiction than from folktales. In sum, theme extraction is a difficult task for children, especially for expository texts and unrealistic fiction; children are generally unsuccessful until the fourth grade.

Regardless of age, themes are not typically included in story recall nor in responses to questions regarding the topic of the story (Brown, Campione & Day, 1981; Goldman, Reyes & Varnhagen, 1984; Mandler & Johnson, 1977; Stein & Glenn, 1979; Taylor, 1986). Main points are not automatically extracted from texts (Afflerbach, 1990; Reder & Anderson, 1980; Williams, 1993) but require generalizing from the literal level of the text (van Dijk & Kintsch, 1978).

The relationship between theme comprehension and general reading comprehension has been studied in regard to its influence on
text comprehension. Dooling & Mullet (1973) varied the point at which a thematic title was presented (before, after, not at all) for prose passages. Those who saw the theme before reading the text performed significantly better on comprehension and recall. Those with more background knowledge about the topic also performed better. Afflerbach (1990) found that experts automatically construct the main idea of texts that are about familiar topics significantly more than for texts about unfamiliar topics. Afflerbach suggests that main idea construction is not automatic or fundamental generally unless the topic is familiar.

Moral story comprehension

The research reported here focuses on the comprehension of moral themes in children's stories. Although there have been a number of studies about moral or theme comprehension, often the authors interpret "moral" to mean the point of the story, whether or not it has to do with morality. Two studies, however, have included moral messages in their stimulus materials. Johnson & Goldman (1987) studied children's ability to recognize rules of moral conduct in stories and their ability to group stories according to the rules of conduct. They used Bible stories that illustrated one of three "rules of conduct": helping, obeying or not being afraid. Young children tended to group stories according to actions and concrete items rather than rule similarity. The other study that has focused on moral story comprehension is Stein & Trabasso (1982). The authors manipulated aspects of intention by varying information about motives and their goal-relatedness; making a story variant of Kohlberg's "Heinz and the drug," they manipulated the goals and severity of consequences. They
tested comprehension and evaluation of outcomes in response to manipulations.

Unlike these studies, the research reported in this article used non-religious, literary stories that concerned getting along with others in particular ways. The present study investigated theme comprehension using several tasks that probed the subject's representation of a moral story. Our definition of "moral" is a philosophical one involving cooperating or getting along with others (Piaget, 1932; Rawls, 1971; Rest, Narvaez, Bebeau & Thoma, 1997). For us, a "moral story" has a theme about a specific aspect of getting along with others. In the selection of texts for this study, we did not choose texts such as Aesop's fables since they focus mostly on types of prudence (e.g., don't be vain, plan wisely, don't be fooled, etc.). In this study we used well-constructed stories (i.e., each story had a beginning, middle, and end) from children's literature which had moral themes about a specific way to cooperate (tolerance, helpfulness, and self-sacrifice).

We expected to find that moral theme extraction would be more difficult for younger readers because theme extraction is generally more difficult for them. We set up three tasks which we expected to be of increasing difficulty: selecting the moral theme when it is found in a multiple-choice context, identifying the story-with-the-same-message when the message is in a new context, and moral message construction (a summarization-type task).

Hypotheses

The main hypothesis was that there would be a developmental difference in the comprehension of the moral message in moral
stories. A secondary set of hypotheses reflected the assumption that there would be a differential level of difficulty in the tasks regarding moral theme identification. The moral generation task should be more difficult than the multiple-choice tasks. It was expected that the story selection task would be more difficult than the theme selection task because it involves extracting the theme of each vignette and comparing it to the original story's theme.

Method

Subjects

The sample consisted of 127 subjects, 61 third grade students (mean age=9) and 38 fifth grade students (mean age=11) from suburban schools who volunteered to participate, and 28 college-level students from an public university who received course credit for participation. All groups had approximately equal numbers of males and females.

Materials

Stories from The Book of Virtues and similar texts, as well as many picture books, were examined in an effort to find stories that (a) had a clear moral (and not prudential) message; (b) had a beginning, middle, and end; (c) had topics to which children could relate; (d) were not familiar stories; (e) had characters in conflict; (f) were coherent, and (g) were short enough for a research study. Three moral stories were selected for this research. "Mother Snowbed" (or "Mother Holly"), a Grimm tale (Bennett, 1993), 825 words, is about two daughters of very different character who, when faced with the chance to help others, react differently. Each daughter is rewarded appropriately. "Damon and Pythias" (DeSpain, 1993) is an ancient
Greek tale (649 words). A man is condemned to death but asks to settle his affairs and say goodbye to his parents before he is executed. The king allows him to do so only because a friend agrees to take his place in prison and be executed if the man does not return within a given time. "Natural Habits" (ibid.) is an African tale about a monkey and a rabbit who have difficulty tolerating each other (473 words). They challenge each other to modify each one’s habits but are unsuccessful. A fourth story, an intercultural picture book story, *How my parents learned to eat* (Friedman, 1984), was used to measure the children's reading comprehension (605 words). A set of 20 true-false statements tested comprehension of this story. All texts approximated a Flesch-Kincaid readability level of grade three.

**Tasks**

There were three tasks to test moral message understanding: 1) The constructed response for each moral story, in which students wrote what they thought was the author's message. It was scored in two ways, as a gist "hit" or a "miss" and also as high or low on cooperation. 2) Readers selected the best moral theme from a selection of six possible themes (the distractors were generated from story material which were superficial points, truisms or minor subthemes). See Table 1 for the distractor and target themes. 3) Readers identified which of three vignettes had the same moral theme as the original story (one story type used the same actions--i.e., the same plot characteristics-- but had different actors and a different message; a second story type used the same actors but involved different actions and messages; the third story type had different actions and actors but the same message). From these three tasks we
<table>
<thead>
<tr>
<th>Story</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Damon and Pythias</td>
<td></td>
</tr>
<tr>
<td>1. It's not over till it's over</td>
<td></td>
</tr>
<tr>
<td>2. Be careful who your friends are</td>
<td></td>
</tr>
<tr>
<td>3. It's important to settle your business and say goodbye to parents before you die</td>
<td></td>
</tr>
<tr>
<td>4. Good luck may save you or your friend's life</td>
<td></td>
</tr>
<tr>
<td>*5. A friend is someone who helps you even when it might hurt</td>
<td></td>
</tr>
<tr>
<td>6. Don't mess around with kings</td>
<td></td>
</tr>
<tr>
<td>Natural Habits</td>
<td></td>
</tr>
<tr>
<td>1. Monkeys and rabbits make rude friends</td>
<td></td>
</tr>
<tr>
<td>2. Be alert, you may get tricked</td>
<td></td>
</tr>
<tr>
<td>3. Never trust a monkey or a rabbit</td>
<td></td>
</tr>
<tr>
<td>4. It's not polite to turn your head or scratch when you are talking</td>
<td></td>
</tr>
<tr>
<td>5. Don't make promises you can't keep</td>
<td></td>
</tr>
<tr>
<td>*6. Accept others as they are and don't try to change them</td>
<td></td>
</tr>
<tr>
<td>Mother Holly</td>
<td></td>
</tr>
<tr>
<td>1. Be good and you won't have to worry</td>
<td></td>
</tr>
<tr>
<td>2. Don't try to be anyone else, be your own person</td>
<td></td>
</tr>
<tr>
<td>*3. Willingness to help others will be rewarded</td>
<td></td>
</tr>
<tr>
<td>4. If someone asks you to do something, do it</td>
<td></td>
</tr>
<tr>
<td>5. Ashes are no comparison to gold</td>
<td></td>
</tr>
<tr>
<td>6. Look before you leap</td>
<td></td>
</tr>
</tbody>
</table>
developed an aggregate score that combined the responses from the three tasks for each of the three moral stories.

The independent variables were age/grade level and, for the children, score on the 20-item true-false comprehension test of the fourth story.

Procedure

The children were tested in groups of 25 to 55 in two 45-minute sessions one week apart at the end of the school year. Each child was given printed booklets containing (a) a practice story ("The boy who cried wolf") in which they were shown how to do all three tasks, (b) three stories which were used to test the student's ability to identify the moral theme, and (c) a fourth story used to assess reading comprehension independent of theme. The stories and test items were presented both on paper and on audiotape so that reading skill differences would be less likely to interfere with story comprehension because oral language skills are sometimes more advanced than written language skills in school-aged children (see McCounaughy, 1985). The tape recordings were made by a former professional voice actress. Children responded to the questions by marking the booklets.

The questions asked after each of the moral stories were the following:

1) What do you think the message or theme of this story is? In other words, what do you think the author would like you to learn from reading this story? You may quickly look back at the story but we will only give you one minute to answer.

2) Please read the following three stories and then decide whether Story A, Story B, or Story C has the same message as
[the story]. Circle the name of the story that best matches the message at the bottom of the page.

3) Some people have suggested one of the following to be the message of [the story]. Please read the following list of possible messages and circle the one you think best fits [the story].

The adults read the material and answered the questions on their own. Adults did not read the comprehension story or complete the associated true-false test.

**Scoring**

The stories and tasks were tested with individual adults and children as the materials were created and revised. With the final protocol, a control group of twenty adults completed the tasks to make sure that the researchers' understanding of the themes was in line with those of adults. The adults read the stories and then performed the two multiple-choice tasks. 1) they selected the story, out of three, that had the same moral message; 2) they selected the best message out of a list of six messages. The adults agreed 100% with the researchers regarding which message and story best reflected the original story.

For the primary analyses, scores for each story task were computed separately and then aggregated across tasks and stories for a moral comprehension score consisting of twelve items. For the moral generation task, answers were scored in two ways for each of the three stories: for getting the gist of the message (hit=1 or miss=0), and for high or low cooperation with others in the answer. Interjudge reliability was calculated at .91 using Cohen's kappa. For the multiple-choice message selection and story selection tasks, participants' choices were coded as hit/miss. A Cronbach alpha was
computed for this 12-item index of moral comprehension, alpha was .65, within the acceptable bounds for group analysis.

Results

Two separate sets of analyses were conducted on the data. In the first analysis, an aggregate score was computed across tasks and stories. A one-way ANOVA was conducted for age level. In the second, Chi-square analyses were conducted across age groups for each of the three tasks for each story.

Moral comprehension

The main hypothesis, that there is a developmental difference in the identification of the moral message in moral stories, was supported. A one-way ANOVA was significant for the aggregate moral comprehension score, $F_{(2,122)}=24.79$ (p<.0001) indicating a developmental difference in moral message comprehension. (See Figure 1.) Concern that this might only be due to reading comprehension prompted a separate ANOVA for children's scores using the 20-item reading comprehension score as a covariate (3rd grade M=17.16, S.D.=1.98; 5th grade M=18.50, S.D.=1.43). This, too was significant, $F_{(1.94)}=11.18$ (p<.001), suggesting that moral comprehension involves more than general reading comprehension skills.

Story and task analyses

A secondary set of hypotheses compared tasks within story across age (see Figure 2). These analyses were based on the assumption that the tasks reflect a differential level of difficulty in moral theme identification. The moral generation task was expected to be the most difficult, followed by the story selection task. The
Figure 1.
Mean Average for Aggregate Moral Comprehension Score by Grade Level (Maximum is 15)*.
Figure 2.

Mean Average for Theme Selection, Story Selection, and Moral Generation Across Stories by Grade.
Table 2
Proportions of Correct Answers by Grade for Each Task by Story and Across Stories

<table>
<thead>
<tr>
<th></th>
<th>Grade 3 (n=61)</th>
<th>Grade 5 (n=38)</th>
<th>Adults (n=28)</th>
<th>Chi-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral generation^</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damon</td>
<td>67</td>
<td>87</td>
<td>96</td>
<td>10.69**</td>
</tr>
<tr>
<td>Natural habits</td>
<td>64</td>
<td>87</td>
<td>100</td>
<td>16.82**</td>
</tr>
<tr>
<td>Mother Holly</td>
<td>87</td>
<td>87</td>
<td>100</td>
<td>4.11</td>
</tr>
<tr>
<td>Average across stories</td>
<td>73</td>
<td>87</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>Theme selection^</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damon</td>
<td>79</td>
<td>95</td>
<td>96</td>
<td>3.04*</td>
</tr>
<tr>
<td>Natural habits</td>
<td>59</td>
<td>97</td>
<td>93</td>
<td>24.66**</td>
</tr>
<tr>
<td>Mother Holly</td>
<td>61</td>
<td>84</td>
<td>96</td>
<td>16.78**</td>
</tr>
<tr>
<td>Average across stories</td>
<td>66</td>
<td>92</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Story selection^</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damon</td>
<td>75</td>
<td>97</td>
<td>100</td>
<td>15.44**</td>
</tr>
<tr>
<td>Natural habits</td>
<td>59</td>
<td>89</td>
<td>96</td>
<td>20.05**</td>
</tr>
<tr>
<td>Mother Holly</td>
<td>85</td>
<td>89</td>
<td>96</td>
<td>2.46</td>
</tr>
<tr>
<td>Average across stories</td>
<td>73</td>
<td>92</td>
<td>97</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-square significant at p < .01

**Chi-square significant at p < .001.

^The subject wrote what he or she thought was the theme of the story.

"The subject selected the theme from a list of possible themes.

'The subject selected the story with the same theme from a choice of three stories.
multiple-choice theme selection task was expected to be easiest. The proportions of correct responses by grade for these tasks are listed in Table 2. All Chi-square comparisons were significant except for two. The proportions indicate a developmental effect for all tasks.

**Distractors**

The generated messages that were misses were categorized into 9 categories. No statistical analyses were conducted. The categories were: prudence concerns (e.g., Damon: "That you should not try to get rid of somebody"); Natural Habits: "To not be alone in the jungle"), equity concerns (e.g., Mother Holly: "You must deserve your reward," "That you should not only love one of your brothers or sisters, you should love them both"), summary statements (e.g., Damon: "It tells about friendship", Natural: "That they were both trying to trick each other"), a focus on results (e.g., Mother: "Don't think that you'll get gold if somebody else does"), a focus on immutable traits (e.g., Natural: "It's natural that if you did something since you were a baby you still do it"), a focus on mutable traits (e.g., Damon: "Always be nice and don't be mean"), truisms (e.g., Damon: "People believe in friendship"), focus on a subtheme (e.g., Damon: "To help your parents you must go sell your house and get the money"), conventionality (e.g., Natural: Don't scratch or look around when eating"). See Table 3 for a summary.

**Discussion**

This study investigated children's ability to identify a moral theme. Advocates of character education suggest that reading stories with moral messages to children is a way to instill good values, but is this advice well-founded? Both types of data analyses used in this
Table 3

Number of Misses by Category for Moral Generation

<table>
<thead>
<tr>
<th></th>
<th>Damon</th>
<th>M. Holly</th>
<th>Natural Habits</th>
<th>Across stories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gr 3</td>
<td>Gr 5</td>
<td>Gr 3</td>
<td>Gr 5</td>
</tr>
<tr>
<td>Prudence</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Summary</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Equity</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Results focus</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Focus on</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>immutability</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Focus on</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mutable traits</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Truism</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Focus on subtheme</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Conventionality</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

study, the aggregate and the story-specific, indicate that the younger children had a more difficult time with all the moral theme identification tasks than did the older children. They made more errors in selecting the theme from a list, in selecting a story with the same theme and in constructing a theme for the story. They also had a more egocentric perspective on the themes that they generated, constructing cooperative themes less often than fifth grade students or adults. A developmental trend emerged on all tasks with adults performing at ceiling levels on the tasks. Contrary to expectation, the story selection task was not more difficult than the theme selection
task. Even after reading comprehension was taken into account, the fifth graders performed significantly better on the composite moral comprehension score. The third grade students performed best on one story, Mother Holly. This story is particularly graphic about the rewards/benefits of good/bad behavior.

The recommendations by Bennett and other proponents of character education who argue that children can learn values by hearing moral stories is not supported by our findings. These research results suggest that children do not necessarily pick up the message intended by the author in a narrative. Reading comprehension skills help but are not sufficient for picking up a moral message. Contrary to the claim of character education proponents, it takes more than reading or hearing a moral story to develop "moral literacy".

Implications

Although this is a preliminary study, we can suggest directions for further research. 1) Selecting literature for moral education: The findings here suggest that materials for moral education need to be selected carefully. More specific recommendations must wait for further research delineating the types and sources of comprehension differences among children. What kinds of themes attract younger children? What are the developmental landmarks for comprehension of certain kinds of themes? 2) What kinds of moral story characteristics facilitate moral theme comprehension? Jetton (1994) found that elementary school children are strongly influenced by narrative story elements, and have great difficulty extracting embedded text elements such as expository information which is not part of the story grammar. What kinds of story structure and affective
focus (Brewer & Lichtenstein, 1982) facilitate moral theme comprehension? How are protagonist goals and their completion (or not) related to moral theme recall (Lutz & Radvansky, 1997). How does causal connection strength of a moral theme relate to its comprehension (Trabasso & van den Brock, 1985)? 3) Instruction on theme comprehension: As theme extraction is difficult for children, some researchers have examined strategies to help them. For example, Williams, Brown, Silverstein, & de Cani (1994) demonstrated that fifth and sixth grade students can understand the theme of a narrative, but only with extended and skillful instruction. Perhaps structured questions before, during and after reading would facilitate comprehension of moral themes as found in other research (Carnine, Stevens, Clements, & Kameenui, 1982). 4) Is moral theme comprehension different than non-moral theme comprehension? The relationship between general theme comprehension itself and reading comprehension skills is not clearly delineated, although there is some evidence that the ability to generalize a theme (that is, to move beyond concrete elements) develops by fourth grade (Goldman et al., 1984). Our data suggest that although most children are able to extract a theme by the end of third grade, a large proportion of students cannot. Rosenblatt (1991) and Britton (1984) state that extracting embedded information is difficult and depends on factors such as point of view and how the reader focuses attention. Extracting the message an author intends may be as difficult as extracting character motives (Shannon, Kameenui & Baumann, 1988).
References


Author note

Darcia Narvaez, Department of Educational Psychology; Jennifer Bentley, Department of Educational Psychology; Tracy Gleason, Institute of Child Development; Jay Samuels, Department of Educational Psychology.

This study was supported by a grant from the Center for Applied Research and Educational Improvement, College of Education and Human Development, University of Minnesota. The authors would like to thank the Stillwater and Bloomington school districts, the third grade teachers of Hubert Olson Elementary School—especially Judy Gentz, and the fifth grade teachers of Forest Hills Elementary School—especially Nancy Hof. Also helpful in collecting data were Gretchen Newman and Irene Stone.

This study was presented at the annual meetings of the American Educational Research Association and the Association for Moral Education.

Correspondence concerning this article should be addressed to the first author at 206 Burton Hall, 178 Pillsbury Drive, Minneapolis MN 55455 or narvaez@turtle.psych.umn.edu.