Gender Differences in Moral Reasoning: A Comparison of the Use of Justice and Care Orientations

Maureen Rose Ford
University of Kentucky

Carol Rotter Lowery
Lexington, Kentucky

The present study examines the adequacy of Lawrence Kohlberg’s cognitive-developmental model as a representation of female moral reasoning. Specifically, this study examines the claims of Gilligan (1982) that there are two different conceptions of morality—one described as a morality of justice, on which Kohlberg’s scheme is based, and one described as a morality of care, seen by Gilligan as more representative of female thinking about moral conflict. A sample of 202 college students (Mage = 19 years) filled out a self-report questionnaire on moral dilemmas they had experienced. They then rated their use of both justice and care orientations in resolving those dilemmas. The use of the two orientations was examined in relationship to subject gender, sex role, and perceptions of the two orientations. Few significant differences were obtained except that female subjects were more consistent in their use of a care orientation, and that male subjects were more consistent in their use of a justice orientation, and more feminine males were more likely to report the use of a care orientation than less feminine males. Male and female reasoning about moral conflict is examined in the light of these two perspectives and the relationship of sex roles to endorsement of each perspective is investigated.

In 1932, Piaget published his first major work on moral development, The Moral Judgment of the Child. In it he outlines a three-step process in which there is a basic progression from an external to an internal locus for evaluating “right” and from a more concrete to a more abstract mode of reasoning. The basic moral structure is seen to parallel the development of logical reasoning, and advancements in the latter are seen as necessary for advancements in the former. This is the structural approach to moral development on which Kohlberg built his cognitive-developmental model.

When discussing the first stage of moral development, Piaget commented:

The relations between parents and children are certainly not only those of constraint. There is a spontaneous mutual affection, which from the first prompts the child to acts of generosity and even self-sacrifice, to very touching demonstrations which are in no way pre-scribed. And here no doubt is the starting point for the morality of good which we shall see developing alongside of the morality of right or duty, and which in some persons completely replaces it. (1932, pp. 193–194)

Thus, in 1932, Piaget noted the existence of two different types of morality, a morality of good and a morality of right or duty. Yet, the work of Carol Gilligan has raised this as an empirical question 50 years later. To understand the 50-year gap, one need only look a few pages further in Piaget’s original work. He goes on to say that though difficulty—we mean the notion of justice. It will therefore be on this point that most of our efforts will be directed. (p. 193)

And thus, the “morality of good” drops out of Piaget’s analysis of moral development.

Almost a quarter of a century later, Lawrence Kohlberg began a longitudinal study of the moral development of adolescent males. Working with Piaget’s theory, he sought and found evidence for a structural developmental model of the type described by Piaget. He proposed his universal, invariant sequence, six-stage model of moral development wherein each stage reflects a more advanced social perspective and logical structure—a structure which parallels Piaget’s logical stages and is “best formulated as a structure” (Kohlberg, 1971, p. 195).

Kohlberg argued that justice is “the basic moral principle” (1971, p. 220). Kohlberg noted that “the only general principle of content, other than justice, seriously advanced by (moral) philosophers, has been the principle variously termed utility or benevolence.” Although admitting that, like justice, benevolence can be universalized, he nevertheless saw benevolence as insufficient because “it cannot resolve a conflict of welfares” (p. 220).

With his dismissal of benevolence as being inadequate for solving dilemmas, Kohlberg nevertheless acknowledged that without such a moral attitude, moral conflict is not even experienced. It is on this point that much of the criticism against Kohlberg turns. Peters (1971), for example, discusses the importance of other issues besides knowledge of justice in the moral development of the child. The child, he says, “might know what justice is, but not care about it overmuch. . . . How do children come to care? This seems to be the most important question in moral education; but no clear answer to it can be found in Kohlberg’s writings” (p. 262).

Many other writers and researchers have taken issue with Kohlberg’s model as to its sufficiency in explaining moral development (e.g., Holstein, 1976; Simpson, 1974). The particular...
criticism of Kohlberg's failure to include the concept of care along with the concept of justice is receiving more attention in the psychological literature recently, principally because of the work of Carol Gilligan. Her interest is the issue that in choosing to focus on an ethic of justice, Kohlberg creates a model of adult moral development whose focus on autonomy, separation, and individuation leaves women at a distinct disadvantage. Traditional female qualities, such as a sense of responsibility to and for others, are qualities almost exclusively relegated to Stage 3 in Kohlberg's model, the stage just above the egocentric concerns of the pre-conventional stage. It is an "unfair paradox," as Gilligan says, "that the very traits that have traditionally defined the 'goodness' of women, are those that mark them as deficient in moral development" (Gilligan, 1982, p. 18).

The question of whether men and women do differ in stages of moral reasoning has been extensively debated in the literature. Langdale (1980) cites many studies which found that women did not score as high on Kohlberg's scale as did men. Several studies (e.g., Bussey & Mangan, 1982; Poppen, 1974) suggest that Stage 3 is the modal stage for women, and that Stage 4 is the modal stage for men. Given the findings that many adults never make it past this conventional level of moral development (Kuhn, Langer, Kohlberg, & Haan, 1977), the issue of whether Stages 3 and 4 are sex-typed and possibly alternate conventional stages is an important one in terms of their use in indicating degree of moral development.

Gilligan, however, objects to the whole process of trying to fit women into models that were developed on men and that in her view, do not adequately represent the development of the female person. In order to construct models of development which represent the "voices" of women, Gilligan believes we must return to the source, much as Piaget did, and listen to women's discourse as they struggle to make decisions about moral conflicts in their lives (1982, p. 20).

Using an open-ended interview, Gilligan asked women considering whether to continue or abort a pregnancy questions regarding how they were thinking about the decision (1977). Her subjects were 29 women, ranging in age from 15 to 33, diverse in ethnic background and social class. What Gilligan found as she listened to these women was that their moral language was replete with words like selfishness, responsibility, care and avoiding hurt and that their decisions were weighed in light of their relationships with others. Using Kohlberg's and Piaget's preconventional, conventional, and postconventional scheme, she traces a developmental path of a morality of care. But unlike Kohlberg's scheme, which ties the development of morality as justice to the changing understanding of equality and reciprocity, the development of morality as care is tied to a changing understanding of responsibility and relationships.

Having "found" this second conception of morality, a morality centered on issues of care, Gilligan and her colleagues sought to investigate its prevalence (Gilligan, Langdale, Lyons, & Murphy, 1982). They used a cross-sectional sample of men and women matched for age, educational experience, and social class. Subjects were asked to describe a personal, real-life experience of moral conflict, and they were then asked a series of standard follow-up questions inquiring into how the subject constructed, resolved, and evaluated the conflict. The results reported were that 75% (12) of the women used a predominantly care orientation, whereas only 25% (4) of the women used a predominantly justice orientation. For men, 79% (11) used a predominantly justice orientation and 14% (2) used a predominantly care orientation, with 7% (1) using both equally. There were no women who failed to present a consideration of care, and no men who failed to present a consideration of rights. However, 36% (6) of the women failed to present any consideration of justice, and 36% (6) of the men failed to present any consideration of care. The author's conclusion was that "in real-life moral conflict, individuals call upon and think about considerations predominantly within one mode which is related to, but not defined by, a person's gender" (Gilligan et al., 1982).

The purpose of the present study was to investigate further these findings of differential use by men and women of the justice and care orientations to moral conflict. Given the observation that these orientations are not completely gender specific, a measure of psychological sex roles was included to investigate whether sex role may be more predictive than gender for the use of either orientation. One study (Pritz & Royer, 1982) has found this to be the case. The present study stays with Gilligan's suggestion that moral decisions should be studied in the context of subjects' real-life dilemmas. However, when Gilligan used a derivative methodology in her second study, she trained raters to code the language used by subjects. To avoid the potential confound of gender-related differences in the semantics of expressing moral decisions rather than in the moral reasoning process itself, the present study uses a more direct methodology. The present study also controls for bias that may be introduced by having subjects report their own dilemmas due to possible gender-related differences in self-disclosure that are incidental to moral judgment.

**Method**

**Subjects**

Subjects were 101 male and 101 female undergraduates from introductory psychology classes who volunteered to participate. The average age of subjects was 19 years, with a range of 18 to 29 years. They had no known prior exposure to the concepts or theories addressed by the study.

**Procedure**

Subjects were asked to describe three important moral conflicts in their lives. After each, they were asked to rate the importance of the conflict in their life at that time and the degree of difficulty they experienced in making their decision about what to do. Both ratings were done on 7-point Likert scales. After describing and rating all three conflicts, subjects read summary descriptions of the justice and care orientations to moral conflict. Each summary consisted of a five-sentence description of the orientation; both had been previewed by Gilligan's colleagues (N. P. Lyons, personal communication, October 1982) and designated as accurate representations of the two orientations. They were then asked to rate, on 7-point Likert scales, the degree to which each orientation was a part of their own thinking about each of the conflicts they described.

Subjects completed two additional measures: the Interpersonal Disposition Inventory (IDI) and selected items from the Semantic Differential. The IDI is an 85-item scale developed by Berzins, Welling, and Wetter (1977) to measure psychological sex roles. Scores are obtained on masculinity and femininity subscales, and median splits on these subscale scores are used to divide subjects into high and low masculinity and
MORAL REASONING

femininity groups, and into four sex role categories (masculine, feminine, androgynous, and indeterminate).

The Semantic Differential is a standardized procedure for measuring the connotation of a concept for an individual (Os Guinness, Suci, & Tanenbaum, 1957). Factorial analyses of the original set of 50 scales revealed that concepts are judged on three major independent dimensions: Evaluative, Potency, and Activity factors. Three sets of adjectives from each of these factors (a total of 9 pairs) made up the differential used in this study. The adjectives pairs chosen were those that had relatively high loadings on the factor in question and relatively low loadings on the other two factors. In order to evaluate the connotation of each orientation, the two paragraphs summarizing the justice and care orientations were presented separately and subjects completed a differential on each, thus yielding ratings of both on all nine bipolar scales.

The entire procedure took about 2 hr, with subjects working in small groups and reporting their responses on forms provided.

Fifty subjects were asked to complete the justice and care ratings on their conflicts again, 3 to 4 weeks later. These ratings were used as a measure of the reliability of the rating procedure.

In a post hoc procedure, 50 subjects' protocols (25 men and 25 women) were selected at random and their first moral conflicts submitted to a content analysis. This was done to assess the extent to which the issues presented in the subjects' dilemmas were embedded in relationship and justice contexts. These ratings were done on 7-point Likert scales by two independent judges trained to be familiar with the care and justice concepts as presented in this study. These ratings focused on the content of the dilemmas only, not on their construction or evaluation. These ratings were thus quite distinct from the subjects' ratings of the extent to which these concepts had been a part of their thinking about the dilemmas.

Results

Because subjects were generating their own conflicts, direct between-subjects comparisons of the ratings of the two conceptions of morality were not possible without some preliminary analyses. Results to be considered first are the reliability of the moral conflict ratings and the relationship of the importance and difficulty ratings to gender and use of moral orientation.

The reliability of the justice and care ratings was assessed by retesting a subset of 50 subjects. Reliability coefficients were calculated separately for men and women, with interesting results. Again summed across all three dilemmas, the correlations obtained were men, $r = .76, p < .001$, on justice ratings, $r = .39, p < .12$, on care ratings; women, $r = .17, p < .34$, on justice ratings, $r = .70, p < .001$, on care ratings. Fisher's $r$ to $z$ transformation showed that men were significantly more consistent in their use of the justice orientation than women (Fisher's $z = 4.88, p < .05$) and that women were significantly more consistent than men in the use of the care orientation (Fisher's $z = 2.69, p < .05$).

The pattern of results shown in Table 1 indicated that only the ratings of the first conflicts were moderately reliable measures of the justice and care orientations for both men and women. Thus, between-subjects comparisons on the justice and care ratings were made only on the first set of ratings.

<p>| Table 1 |
| Test-Retest Correlations of Justice and Care Ratings for the Three Conflict Situations |</p>
<table>
<thead>
<tr>
<th>Situation</th>
<th>Men</th>
<th>Men</th>
<th>Women</th>
<th>Women</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justice</td>
<td>$r$</td>
<td>$p$</td>
<td>$r$</td>
<td>$p$</td>
<td>$r$</td>
</tr>
<tr>
<td>Conflict 1</td>
<td>.61</td>
<td>.009</td>
<td>.40</td>
<td>.023</td>
<td>.41</td>
</tr>
<tr>
<td>Conflict 2</td>
<td>.43</td>
<td>.086</td>
<td>.13</td>
<td>.489</td>
<td>.21</td>
</tr>
<tr>
<td>Conflict 3</td>
<td>.57</td>
<td>.018</td>
<td>.17</td>
<td>.350</td>
<td>.27</td>
</tr>
<tr>
<td>Overall</td>
<td>.76</td>
<td>.001</td>
<td>.17</td>
<td>.340</td>
<td>.31</td>
</tr>
</tbody>
</table>

Care

<table>
<thead>
<tr>
<th>Situation</th>
<th>Men</th>
<th>Men</th>
<th>Women</th>
<th>Women</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict 1</td>
<td>.59</td>
<td>.013</td>
<td>.51</td>
<td>.003</td>
<td>.51</td>
</tr>
<tr>
<td>Conflict 2</td>
<td>.20</td>
<td>.433</td>
<td>.65</td>
<td>.001</td>
<td>.52</td>
</tr>
<tr>
<td>Conflict 3</td>
<td>.35</td>
<td>.168</td>
<td>.53</td>
<td>.001</td>
<td>.47</td>
</tr>
<tr>
<td>Overall</td>
<td>.39</td>
<td>.120</td>
<td>.70</td>
<td>.001</td>
<td>.57</td>
</tr>
</tbody>
</table>

Note. In the justice and care orientations, $N = 17$ for men, $N = 33$ for women, and $N = 50$ combined.

Because subjects were generating their own conflicts, direct between-subjects comparisons of the ratings of the two conceptions of morality were not possible without some preliminary analyses. Results to be considered first are the reliability of the moral conflict ratings and the relationship of the importance and difficulty ratings to gender and use of moral orientation.

The reliability of the justice and care ratings was assessed by retesting a subset of 50 subjects. Reliability coefficients were calculated separately for men and women, with interesting results. Again summed across all three dilemmas, the correlations obtained were men, $r = .76, p < .001$, on justice ratings, $r = .39, p < .12$, on care ratings; women, $r = .17, p < .34$, on justice ratings, $r = .70, p < .001$, on care ratings. Fisher's $r$ to $z$ transformation showed that men were significantly more consistent in their use of the justice orientation than women (Fisher's $z = 4.88, p < .05$) and that women were significantly more consistent than men in the use of the care orientation (Fisher's $z = 2.69, p < .05$).

The pattern of results shown in Table 1 indicated that only the ratings of the first conflicts were moderately reliable measures of the justice and care orientations for both men and women. Thus, between-subjects comparisons on the justice and care ratings were made only on the first set of ratings.

Importance and Difficulty as Covariates

In order to assess how properties of the content of the subjects' dilemmas may have been related to their preference for one or the other moral orientations, subjects were asked to rate the importance and difficulty of each conflict. The correlation between importance and difficulty for the first dilemma was $r = .53, p < .001$. Though statistically significant, the correlation was too weak to justify combining importance and difficulty as a single measure of "significance," so both were retained as separate measures. Subjects' ratings of the importance and difficulty of their first conflicts were compared with their ratings of their use of the justice and care orientations. These correlations are given for men and women, combined and separately; in Table 2. Both measures of significance were more highly correlated with care ratings than with justice ratings.

A one-way analysis of variance (ANOVA) on importance by sex yielded a significant sex difference, $F(1, 200) = 4.86, p < .05$. The same analysis on difficulty ratings yielded a marginally significant result, $F(1, 200) = 2.95, p < .05$. In both cases, female mean scores were higher than the male mean scores (see Table 3). These results indicate that men and women may have been providing themselves with different stimuli; that is, women rated their self-reported conflicts as significantly more important in their lives and their decisions as more difficult to make.

To test whether men and women would differ in their use of the justice and care orientations, the effects of importance and difficulty just noted needed to be controlled. Therefore, a multivariate analysis of covariance (MANCOVA) was used (see Table 3). Analyzing justice and care ratings by sex, with importance and difficulty as covariates, yielded nonsignificant sex differences, $F(2, 197) = .398, p < .67$.

To evaluate whether psychological sex role may be a mediating variable in man/woman use of the different conceptions of morality, subjects were divided into high and low masculinity, and into high and low femininity categories (Berlins et al., 1977). Analyses were also performed with subjects classified into the four sex role categories: masculine (high MASCUL, low FEMIN), feminine (high FEMIN, low MASCUL), androgynous (high MASCUL, high FEMIN) and indeterminate (low MASCUL, low FEMIN).

A 2 x 2 MANCOVA on the justice and care ratings, using sex
and high/low masculinity as the between-subjects variables, and importance and difficulty as covariates, showed no multivariate effects for sex or masculinity classification and no interaction. The same analysis with high/low femininity as the between-subjects variable yielded a significant multivariate effect for femininity classification, $F(2, 195) = 3.00, p < .05$, and a trend toward significance for the interaction of sex and femininity classification, $F(2, 195) = 2.31, p < .10$.

The significant effects of the latter MANCOVA were subsequently analyzed by univariate ANOVAs. For care ratings, there was a significant main effect for femininity classification, $F(1, 196) = 6.0, p < .02$, moderated by a significant sex by femininity classification interaction, $F(1, 196) = 4.56, p < .03$. Mean comparisons were conducted using the Tukey HSD procedure (Wilk, 1971) and indicated that men classified as high femininity had higher care ratings than men classified as low femininity. Univariate ANOVAs on justice ratings showed no significant effects for femininity classification or the interaction of sex and femininity classification. In summary, masculinity classification did not significantly affect justice or care ratings. Femininity classification affected only care ratings and only for men. Mean justice and care ratings by sex and high/low masculinity and femininity are listed in Table 4.

Analyses were also performed by the four sex role categories, because the four categories reflect all possible combinations of scores on gender orientation. A $2 \times 4$ MANCOVA on the justice and care ratings, using sex and sex role as the between-subjects variables, with importance and difficulty as covariates, showed nonsignificant differences between sex role categories, no significant main effect for sex, and a nonsignificant interaction between sex and sex role.

### Table 3
**Mean Importance, Difficulty, Care, and Justice Ratings For the Three Conflict Situations**

<table>
<thead>
<tr>
<th>Situation</th>
<th>Men</th>
<th>Women</th>
<th>Men</th>
<th>Women</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance</td>
<td>.527</td>
<td>.575</td>
<td>.479</td>
<td>.522</td>
<td>.529</td>
<td>.558</td>
</tr>
<tr>
<td>Difficulty</td>
<td>.529</td>
<td>.575</td>
<td>.479</td>
<td>.522</td>
<td>.529</td>
<td>.558</td>
</tr>
<tr>
<td>Care</td>
<td>.469</td>
<td>.457</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice</td>
<td>.162</td>
<td>.151</td>
<td>.175</td>
<td>.178</td>
<td>.168</td>
<td>.152</td>
</tr>
</tbody>
</table>

In all situations, $N = 101$.

### Table 4
**Mean Justice and Care Ratings by Sex and High/Low Masculinity and Femininity Classification**

<table>
<thead>
<tr>
<th>Rating</th>
<th>Men</th>
<th>SD</th>
<th>n</th>
<th>Women</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Masculinity</td>
<td>4.52</td>
<td>1.43</td>
<td>63</td>
<td>4.39</td>
<td>1.95</td>
<td>38</td>
</tr>
<tr>
<td>Low Masculinity</td>
<td>5.0</td>
<td>1.49</td>
<td>36</td>
<td>4.68</td>
<td>1.59</td>
<td>63</td>
</tr>
<tr>
<td>High Femininity</td>
<td>4.59</td>
<td>1.43</td>
<td>37</td>
<td>4.56</td>
<td>1.72</td>
<td>69</td>
</tr>
<tr>
<td>Low Femininity</td>
<td>4.59</td>
<td>1.43</td>
<td>64</td>
<td>4.59</td>
<td>1.79</td>
<td>32</td>
</tr>
</tbody>
</table>

* Means a and b differ at the $p < .05$ level.

### Semantic Differential

Subjects rated the justice and care paragraphs on a 7-point differential made up of three independent semantic dimensions: Evaluative, Potency, and Activity factors. Mean dimension ratings for each orientation are listed in Table 5. Because Osgood's dimensions are orthogonal, a repeated measures ANOVA was used. This mixed design ANOVA, using gender as the between-subjects factor, and moral orientation and the three semantic dimensions as the within-subjects factors, showed a significant main effect for the semantic dimensions, $F(2, 394) = 484.95, p < .001$, a significant interaction between the dimensions and the two orientations, $F(2, 394) = 21.67, p < .001$, and a significant sex by dimensions interaction, $F(2, 394) = 6.63, p < .01$.

The interactions noted were further evaluated by the Bonferroni $t$ post hoc comparison procedure (Wilk, 1971). The care and justice orientations differed significantly on the potency factor, $t(201) = 6.17, p < .01$, with the justice orientation being rated as more potent than the care orientation. The two orientations were not significantly different on the evaluative and activity factors, which were seen as equally good and active approaches to moral conflict. Investigation of the sex by dimensions interaction indicated nonsignificant differences between men and women in their ratings on the three dimensions. Mean ratings of the two orientations on the potency scale are shown in Table 6.

### Content Analysis

Fifty subjects' first dilemmas were rated by two independent judges to assess the extent to which the content of the dilemmas was imbedded in care and justice contexts. Interrater reliability was computed by the Pearson formula and found to be adequate for research purposes: for care ratings, $r = .72, p < .01$, and for justice ratings, $r = .79, p < .01$. The mean ratings across the two raters on justice and care for men, and justice for women, were identical ($M = 4.56$). The mean rating for women on care was somewhat higher ($M = 4.94$), but not significantly different from the justice ratings for women (as evaluated with a $t$ test). These
Table 5
Mean Semantic Dimension Ratings for the Justice and Care Orientations*

<table>
<thead>
<tr>
<th></th>
<th>Justice</th>
<th></th>
<th>Care</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Men</td>
<td>Women</td>
<td>Combined</td>
<td>Men</td>
</tr>
<tr>
<td>Evaluate</td>
<td>M</td>
<td>2.12</td>
<td>2.03</td>
<td>2.08</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.08</td>
<td>1.05</td>
<td>1.06</td>
</tr>
<tr>
<td>Potency</td>
<td>M</td>
<td>3.55</td>
<td>3.71</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.01</td>
<td>1.17</td>
<td>1.10</td>
</tr>
<tr>
<td>Activity</td>
<td>M</td>
<td>3.96</td>
<td>3.85</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.90</td>
<td>.81</td>
<td>.86</td>
</tr>
</tbody>
</table>

Note. In the justice and care orientations for men and women, N = 101; N = 202 in the justice and care orientations combined.
*The semantic differential is constructed in such a way that lower scores indicate higher evaluation, potency, and activity ratings.

results suggest that the contents of male and female dilemmas did not substantially differ in the extent to which they were embedded in contexts of care and justice.

Discussion

One of the purposes of this study was to examine whether the gender differences Gilligan reported in her research could be found across a set of conflicts. Gilligan's conclusions were based on single conflict investigations (Gilligan, 1982; Gilligan et al., 1982). Generalizations of her findings across three conflict situations would provide stronger support for her claim that men and women approach moral conflicts from basically different orientations. Unfortunately, such direct comparisons could not be made, as test-retest coefficients showed the ratings to be unreliable measures of the justice and care orientations for all but the first conflict. However, the finding that there were very different reliability patterns for men and women on the justice and care ratings provides indirect support for Gilligan's assertions. For women, the care ratings of the three conflicts were significantly more stable from one time to the next than were the justice ratings. Exactly the reverse was true for men. It would seem, then, that despite findings of differences or similarities in actual ratings of the two orientations for a given conflict, the care orientation is a consistent consideration for women, and the justice orientation is a consistent consideration for men.

Worell and Worell (1965), in a study of personality conflict, suggest such intradividual variability can be viewed as a lawful product. Highly variable responses to the same stimuli are seen as indications of competing response tendencies. In a situation where none of the responses are clearly dominant, the subjects respond with any one of their closely competing responses. In contrast, the person with a dominant response is less conflicted and will tend to repeat the response in similar situations and thus show less variability. In terms of this study, this perspective would suggest that for women, considerations of care are consistent dominant responses. Thus their sense of the rule care plays in their moral dilemmas is nonconflicted and therefore their ratings of the care orientation less variable from one time to the next. In contrast, considerations of justice are less dominant for women, and their sense of the role of justice in their moral dilemmas less clear and more conflicted. This leads to greater variability in justice ratings in response to the same stimulus, and thus lower coefficients of reliability. For men, the justice perspective would be the dominant response tendency and thus shows less variability. The care orientation would reflect more competing response tendencies and thus be more variable.

The decision to allow subjects to generate and rate their own conflicts rather than respond to standardized dilemmas made it necessary to evaluate in some manner the relative meaning or significance of the conflicts in their lives. This was necessary so that any differences along this dimension could be controlled when assessing differences in the justice and care orientations. Differences in the importance and difficulty of the conflicts were also of interest relative to the use of the different orientations.

The results showed that the importance of the conflict in the subject's life and the difficulty of the decision were in fact significantly associated with care ratings but not with justice ratings. It seems, then, that for both men and women, the more important and more difficult moral decisions they have made up to this point in their lives are more related to issues of care than to issues of justice. Recalling that Kohlberg's dilemmas are constructed as abstract problems of opposition—competing rights, competing values—it is worth noting that such abstract issues were not at the heart of the important moral conflicts for the subjects in this sample.

The correlations between care and importance help in clarifying a possible confounding variable in this study, which is that women generally rated their conflicts as more important and more difficult than men did. This suggests that the two groups of subjects were providing themselves with different stimuli, because all subjects were asked to describe the most significant conflicts they had faced up to that point in time, an obvious observation is that women perceived themselves as having had to make more important and difficult moral decisions than men did. Given the relationship between importance, difficulty, and

Table 6
Mean Ratings of the Justice and Care Orientations on the Three Potency Scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Justice</th>
<th></th>
<th>Care</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Severe-leniency</td>
<td>M</td>
<td>3.99</td>
<td>3.72</td>
<td>4.28</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.52</td>
<td>1.59</td>
<td>1.63</td>
</tr>
<tr>
<td>Masculine-feminine</td>
<td>M</td>
<td>3.24</td>
<td>3.80</td>
<td>3.96</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.11</td>
<td>1.19</td>
<td>1.26</td>
</tr>
<tr>
<td>Hard-soft</td>
<td>M</td>
<td>3.41</td>
<td>3.60</td>
<td>4.16</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.44</td>
<td>1.59</td>
<td>1.53</td>
</tr>
</tbody>
</table>

Note. In the justice and care orientations, N = 101.
care already mentioned, it would seem possible that the reason women rated their conflicts at more important and difficult was because the conflicts they wrote about were those in which issues of care were the central concern, more so than in the conflicts of male subjects. There is support in the literature for such a speculation (Block, 1973; Maccoby & Jacklin, 1974); however, a post hoc content analysis of a sample of subjects’ dilemmas did not suggest that this was the case. Though females’ dilemmas were rated as presenting issues somewhat more embedded in care or relationship contexts than were males’ dilemmas for the sample selected, this difference was not statistically significant.

The question addressed in part by this content analysis is whether women select and report dilemmas defined by relationship issues out of a pool of dilemmas similar to the pool that men experience, or whether women in actuality experience more relationship dilemmas because of their social context. Kohlberg believed the second hypothesis was true and that that was why women gravitated toward Stage 3. He felt women would progress to higher levels of moral development if they experienced more varied contexts (Kohlberg & Kramer, 1969). Gilligan believes the first is true, that women identify relationship issues in their conflicts more than men. Gilligan = Yes!

Turning to the major purpose of this study, it was expected that men would show greater usage of the justice orientation than women, and women greater usage of the care orientation than men. Additionally, it was expected that psychological sex role would mediate the relationship somewhat, with high femininity being more associated with the use of care and high masculinity being more associated with the use of justice. Men and women showed a tendency to differ in the predicted directions in the use of justice and care, but the differences were very small and statistically nonsignificant. Both sexes apparently considered questions of relationship, care, and responsibility, as well as questions of fairness, justice, and rights, and they considered them fairly equally. In accounting for these results in light of the very different social patterns already mentioned, one might speculate as to whether subjects could have a primary orientation but, when presented with the other, see it as a reasonable and valuable approach and rate their conflicts in that light. Regardless, this pattern of results differs considerably from the more extreme findings of Gilligan and her colleagues. In their studies, subjects’ descriptions of their conflicts were coded by the researchers for considerations of justice and care. The researchers, in essence, were coding the language used by the subjects to describe their conflicts. It is possible that the language used by the subjects in their own descriptions truly differed and was not tapped by the methodology of this study. However, Gilligan may have encountered the same confound that was a concern in this study: women were perhaps presenting somewhat different dilemmas than men, and ones in which the nature of the dilemma pulled more for one orientation than the other. The content analysis in this study did not indicate significant differences in the extent to which the contents of male and female dilemmas were embedded in contexts of care and justice. However, it is difficult to guarantee the equivalence of subject generated dilemmas, an equivalence that would be necessary to address this question. To sort out the influence of the content of the dilemmas, and concentrate on the issue of subject identification of care or justice issues, it would seem necessary to present standardized dilemmas that are equated or balanced for the extent to which the content is embedded in justice or care contexts.

Moving from gender to sex roles, levels of femininity proved an important influence on considerations of care issues. High scorers on the IDI femininity scale have been characterized as nurturant, affiliative, and self-subordinating (Berzins et al., 1977). It was not surprising, therefore, to find that persons with high levels of femininity rated considerations of care more highly than persons with low levels of femininity. It was interesting, however, that this was true almost exclusively for men, with women high and low on femininity rating considerations of care about the same. Levels of masculinity did not significantly influence care ratings for men or women.

In regard to justice ratings, levels of masculinity did not prove to be an important influence on use of the justice orientation. In understanding this result, we need to return to the definition of the masculinity construct as defined for the IDI. High scorers on the masculinity scale have been described as dominant, instrumental, autonomous, and oriented toward physical risk. Though the characteristic of autonomy can be viewed as part of the noncontextual, removed-in-judgment approach to moral conflict described as the justice orientation, characteristics such as dominance and orientation toward risk are less theoretically related to the justice orientation than the construct of femininity is to the care orientation.

The semantic differential was used in this study to examine the meaning of the justice and care orientations to men and women. The first question of interest was whether the subjects would indicate one orientation to be clearly “better” than the other on any of these dimensions. Examination of the three factors indicates that both orientations were seen as highly positive and of average activity by both women and men. The only significant difference was on the potency factor with the justice orientation being seen as more potent than the care orientation. The three pairs of adjectives that made up this dimension were severe—lesser, masculine—feminine, hard—soft, a very sex role stereotypic set of adjectives. A check of the mean scores revealed that the justice orientation was clearly seen as more masculine and the care orientation as more feminine. Thus, though both orientations were equally valued, one was perceived as a masculine response and one as a feminine response.

Taken together, the findings from this study suggest some gender differences in use of the justice and care orientations but raise questions about the strength and nature of these differences. The complex pattern of findings also raises the question of whether women reported more relationship concerns because they generally cast the conflicts they encounter in those terms or whether their social context means they experience more conflicts dealing with issues of care than men do.

In the only study in which Gilligan and her colleagues directly compared males and females (Gilligan et al., 1982), subjects’ dilemmas were coded for considerations of response and rights but were not equated for content. Thus the conclusions about differences in orientation between males and females may also be largely reflective of the kinds of conflicts her male and female subjects chose to discuss. That female conflicts reflected more considerations of care, and male conflicts reflected more consid-
erations of rights is a significant finding, and should not be mini-
mized, but it may be no more than a reflection of the different
arenas and self-definitions out of which the majority of males
and females operate. It does not directly address the question of
whether, in a given conflict situation, females would focus on
issues of relationship, responsibility, and care, and males would
focus on issues of rights, rules, and justice. Again, to get at
this latter question it seems necessary to return to a standardized
dilemma format.

This study provides some support for Gilligan's assertions that
females are more attuned to issues of care in moral conflicts and
males more attuned to issues of justice. However, it also supports
the conclusion that the realm of care is not an exclusively female
realm nor justice an exclusively male realm. Brabec (1983),
after surveying the literature on moral reasoning, empathy and
altruism, concludes that sex differences in morality are at best
minimal and are not consistently found. She raises the question
as to why so many people find Gilligan's claims intuitively ap-
pealing and believe they speak to an essential truth, even when
there is no clear empirical support, or in the case of moral rea-
soning, there is evidence which contradicts her claims. Her answer
is that we may be dealing with a mythic truth rather than an
empirical truth. Mentioning studies that found boys and girls to
be perceived differently in their helping behaviors when behaving
very much the same (Shigetomi, Hartmann, & Gilford, 1981),
she suggests we may have some need to perceive males and fe-
males as morally different. Indeed, in the present study, males
and females did not significantly differ in their use of the two
orientations, yet they rated the justice orientation as masculine
and the care orientation as feminine. Brabec suggests this myth
meets our emotional and cultural needs.

There is an essential tension between autonomy and interde-
pendence, between the requirements of justice and the demands of mercy, be-
tween absolute moral principles and situation specific moral action,
between reason and affect. To resolve this tension by assigning half
to males and half to females when evidence does not support that
division is to reduce the complexity of morality, to cloud truth with
myth, to do an injustice to the capacities of both sexes and to lose
an opportunity to revise and modify our theories of morality. (Brabec,
1983, p. 287)

Though this study indicates there may in fact be some differ-
ences in the moral orientations of men and women, the inves-
tigator agrees with Brabec that Gilligan's significant contribution
cannot be in suggesting that men and women differ in their orienta-
tions to moral conflict, but in broadening our definition of
what constitutes an adequate description of the moral rea-
soning process.

References

scale users manual-revised. Unpublished materials, University of Ken-
tucky.

Block, J. H. (1973). Conceptions of sex role: Some cross-cultural, lon-


Gilligan, C. (1977). In a different voice: Women's conception of the self

Gilligan, C. (1982). In a different voice. Cambridge, MA: Harvard Uni-
versity Press.

Gilligan, C., Langdale, S., Lyons, N., & Murphy, M. (1982). The contribu-
tion of women's thought to developmental theory: The elimina-
tion of sex bias in moral development research and education. Final Report:
National Institute of Education.

Holstein, C. B. (1976). Irreversibility, stepwise sequence in the develop-
ment of moral judgment: A longitudinal study of males and females. Child
Development, 47, 51-61.

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic
fallacy and get away with it in the study of moral development. In T. Mischel

Kohlberg, L. (1976). Moral stages and moralization: The cognitive de-
velopmental approach. In T. Lickona (Ed.), Moral development and be-
havior: Theory, research, and social issues. New York: Holt, Rinehart
& Winston.

Kohlberg, L., & Kramer, R. (1969). Continuities and discontinuities in
childhood and adult moral development. Human Development, 12,
92-120.

of formal operations in logical and moral judgment. Genetic Psychology
Monographs, 95, 97-188.

Langdale, S. (1980). Conceptions of morality in developmental psychology:
Is there more than justice? Unpublished manuscript, Harvard Uni-
versity.


Osgood, C., Suci, G., & Tannenbaum, P. (1957). The measurement of

Mischel (Ed.), Cognitive development and epistemology. New York:
Academic Press.

Piaget, J. (1932). The moral judgment of the child (M. Gahain, Trans.).

Poppes, P. J. (1974). The development of sex differences in moral judg-
ment for college males and females. Dissertation Abstracts International,
35(2-B), 1108.

don't mix: Sex and sex-role patterns in moral judgment orientation.

Shigetomi, C. C., Hartmann, D. P., & Gilford, D. M. (1981). Sex differ-
ces in children's altruistic behavior and reputations for helpfulness.
Developmental Psychology, 17(4), 434-437.

and cultural bias. Human Development, 17, 81-106.


Worel, J., & Worel, L. (1965). Personality conflict, originality of response,

Received March 19, 1985
Revision received August 6, 1985