

Gender Differences in Moral Development

Geri R. Donenberg and Lois W. Hoffman

University of Michigan

Sixty-nine Midwestern middle-class children and adolescents were tested on justice and care orientations when reasoning abstract and interpersonal moral dilemmas. Nona Lyons' ("Two Perspectives on Self, Relationships and Morality," Harvard Educational Review, 1983, 53, 125-145) scoring method was used to score subjects' responses. A 2(sex) × 2(age) analysis of variance run on the total justice and care scores, as well as each individual dilemma, supported Carol Gilligan's (In a Different Voice: Psychological Theory and Women's Development, Cambridge, MA: Harvard University Press, 1982) theory that two distinct ways of thinking about moral problems exist—justice and care—and are differentially related to gender. Girls emphasized the morality of care significantly more than justice. Contrary to Gilligan (1982) and Lyons (1983), however, boys in both age groups emphasized the morality of justice and care equally. Data from the interpersonal dilemmas using Lyons's (1983) coding scheme are consistent with J. Piaget (The Moral Judgment of the Child, New York: Free Press, 1966) and Lawrence Kohlberg ["The Cognitive-Developmental Approach," in D. A. Goslin (Ed.), Handbook of Socialization Theory and Research, Chicago: Rand McNally, 1969]: older subjects became more justice oriented and younger subjects emphasized the morality of care. Sex differences on Kohlberg's stage theory were not significant and the protagonist's gender in the Heinz dilemma had no effect on moral reasoning.

The study of moral development is currently expanding to include new ideas and considerations involving the process of moral development. In the past, Kohlberg's stage theory has been widely accepted as a conceptualization of an individual's moral growth. However, over the past decade many criticisms arose over the generalizability of his theory. Carol Gilligan (1977, 1982) suggested that people consider more than justice when reasoning moral conflicts, and found another "voice" concerned with care, relationships, and connections with other people especially salient in women. A great deal of con-

troversty has arisen over the universality of Kohlberg's theory, as some research has shown distinct sex differences where females score lower than males on Kohlberg's stages. This research will examine the suggestions that the failure of women to fit existing models of growth and testing may indicate a problem with representative theories of human development, and that an additional strategy for reasoning moral dilemmas, the Morality of Care, is used by males and females in moral reasoning, but has been excluded from popular developmental theories.

Influenced by Piaget (1966), Lawrence Kohlberg (1969) created a universal stage model to explain moral development. The scale was derived from extensive case analyses and two-hour interviews with boys ranging in age from 10 to 16 years. Kohlberg presented them with hypothetical abstract moral dilemmas where obedience to law and rules or commands of authority conflict with the needs or welfare of other persons (Kohlberg & Kramer, 1969). The idea of justice reflected in Kohlberg's model closely resembles Piaget's theory that children move toward "autonomous morality."

Kohlberg created a six-stage theory of moral development divided into three levels from the results of his study. In Stage 1, the child is egocentric and views everything as good and bad based on the individual's obedience to rules and authority. The second stage includes children whose behavior is motivated by self-satisfaction and occasionally satisfaction to others. These two stages comprise the Pre-conventional Level of morality. The next level of moral development according to Kohlberg, emphasizes conforming to societal values and norms, which he labels the Conventional Level. In Stage 3, children focus on gaining approval and pleasing others, while in Stage 4 they are simply "doing their duty" and maintaining the social order for its own sake. The last level of development, Moral Principles, encompasses people who behave according to a universal principle of justice. Stage 5 is defined by laws and rules that determine right and wrong behavior, and where duty and obligation are in terms of contract not individual needs. The last stage in Kohlberg's scheme is comprised of people who resolve conflicts by applying a universal principle (Kohlberg, 1969). Progression through stages depends on resolving cognitive disequilibrium and is fostered by role-taking opportunities. According to Kohlberg, the model is culturally universal, the stage sequence is invariant and hierarchical, and each stage is homogeneous (Kohlberg & Kramer, 1968). The theory assumes that increased autonomy and individuation lead to more advanced moral thinking. There have been several revisions in the scoring system over the years, and in the recent revision, Stage 6 has been eliminated (Colby, Gibbs, Kohlberg, Speicher-Dubin, & Candee, 1980).

A great deal of the literature supports Kohlberg's theory (Holstein, 1976; Rest, 1983; Rest, Davison, & Robinson, 1978; Sternlieb & Youniss, 1975;

Turiel, 1966; Walker, 1984). However, many people have challenged its validity (Bussey & Maugham, 1982; Fishkin, Keniston, & MacKinnon, 1973; Gilligan, 1977, 1982; Haan, 1975; Hoffman, 1984; Holstein, 1976; Rest, 1983; Simpson, 1974; Sullivan, 1977).

James Rest (1983) reviewed a dozen cross-sectional and longitudinal studies that test Kohlberg's theory. While he found general support for the model, he discovered some important qualifications: there was no improvement over time, and no clear evidence that Stage 6 follows Stage 5. Furthermore, he found inconclusive evidence that cognitive disequilibrium leads to moral growth and that the stage sequence is invariant.

Holstein (1976) found supportive evidence for a stepwise sequence in moral judgment from level to level, but not stage to stage, and only for the first two levels. Regression in Kohlberg's stages has been found as well (Bussey & Maugham, 1982; Fishkin et al., 1973; Holstein, 1976; Kohlberg & Kramer, 1969). Another problem may be that people use different strategies for reasoning real-life or interpersonal as opposed to hypothetical dilemmas (Haan, 1975).

Kohlberg's stage theory has been challenged as being sexually biased in favor of males (Baumrind, 1986; Bussey & Maugham, 1982; Gilligan, 1977, 1982; Haan, Smith, & Block, 1968; Holstein, 1976; Kohlberg & Kramer, 1969). Many people question the model's generalizability because it was derived from an all-male sample (Brabeck, 1983; Bussey & Maugham, 1982; Garwood, Levine, & Ewing, 1980; Gilligan, 1982) and also because the protagonist in each dilemma is male, making it difficult for females to relate to. Furthermore, in many investigations girls have been found to score lower than boys on Kohlberg's scale (Baumrind, 1986; Bussey & Maugham, 1982; Ford & Lowery, 1986; Gibbs, Arnold, & Burkhart, 1984; Haan et al., 1968; Holstein, 1976; Kohlberg & Kramer, 1968). Bussey and Maugham (1982) attribute this disparity to the nature of Stages 3 and 4, where Stage 3 has traditionally "feminine" characteristics of expressiveness and need to do good to gain approval, and Stage 4 accommodates the typical "masculine" role of law and order.

Other researchers have reported no sex bias on Kohlberg's scale (Brabeck, 1983; Gibbs et al., 1984; Hoffman, 1975; Kerber, Greeno, Maccoby, Luria, & Stack, 1986; Maccoby & Jacklin, 1974; Rest, 1979, 1983; Thomas, 1986; Walker, 1984), or that the females even scored higher than males (Garwood et al., 1980; Rest, 1979; Thomas, 1986). In a recent review of the literature, Lawrence Walker (1984) found 10 studies of adolescents and youths out of 108 that yielded significant sex differences. Of these, 8 showed differences favoring males. Walker attributes other sex difference findings to poor methodology and reliance on early stage theory definitions, and cites only a few consistent sex differences among children on Kohlberg's scale.

Thomas (1986) found girls were at no disadvantage when measured by Rest's (1979) Defining Issues Test, a measure of moral reasoning including six hypothetical dilemmas that involve competing social claims modeled after Kohlberg's dilemmas. However, Bussey and Maugham (1982) varied the sex of the protagonist in a series of dilemmas using Kohlberg's measure, and discovered that males were affected and females were not. Males dropped to Stage 3 reasoning with a female protagonist but maintained Stage 4 thinking with a male protagonist. Females, however, remained at Stage 3 reasoning regardless the protagonist's gender.

Thus sex differences on Kohlberg's measure have revealed ambiguous results. In addition, other concepts of morality have shown a very different pattern of sex differences, but these are also controversial. Empathy, guilt, altruism, and other helping behaviors also have been used to measure morality. Much of this research indicates sex differences favoring females (Brabeck, 1983; Eisenberg-Berg & Mussen, 1978; Hoffman, 1984). This body of research seems to support Gilligan, as prosocial behaviors seem to reflect caring considerations.

However, some researchers claim there are no sex differences in prosocial behaviors, and attribute the apparent evidence to poor methodology (Eisenberg & Lennon, 1983; Maccoby & Jacklin, 1974; Tavis & Offir, 1977). Tavis and Offir (1977) argue that girls have a reputation for being more altruistic, although in fact the sex differences are really quite small. Eisenberg and Lennon (1983) claim that most sex-disparity empathy findings occur on self-report standards of affective responses to others, and are not reliable.

The present investigation was derived mainly from Gilligan's (1982) work on understanding women's development as different from men's, and emphasizes the distinction between two moral voices, the Morality of Care and the Morality of Justice. Gilligan asserts that females develop differently than males, and as long as the categories by which development is assessed are derived from research on men, divergence from the masculine standard will inevitably be seen as a failure of development. Gilligan offers a new theory of development to account for women's experience, and identifies a distinct moral language among women that emphasizes the obligation to exercise care and avoid hurting others. Expanding on designs traditionally used to research moral development, Gilligan includes space for the differentiated, comprehensive, and reflective thought patterns of women. She found that women define themselves through others and their relationships, while men tend to separate themselves from the world, making it easier for them to relate to the hypothetical abstract dilemmas currently employed to assess moral development. For men, the moral imperative is to respect rights and protect the rights to life and self-fulfillment. Women, however, are more concerned

with the welfare of other people. Gilligan describes the stages of women's development to represent a more complete understanding of the relationship between self and other, where each transition involves a critical reinterpretation of the conflict between selfishness and responsibility.

People have criticized Gilligan's (1982) theory on various levels. Kerber et al. (1986) suggest that Gilligan's model contains an inherent assumption that women feel and men reason, which has traditionally kept women out of more "serious" roles in society. They also argue that Gilligan's method of analysis was too interpretative and not empirical enough.

Brabeck (1983) compares Gilligan's and Kohlberg's theories of moral development on their central points. For Gilligan, morality arises from conflicting responsibilities for women, rather than competing rights as it does for men. Hurt is women's central moral concern, superseding issues of fairness. Furthermore, Brabeck suggests that Gilligan's theory explains a morality of responsibility based on nonviolence and harmony, whereas Kohlberg's is based on fairness and reciprocity. While Gilligan distinguishes issues of attachments, self-sacrifice and selfishness, and considerations of relationships, Kohlberg is more concerned with separateness, rules, legalities, and considerations of the individual. For women, Brabeck continues, moral dilemmas are contextual and resolved through inductive thinking, but in Kohlberg's scheme moral principles are universal and applied to moral dilemmas through formal and abstract thinking. In sum, Kohlberg assumes people develop through stages that are universal, invariantly sequential, and hierarchically arranged, while Gilligan's stage theory is hierarchically arranged but not invariantly sequential and reflects women's growth. Brabeck concludes that both theories are essential and valuable in gaining an understanding of people's moral growth process.

Recent empirical work on sex differences in moral development has been done by Nona Lyons (1983). She interviewed male and female children, adolescents, and adults, and found supporting data for two distinct kinds of considerations used by people when making moral decisions – justice and care. In response to many criticisms of the lack of empirical evidence for Gilligan's (1982) model, Lyons created a methodology for testing and coding the data. Lyons more clearly defines the two moral voices used by men and women. According to Gilligan and Lyons, there are two ways of perceiving others and relating to others: (a) The perspective of the separate/objective self, which Lyons has labeled "reciprocity," is based on impartiality, objectivity, and the distancing of the self from others. (b) The perspective of the connected self, which Lyons calls "response," is based on interdependence and concern for another's well-being. According to Lyons (1983), the Morality of Justice emphasizes the rules and laws in society and reasons moral conflicts according to a set conception of "right and wrong." The Morality of Care, however,

focuses on the people involved, their feelings, and the respondent's relationship/responsibility to their well-being. Lyons (1983) and Gilligan (1982) found these two voices differentially related to gender but not exclusive to one sex.

To examine some of these issues, an empirical investigation was undertaken in which children and adolescents in a middle-sized Midwestern college community were interviewed to see how they solved abstract and interpersonal moral dilemmas. This study sought to answer the following questions: (a) Do girls emphasize the morality of care while boys emphasize the morality of justice, as suggested by Gilligan? (b) Are there age-related changes, as implied by Kohlberg, such that older subjects are more likely to emphasize justice while younger subjects are more likely to emphasize care? (c) Does the gender of the protagonist in the story affect moral reasoning for boys or girls?

METHOD

Subjects

Seventy-one Midwestern middle-class boys and girls from a college community participated in the study. Names of potential subjects were obtained through an elementary and a secondary school in the same district. Subjects selected were Caucasian and middle-class. One hundred and sixty letters were sent to potential subjects followed by a phone call to their homes. Of those contacted, thirty-eight 5th and 6th graders and thirty-three 10th and 11th graders agreed to participate. Six subjects were dropped from the study due to tape recording failures. Two 10th-grade interviews and four 5th-grade interviews were inaudible and could not be coded.

Design and Procedure

All data were obtained in individual interviews lasting from 15 to 20 minutes. These interviews were conducted either in a classroom or the school library during free time or after school. A tape recorder was used to (a) preserve the interview, (b) allow coding to be done at a later date, and (c) allow for smooth uninterrupted interviews. Subjects were assured that no right or wrong answer existed, and they were not expected to answer in any particular fashion; they were encouraged to answer whatever questions they could, and to refuse any they did not want to answer. A series of four moral dilemmas were read to each subject and s/he was asked to solve them. Fol-

lowing the procedure used by Lyons (1983), the experimenter asked probing questions in order to examine the subject's reasoning process, i.e., "Why do you choose to do that?" or "What would Kathy think about in trying to make that decision?" The probing questions following the dilemmas were divided into two categories of thought: how the subject constructed the problem and how s/he resolved the problem.

Measures

The first dilemma presented in the interview was Kohlberg's (1969) Heinz dilemma [see Table I(a)]. The experimenters could not apply Lyons's (1983) scoring method to this dilemma because of its abstract content. Thus, the Heinz dilemma was mainly used to assess sex and age differences on Kohlberg's stage theory of moral development, and to see if the sex of the protagonist affected the level of the response. Two versions of the Heinz dilemma were employed randomly to subjects in order to test for any effects of the protagonist's gender on the subject's moral reasoning. One version of the story read that a man's wife was dying and in need of a drug, while the other version read that a woman's husband needed the drug to live.

Second, an interpersonal moral dilemma created by Selman (1971) was employed [see Table I(b)]. Selman's adapted version of this interpersonal dilemma was used to accommodate the adolescent subject group. The conflict in the story remains basically unchanged [see Table I(c)]. Questions involving conflict resolution between the two actors in the story were asked following the discussion of the Selman dilemma.

Next, the subject was asked to describe a situation where s/he was not sure what to do. This is considered a real-life or personal dilemma provided by the subject in this study. While most subjects related a problem involving other people, if they did not they were encouraged to think of one involving others [see Table I(d)]. The research asked the same questions following the personal dilemma — i.e., "What did you think about in trying to decide what to do?" — that were asked for the other interpersonal dilemmas. The subject was also asked to evaluate his/her decision in the real-life dilemma.

The last dilemma presented in the interview was another hypothetical interpersonal situation created for this study and intended to be a common experience for all of the subjects [see Table I(e)]. The same follow-up questions were asked to assess the subject's reasoning process. Finally, some general questions about responsibility, caring, and morality were asked at the end of the interview, so that subjects' reasoning of the stories would not be affected by them.

Table I. Moral Dilemmas

(a) Heinz dilemma

A man's/woman's wife/husband is almost certain to die without a certain drug. The druggist who invented the drug won't sell it to the husband/wife except at a very high price, more than the husband/wife can possibly pay. He will not consider the possibility of receiving part of the money later. Should the husband/wife steal the drug to save his/her wife's/husband's life? Why?

(b) Selman friendship dilemma (children's version):

Kathy and Becky have been best friends since they were five years old. They went to the same kindergarten and have been in the same class ever since. Every Saturday they would try to do something special together, go to the park or the store, or play something special at home. They always had a good time with each other.

One day a new girl, Jeanette, moved into their neighborhood and soon introduced herself to Kathy and Becky. Right away Jeanette and Kathy seemed to hit it off very well. They talked about where Jeanette was from and the things she could be doing in her new town. Becky, on the other hand, didn't seem to like Jeanette very well. She thought Jeanette was a showoff, but was also jealous of all the attention Kathy was giving Jeanette.

When Jeanette left the other two alone, Becky told Kathy how she felt about Jeanette. "What did you think of her Kathy? I thought she was kind of pushy, butting in on us like that."

"Come on Becky. She's new in town and just trying to make friends. The least we could do is be nice to her."

"Yeah, but that doesn't mean we have to be friends with her," replied Becky. "Anyway, what would you like to do this Saturday? You know those old puppets of mine, I thought we could fix them up and make our own puppet show."

"Sure, Becky, that sounds great," said Kathy. "I'll be over after lunch. I better go home now. See you tomorrow."

Later that evening Jeanette called Kathy and surprised her with an invitation to the circus, the last show before it left town. The only problem was that Kathy had already promised to go to Becky's. Kathy didn't know what to do, go to the circus and leave her best friend alone, or stick with her best friend and miss a good time.

(c) Selman friendship dilemma (adolescent version):

Charlene and Joanne have been good friends since they were five. Now they were in high school and Joanne was trying out for the school play. As usual she was nervous about how she had done, but Charlene was there to tell her she was very good and give her moral support. Still Joanne was worried that a newcomer in the school would get the part. The new girl, Tina, came over to congratulate Joanne on her performance and then asked if she could join the girls for a snack. Right away Charlene and Tina seemed to hit it off very well. They talked about where Tina was from and the kinds of things she could do in her new school. Joanne, on the other hand, didn't seem to like Tina very well. She thought Tina was a little pushy, and maybe she was a bit jealous over all the attention Charlene was giving Tina.

When Tina left the other two alone, Joanne and Charlene arranged to get together on Saturday, because Joanne had a problem that she would like to talk over with Charlene. But later that day, Tina called Charlene and asked her to go see a play on Saturday.

Charlene had a dilemma. She would have jumped at the chance to go with Tina, but she had already promised to see Joanne. Joanne might have understood and been happy that Charlene had the chance to go, or she might feel like she was losing her best friend when she really needed her.

(d) The personal dilemma:

Have you ever been in a situation when you weren't really sure what to do? Can you describe it to me?

(e) The third dilemma:

Suppose two of your closest friends got into an argument and both were trying to convince you that their side was right. Both friends said they would be angry with you if you didn't join their side.

Scoring Moral Dilemmas

Two different scoring methods were used to assess (a) the Kohlberg dilemma and (b) the interpersonal dilemmas. The first method was Kohlberg's (1969) original system to determine subjects' stage score on the Heinz dilemma. The second method employed to score subjects' responses was quantitative increments developed by Lyons (1983) to allow for a parametric statistical analysis of the results. This scoring method was used to code subjects' reasoning of the three interpersonal dilemmas—Selman's, the real-life/personal (provided by the subject), and the third (created for this study).

Lyons's coding scheme divides people's moral reasoning into three categories: the construction, resolution, and evaluation of the problem. Subjects' responses were coded according to these distinctions and also determined to be either a caring or justice consideration as defined by Lyons's model. Lyons's scheme details what constitutes justice or rights concerns, as well as response or care considerations. Following this method, subjects were given a point in the appropriate response or rights category according to their construction, resolution, or evaluation of the dilemma. In other words, for each statement given by the subject, the coders assigned a category—construction, resolution, or evaluation of the problem—and then determined whether it was a caring or justice consideration. This method was applied to the Selman, real-life/personal, and third dilemmas. Finally, the number of caring responses and justice responses were totaled for each part of the problem (construction, resolution, and evaluation), as well as the total number of justice and care responses overall for each subject. The subject's dominant mode of thought was determined by the category, justice or caring, with the higher score.

Conflict Resolution

The series of conflict resolution questions that followed the Selman dilemma were scored on a 4-point scale according to a prepared set of possible answers (i.e., yes/depends, no, other). Another attempt to assess the degree to which subjects were oriented toward care or justice was a question about whether a right and/or kind thing to do exists in situations of moral conflict, and which is more important for society, a hospital or a police department.

General Questions

The general questions were coded based on two possibilities—internal vs. external motivation and self vs. others. When subjects defined responsi-

bility and caring as self-motivated (e.g., "If I promise to babysit my little brother, I am responsible to do that") rather than externally imposed (e.g., "If my mom tells me to clean my room, I am responsible for my room being clean"), then they were given an internal score rather than an external one. Additionally, when subjects told of a personal moral dilemma that presented a conflict between self and others, and chose others when responsibility to both was equally important, they were given an "other" score.

Coding

All of the interviews were coded by the senior author and one other person to reduce researcher bias. The other coder was trained to use Lyons's (1983) coding scheme, and was blind to the subject's gender and age while coding the interviews. The reliability between coders was .88. The general procedure was for both coders to listen to each interview together and assign a separate code. Any discrepancies were resolved in discussion and then assigned a code based on a joint decision. A new code was assigned with each new statement the subject made, but codes were not given twice when the subject merely repeated a previous statement in an attempt to explain the same thought.

RESULTS

A 2(sex) \times 2(age) \times 2(version) analysis of variance (ANOVA) was run on the Kohlberg stage score for the Heinz dilemma. The data indicate no significant effects due to age, sex, or the version of the story, and there was no interaction between any of the variables. Sex and sex by age effects were not significant. Although the differences were not significant, the total mean scores indicate that the older girls scored higher ($M = 4$) than all of the other groups on the version with a male protagonist. Younger boys reached the next highest stage for the same version ($M = 3.4$). Finally, the older girls also reached the highest stage for the version with a female protagonist ($M = 3.8$) compared with the older boys ($M = 3.5$) and younger boys ($M = 3.7$). The mean stage score for both sexes and ages was 3.5 (Table II).

A second ANOVA was run 2 (sex) \times 2 (age) with repeated measures, using the three dilemma stories to examine the effects of sex and age on moral voice. Statistically significant sex differences were found on the number of care vs. justice responses given by male and female subjects. Girls gave significantly more care-oriented responses than boys [$F(1, 62) = 23.19, p < .001$] and a greater number of responses overall than the boys [$F(1, 62) =$

Table II. Kohlberg Stage by Sex, Age, and Version of Story

	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls
Version 1	3.25	3.42	4.00	3.10
Version 2	3.50	3.75	3.77	3.50

3.97, $p < .10$). While older subjects gave more responses than younger subjects [$F(1, 62) = 10.01$, $p < .01$], more care than justice responses were given overall [$F(1, 62) = 26.16$, $p < .001$], but the excess of care over justice responses was most marked for girls (see Table III).

Scores were further broken down into two parts: the construction and resolution of the interpersonal dilemmas. A third ANOVA with repeated measures was run with 2 (sex) \times 2 (age) and construction and resolution as the within-subject variables. As expected, there were statistically significant sex and age differences, paralleling the previous results. Based on the percentage of justice responses, cell means indicate that children were more oriented toward care for the construction of the dilemma and more justice oriented for the resolution of the conflict ($M = .29$ for construction and $M = .41$ for the resolution; see Table IV). The data analysis revealed that boys were more justice oriented than girls [$F(1, 56) = 6.36$, $p < .05$], and older children emphasized the Morality of Justice more often than younger children [$F(1, 56) = 4.49$, $p < .05$]. Statistically significant evidence also indicated that both boys and girls use care considerations over justice for the construction of the dilemma more than they do for the resolution of the

Table III.

Number of care and justice responses by sex and age				
	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls
Care	7.61	6.05	11.50 ^a	10.41 ^a
Justice	7.94	5.35	5.71	3.82
NOVA for voice (no. of responses) by sex and age				
Source	<i>F</i> ratio	Significance level		
Sex (S)	3.97	$p < .10$		
Age (A)	10.01	$p < .01$		
Voice (V)	26.16	$p < .001$		
VS	23.19	$p < .001$		

^aSignificant at $p < .10$.

problem [$F(1, 56) = 9.13, p < .01$]. There were no significant sex and age interaction effects (Table IV).

The separate analyses of each dilemma yielded findings consistent with the general results. In each dilemma, girls used the Morality of Care more often than Justice, and more often than boys: Selman— $F(1, 62) = 13.71, p < .001$; Personal— $F(1, 62) = 13.20, p < .001$; Third— $F(1, 63) = 3.59, p < .10$. On the Selman and Personal dilemmas, we found older subjects gave more responses than younger subjects: Selman— $F(1, 62) = 4.22, p < .10$; Personal— $F(1, 62) = 14.76, p < .001$. In addition, complementing the general results, we found that on the Third and Selman dilemmas all subjects emphasized the Morality of Care over Justice: Selman— $F(1, 62) = 33.84, p < .001$; Third— $F(1, 63) = 5.78, p < .05$. Girls gave more responses than boys on the Selman dilemma [$F(1, 62) = 9.39, p < .01$] and older subjects showed a greater justice orientation on the Third dilemma than younger subjects [$F(1, 63) = 9.91, p < .01$]. (See Tables V–VII.)

No significant effects were found by age or sex on the conflict resolution or general questions. The data on the general questions paralleled the overall results of care and justice orientations. Subjects were generally more caring than justice oriented, and almost everyone believed a hospital was more important for society than the police department.

DISCUSSION

The data support in part and contradict in part the hypotheses of Kohlberg and Gilligan. There were no significant sex differences on Kohlberg's

Table IV.

Percentage of justice responses in the construction and resolution of the dilemmas by sex and age					
	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls	Marginal
Construction	0.38426	0.24305	0.33074	0.18437	0.29022 ^a
Resolution	0.54186	0.47660	0.34812	0.27154	0.41152 ^a
NOVA for voice in construction and resolution by sex and age					
Source	<i>F</i> ratio	Significance level			
Sex	6.36	$p < .05$			
Age	4.49	$p < .05$			
Task ^b	9.13	$p < .01$			

^a $p < .01$.

^bTask—the construction or resolution of the problem.

Table V.

Cell means for the Selman dilemma by voice, sex, and age				
	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls
Care	3.44	2.94	6.14	5.12
Justice	2.61	2.12	2.21	1.59
NOVA for Selman dilemma—voice—(no. of responses) by sex and age				
Source	<i>F</i> ratio	Significance level		
Sex (S)	9.39	$p < .01$		
Age (A)	4.22	$p < .10$		
Voice (V)	33.84	$p < .001$		
VS	13.71	$p < .001$		

scale as applied to the single story. This result agrees with some of the literature (Brabeck, 1983; Gibbs et al., 1984; Hoffman, 1975; Kerber et al., 1986; Maccoby & Jacklin, 1974; Rest, 1979, 1983; Thomas, 1986; Walker, 1984), but contradicts other reports (Baumrind, 1986; Bussey & Maughan, 1982; Gilligan, 1977, 1982; Haan et al., 1968; Holstein, 1976; Kohlberg & Kramer, 1968). People have criticized research that supports sex differences on Kohlberg's stage theory as wrought with methodological difficulties. Baumrind (1986), for example, cites evidence that a disparity between the sexes on Kohlberg's scale can be found when educational level, culture, and class are not controlled, and these variables were considered here in the choice of our subjects. This may explain why we found no sex differences. It is also possible that the failure to find sex differences reflects the fact that sex differences when found occur mainly in the last three stages, whereas means in this sample are between three and four.

Table VI.

Cell means for the personal dilemma by voice, sex, and age				
	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls
Care	2.17	1.00	3.71	2.06
Justice	3.00	1.76	1.79	1.41
NOVA for personal dilemma—voice—(no. of responses) by sex and age				
Source	<i>F</i> ratio	Significance level		
Age	14.76	$p < .001$		
Voice \times Sex	13.20	$p < .001$		

Table VII.

Cell means for the third dilemma by voice, sex, and age				
	10th- and 11th- grade boys	5th- and 6th- grade boys	10th- and 11th- grade girls	5th- and 6th- grade girls
Care	2.00	2.00	1.64	3.24
Justice	2.33	1.39	1.71	0.82
NOVA for third dilemma—voice—(no. of responses) by sex and age				
Source	<i>F</i> ratio	Significance level		
Voice (V)	5.78	<i>p</i> < .05		
VS	3.59	<i>p</i> < .10		
VA	9.91	<i>p</i> < .01		

Furthermore, our results did not show that scores are affected by the sex of the protagonist. The tendency for girls to reach a higher stage score when the protagonist was male may indicate a high ability to role play, and it is feasible that when girls are reasoning from a male perspective, they would role play from a more “masculine” point of view—thus, Stage 4. This was found among Bussey and Maugham’s (1982) subjects; however, the girls in their study remained the same and boys vacillated depending on the protagonist’s gender.

Contrary to Kohlberg, this study also failed to find significant age differences on the Kohlberg scale. This could, however, reflect the fact that only one of his dilemmas was included in the research. The data from the interpersonal dilemmas did indicate that older subjects were more justice oriented while younger subjects emphasized the morality of care.

As suggested by Gilligan, the data reveal two distinct modes of moral thought, justice and care, and while males and females use both strategies in making moral choices, there is a propensity for the two voices to be differentially related to gender. These results support the research that suggests males and females emphasize different concerns when thinking about moral conflicts (Bussey & Maugham, 1982; Gilligan, 1982; Huston, 1983; Johnston, 1985; Lyons, 1983). It may seem surprising, however, that in neither age group did boys emphasize the Morality of Justice significantly more than Care. This could be because of the nature of the dilemmas chosen here, which are more personal than abstract. However, it might also reflect the growing challenge to the male stereotype, where it has become easier for boys to be caring today than in the past. Hoffman (1977) has argued that both sexes are socialized into the “expressive” role, but upon entering adolescence, boys are subsequently pressured into the “masculine role” to succeed. Males, then, succumb to the pressure to achieve at the expense of being car-

ing. However, because the older male subjects here still young, it is possible that they have not yet been socialized into the masculine role in this sense, and may still maintain a caring approach to the moral dilemmas presented in this investigation.

It is interesting that, while girls emphasized care on the interpersonal dilemmas more than boys, it was the older girls who obtained the highest mean score on the Kohlberg Heinz dilemma. It seems likely that the content of a story's dilemma affects an individual's approach to solving the problem. Thus, if the conflict involves societal rules and expectations including the law, i.e., the Heinz dilemma, it is appropriate for the individual to concern her/himself with social principles when resolving the dilemma, but if the problem deals only with interpersonal relationships, caring concerns would seem to have precedent over abstract rules. Flexibility in solving moral dilemmas according to its particular content seems appropriate and even expected. In this study, the adolescent girls tended to score higher on Kohlberg's justice measure when the problem involved issues of rights and responsibilities to the law, but used a caring approach when presented with conflicts between friends, suggesting perhaps that compared to boys and younger girls, they showed more content-based flexibility in reasoning about the dilemmas.

The major findings of this study come from the interpersonal moral dilemmas scored according to the method developed by Lyons (1983). They support the hypothesis that there are two ways of thinking about moral issues, justice and care, and that these are related to both gender and age. But while girls emphasize the morality of care more than justice in responding to the interpersonal dilemmas, boys emphasize both equally. While both boys and girls show a shift toward justice responses with age, the older as well as the younger are clearly concerned with care as well as justice. Thus the data indicate that both boys and girls, in grade school and in high school, use both moral voices in reasoning moral conflicts. This result is also reported in a study of college students by Rothbart, Hanley, and Albert (1986). It is reasonable to assume that each of these moral orientations is valuable and that different situations may call for one more than the other. It may be, then, that future research should focus less on gender differences per se and investigate instead the process by which each moral voice, and the sensitivity to appropriately differentiate them, develop in both boys and girls.

REFERENCES

- Baumrind, D. Sex differences in moral reasoning: Response to Walker's (1984) conclusion that there are none. *Child Development*, 1986, 57, 511-521.
- Brabeck, M. Moral judgment: Theory and research on differences between males and females. *Developmental Review*, 1983, 3, 274-291.

- Bussey, K., & Maugham, B. Gender differences in moral reasoning. *Journal of Personality and Social Psychology*, 1982, 42, 701-706.
- Colby, A., Gibbs, J., Kohlberg, L., Speicher-Dubin, B., & Candee, D. *Standard form scoring manual*. Center for Moral Education, Harvard University, 1980.
- Eisenberg, N., & Lennon, R. Sex differences in empathy and related capacities. *Psychological Bulletin*, 1982, 94, 100-186.
- Eisenberg-Berg, N., & Mussen, P. Empathy and moral development in adolescence. *Developmental Psychology*, 1978, 14, 185-186.
- Fishkin, J., Keniston, K., & MacKinnon, C. Moral reasoning and political ideology. *Journal of Personality and Social Psychology*, 1973, 50, 109-119.
- Ford, R. M., & Lowery, R. C. Gender differences in moral reasoning: A comparison of the use of justice and care orientations. *Journal of Personality and Social Psychology*, 1986, 50, 777-785.
- Garwood, S. G., Levine, D. W., & Ewing, L. Effect of protagonist's sex on assessing gender differences in moral reasoning. *Developmental Review*, 1980, 16, 677-678.
- Gibbs, J. C., Arnold, K. D., & Burkhart, J. E. Sex differences in the expression of moral judgment. *Child Development*, 1984, 55, 1040-1043.
- Gilligan, C. In a different voice: Woman's conception of self and morality. *Harvard Educational Review*, 1977, 47, 481-517.
- Gilligan, C. *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press, 1982.
- Haan, N. Hypothetical and actual moral reasoning in a situation of civil disobedience. *Journal of Personality and Social Psychology*, 1975, 32, 255-270.
- Haan, N., Smith, M. B., & Block, J. Moral reasoning of young adults: Political-social behavior, family background, and personality correlates. *Journal of Personality and Social Psychology*, 1968, 10, 183-201.
- Hoffman, M. L. Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Developmental Psychology*, 1975, 11, 607-622.
- Hoffman, M. L. Sex differences in empathy and related behaviors. *Psychological Bulletin*, 1977, 84, 712-722.
- Hoffman, M. L. Moral development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook*. Hillsdale NJ: Lawrence Erlbaum Associates, 1984.
- Hoffman, M. L. Sex differences in moral internalization and values. *Journal of Personality and Social Psychology*, 1975, 32, 720-729.
- Holstein, C. Development of moral judgment: A longitudinal study of males and females. *Child Development*, 1976, 47, 51-61.
- Huston, A. C. Sex-typing. In J. Flavell & E. Makhani (Eds.), *Carmichael's manual of child psychology*, 4th ed. New York: Wiley, 1983.
- Johnston, K. *Two moral orientations—Two problem solving strategies: Adolescents' solutions to dilemmas in fables*. Dissertation, Education Department, Harvard Graduate School of Education, 1985.
- Kerber, L., Greeno, G. C., Maccoby, E. E., Luria, Z., & Stack, C. On *In a Different Voice*: An interdisciplinary forum. *Signs*, 1986, 11, 304-324.
- Kohlberg, L. The cognitive-developmental approach. In D. A. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally, 1969.
- Kohlberg, L., & Kramer, R. Continuities and discontinuities in child and adult moral development. *Human Development*, 1969, 12, 93-120.
- Lyons, N. Two perspectives on self, relationship and morality. *Harvard Educational Review*, 1983, 53, 125-145.
- Maccoby, E. E., & Jacklin, C. *The psychology of sex differences*. Stanford, CA: Stanford University Press, 1974. (See Chap. 8, "Sex Typing and the Role of Modeling.")
- Piaget, J. *The moral judgment of the child*. New York: Free Press, 1966.
- Rest, J. R. *Development in judging moral issues*. Minneapolis: Minnesota University Press, 1979, pp. 120-124.
- Rest, J. R. Morality, cognitive development. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*, 4th ed., New York: Wiley, 1983.

- Rest, J. R., Davison, M. L., & Robbins, S. Age trends in judging moral issues: A review of cross-sectional and longitudinal studies of the defining issues test. *Child Development*, 1978, 49, 263-249.
- Rothbart, M. K., Hanley, D., & Alberg, M. Gender differences in moral reasoning. *Sex Roles*, 1986, 15, 645-653.
- Selman, R. Taking another's perspective: Role-taking development in early childhood. *Journal of Experimental Child Psychology*, 1971, 24, 343-357. (a)
- Selman, R. The relation of role-taking to the development of moral judgements in children. *Child Development*, 1971, 42, 79-91. (b)
- Simpson, E. L. Moral development research: A case study of scientific culture bias. *Human Development*, 1974, 17, 81-106.
- Sternlieb, J. L., & Youniss, J. Moral judgements one year after intentional or consequence modeling. *Journal of Personality and Social Psychology*, 1975, 31, 895-897.
- Sullivan, E. V. A study of Kohlberg's structural theory of moral development: A critique of liberal social science ideology. *Human Development*, 1977, 20, 352-376.
- Tavris, C., & Offir, C. *The longest war*. New York: Harcourt Brace Jovanovich, 1977.
- Thomas, S. J. Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review*, 1986, 6, 165-180.
- Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. *Journal of Personality and Social Psychology*, 1966, 3, 611-618.
- Walker, L. D. Sex differences in the development of moral reasoning: A critical review. *Child Development*, 1984, 55, 677-691.