

REAL-LIFE MORAL JUDGMENT

by

Gillian R. Wark

B.A. (Hon.), Simon Fraser University, 1991

M.A., Simon Fraser University, 1993

Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy
in the Department
of
Psychology

© Gillian R. Wark 1996

Simon Fraser University

March, 1996

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Gillian R. Wark
Degree: PhD
Title of Thesis: Real-Life Moral Judgment

Examining Committee:

Chair: David Cox, PhD

~~Dennis Krebs, PhD~~
Professor
Senior Supervisor

Kim Bartholomew, PhD
Assistant Professor

~~Patricia Kerig, PhD~~
Assistant Professor

Meredith Kimball, PhD
Associate Professor
Internal/External Examiner

Lawrence Walker, PhD
Professor
Dept. of Psychology, U.B.C.
External Examiner

Date Approved _____

3-28-96

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Real-Life Moral Judgment

Author: _____
(signature)

Gillian R. Wark

(name)

• April 1st, 1996

(date)

Abstract

In the context of Kohlberg's (1984) and Gilligan's (1982, 1988) models of moral judgment, three studies investigated the relation of gender and type of moral dilemma to moral judgment. In Study 1, 30 male and 30 female participants listed the main issues they saw to be involved in descriptions of real-life dilemmas. The types of issue listed were determined by both the dilemma and within-person factors, but the latter were not gender-related. In Study 2, 15 male and 15 female participants made moral judgments regarding some of the dilemmas. In Study 3, 40 male and 40 female participants made moral judgments about dilemmas on Kohlberg's test and about two antisocial and two prosocial real-life moral dilemmas they had experienced. There were no overall gender differences in moral judgment. Moral judgment varied across dilemmas. A new approach to the study of real-life moral judgment is discussed.

Table of Contents

Abstract	iii
List of Tables	v
I. Introduction	1
II. Study 1	6
Method	7
Participants	7
Procedure	8
Classification of issues	9
Scoring of issues	10
Results	10
Discussion	14
III. Study 2	22
Method	23
Participants	23
Procedure	24
Scoring	25
Results	27
Discussion	28
IV. Study 3	32
Method	35
Participants	35
Procedure	36
Scoring	38
Results	39
Discussion	46
V. General Discussion	56
References	60
Author's notes	65
Footnotes	66

List of Tables

Table	Page
1. Descriptions of issue categories	69
2. Types of issue listed as a function of type of dilemma	73
3. Types of issue listed consistently across dilemmas	76
4. Examples of classified moral judgments	77
5. Moral orientation as a function of gender and type of dilemma	78
6. Examples of classified real-life moral judgments	79
7. Percent of participants obtaining the same or adjacent stage across types of dilemma	82
8. Mean moral maturity and moral orientation scores as a function of gender and type of dilemma	83
9. Percent of participants at each moral stage as a function of type of dilemma	85
10. Percent of participants obtaining the same or adjacent moral orientations score across types of dilemma	86
11. Mean care and justice likert scores	87
12. Mean ratings as a function of type of dilemma and question	88

Real-Life Moral Judgment

Kohlberg's (1984; Colby & Kohlberg, 1987) and Gilligan's (1982, 1988) cognitive-developmental models of moral judgment share the assumption that differences in moral judgment stem primarily from internal differences among people. According to Kohlberg, moral judgment is determined by cognitive "structures of the whole" that define stages of moral development. According to Gilligan, moral orientation--defined as the tendency to construct moral problems in care-oriented or justice-oriented ways--stems from gender-related differences in cognitive orientation and in identity acquired during early socialization. Both theorists assume that moral judgment is consistent within people across varying content or types of moral dilemma. For Kohlberg, people are expected to display only the form of moral judgment characteristic of their current stage of moral development or, if they are in transition, of their current stage and the next, more advanced, stage. For Gilligan, people are expected to view moral problems in predominantly justice-oriented or in predominantly care-oriented ways.

In support of Kohlberg's (1984) assumption that moral judgment stems primarily from internal, cognitive structures, Colby and Kohlberg (1987) have demonstrated that the structure (moral stage) of people's moral judgments across the dilemmas on Kohlberg's test is highly consistent--91% of Kohlberg's participants made moral judgments at the same or an adjacent stage across all nine dilemmas on his test. However, Kohlberg's test and scoring system may exaggerate stage consistency (Krebs, Vermeulen, & Denton, 1991). Research on the structural consistency of moral judgment has been reviewed by Krebs, Vermeulen, Carpendale, and Denton (1991) and Krebs, Vermeulen, and Denton (1991). Some studies have supported Kohlberg's assumption that moral judgment is structurally homogeneous (Bush, Krebs, & Carpendale, 1993; Higgins, Power, & Kohlberg, 1984; Linn, 1987a; Walker, de Vries, & Trevethan, 1987), but several studies have found that people invoke significantly lower stage forms of moral judgment in response to nonKohlbergian moral dilemmas than they do in response to the dilemmas on

Kohlberg's test (Bartek, Krebs, & Taylor, 1993; Carpendale & Krebs, 1992, in press; Denton & Krebs, 1990; Higgins et al., 1984; Linn, 1984, 1987b; Pratt, Golding, & Kerig, 1987; Wark & Krebs, 1996). In particular, people have been found to invoke lower stage structures when reasoning about dilemmas they have actually experienced (e.g., Wark & Krebs, 1996).

According to Kohlberg (1984), moral development involves the transformation and displacement of old stages by new, more advanced stages of moral reasoning. Transformational-displacement models of moral development maintain that people cannot invoke lower stage forms of moral reasoning because they are no longer available. The evidence, however, instead supports an "additive-inclusive" model of stage change (see Levine, 1979) in which older (lower) stage-structures are retained and invoked in response to certain types of moral dilemma (Carpendale & Krebs, 1992; Krebs, Vermeulen, Carpendale, & Denton, 1991; Krebs, Vermeulen, & Denton, 1991). If the stage-structures people invoke in response to the hypothetical dilemmas on Kohlberg's test are different from the forms of moral judgment they invoke in response to different types of dilemma, it becomes important to investigate how and why people invoke certain forms of moral judgment in response to certain moral dilemmas, particularly in the context of everyday life.

In a recent study, Wark and Krebs (1996) asked a sample of males and females to make moral judgments about the dilemmas on Kohlberg's test, and to report and to make judgments about an impersonal real-life dilemma and a personal real-life moral dilemma¹. Wark and Krebs classified each real-life moral dilemma into one of the following four main types: philosophical dilemmas involving classic moral issues such as abortion, war, euthanasia, and capital punishment; antisocial dilemmas involving reacting to transgressions or resisting temptation; prosocial dilemmas involving issues about loyalty or helping; and social pressure dilemmas involving reacting to pressure from peers or parents². The authors found that the hypothetical dilemmas on Kohlberg's

test and the impersonal, philosophical types of real-life dilemma evoked Stage 3/4 and higher moral judgments. In contrast, the personal, prosocial dilemmas evoked Stage 3 moral judgments and the personal, antisocial dilemmas evoked Stage 2 and 2/3 moral judgments.

Thus, contrary to Kohlberg's (1984) assumption of within-person structural consistency, Wark and Krebs (1996) found that the structure of moral reasoning was determined primarily by the type of dilemma about which people made moral judgments. Indeed, when all the moral judgments made by each participant across the Kohlbergian, impersonal, and personal dilemmas were examined, Wark and Krebs found that 85% of their participants made moral judgments that spanned between three and six substages on a 9-point scale (e.g., Stage 1, Stage 1/2, Stage 2, etc.), with 25% of the participants making moral judgments that spanned five or more substages.

Like Kohlberg, Gilligan attributed differences in moral judgment to internal differences between people. Unlike Kohlberg, Gilligan (1982, 1986, 1988) claimed that males and females are inclined to think about moral problems in qualitatively different ways. In particular, Gilligan (1988) argued that males think about moral problems primarily in terms of equality or justice, "symbolized by the balancing of scales" (p. xviii), and females think about moral problems primarily in terms of care, "symbolized by a network or web" (p. xviii). According to Gilligan (1982, 1988), this gender difference in moral orientation stems from gender-related differences in identity acquired during early socialization. The controversial nature of Gilligan's (1982) assertion stems primarily from her claim that Kohlberg's conception of moral development is biased against females and people with a feminine gender-role identity because it devalues their care-based moral orientation. According to Gilligan (1982), Kohlberg's scoring system assigns the care-based moral judgments females prefer to lower stages (Stage 3) than the justice-based moral judgments preferred by males (which are classified as Stage 4 or

higher). Walker (1984; 1991), however, reviewed the research on moral development and concluded that there are no unqualified gender differences.

Gilligan and Attanucci (1988a) asked a sample of people to report and to make judgments about moral dilemmas they had experienced in their everyday lives. The authors scored the judgments as either justice based or care based and, for each participant, determined the relative number of justice-based to care-based judgments made. In support of Gilligan's (1982, 1988) theorizing, Gilligan and Attanucci found that the majority of their participants made either predominantly care-based moral judgments or predominantly justice-based moral judgments about the real-life dilemmas they had reported. Those who made predominantly care-based moral judgments were primarily female and those who made predominantly justice-based moral judgments were primarily male. Although Gilligan and Attanucci (1988a) interpreted these findings as support for Gilligan's position, other investigators have questioned their methods and conclusions (see Vasudev, 1988, and Gilligan & Attanucci, 1988b, for a rejoinder). Particularly problematic, though neglected by Vasudev (1988), was a possible confound between gender and type of real-life dilemma. In the Gilligan and Attanucci study, each participant reported a different real-life dilemma. It is possible that the observed gender difference in moral orientation occurred because the female participants reported more moral dilemmas involving care-based issues than the male participants reported, and male participants reported more moral dilemmas involving justice-based issues than the female participants reported.

Indeed, in all the studies that have examined gender or gender-related (e.g., gender-role) differences in stage of moral judgment and/or moral orientation on real-life moral dilemmas (Donenberg & Hoffman, 1988; Ford & Lowery, 1986; Gilligan & Attanucci, 1988a; Lyons, 1983; Pratt et al., 1987; Pratt, Golding, Hunter, & Sampson, 1988; Rothbart, Hanley, & Albert, 1986), it is unclear whether the gender differences that were observed stemmed from internal, cognitive differences between females and males or from the tendency for females and males to report, and therefore to make judgments about, different types of moral dilemma. Post hoc analyses by Pratt et al. (1987, 1988), Walker et al. (1987), and Yussen (1977) revealed that females reported more personal

and relationship-based real-life dilemmas than males reported, and Walker et al. (1987) found that these types of real-life dilemma evoked more care-based moral judgments than other types of dilemma.

Wark and Krebs (1996) examined this issue and found that when males and females were asked to report and to make judgments about a personal real-life dilemma, females made more care-based moral judgments than males, which is consistent with Gilligan's (1982, 1988) theorizing. However, the authors also found a gender difference in the types of personal dilemma males and females reported: five times as many females as males reported prosocial types of dilemma, and twice as many males as females reported antisocial types of dilemma. (An equal number of males and females reported social pressure dilemmas.) The prosocial dilemmas evoked more care-based moral judgments than other types of dilemma from both males and females, and when type of dilemma was controlled, the gender difference in moral orientation disappeared. There was a tendency for females to make more care-based moral judgments than males to the social pressure dilemmas, but the difference was not statistically significant.

Gilligan's (1982, 1988) assertion that people tend to construct moral problems in terms of one moral orientation also was not supported by the results of the Wark and Krebs (1996) study. When the authors examined all the moral judgments made by each participant across the Kohlbergian, impersonal, and personal dilemmas, none of the participants viewed all dilemmas in predominantly care-based or justice-based ways, and only 8.5% of the participants obtained the same moral orientation score on all three dilemmas.

One of the limitations of the Wark and Krebs (1996) study, as with virtually all studies on gender differences in real-life moral judgment, is that the different real-life dilemmas reported by their participants were treated as equivalent. For example, Wark and Krebs equated all the dilemmas within categories (e.g., prosocial and antisocial) even though each dilemma in each category was unique. The primary purpose of the present research was to determine whether male and female participants would view the same stimuli--the same prosocial, antisocial, and social pressure real-life dilemma--in the same ways, or

whether females would view them in more care-based ways than males would view them. Would the type of dilemma determine the way in which it was viewed, or would internal cognitive orientations produce gender differences in moral judgment? To examine this issue, I conducted three studies. In the first two studies, the relation of gender and type of moral dilemma to moral orientation was investigated. In the third study, the relationship of gender and type of moral dilemma to moral maturity and moral orientation was investigated.

Study 1

In this study, I investigated (a) the extent to which people perceive the issues in dilemmas in terms of internal cognitive structures or "moral orientations" and the extent to which characteristics of dilemmas determine how they are perceived, and (b) Gilligan's (1982, 1988) claim that females view or construct moral problems in more care-based terms than males do. To determine whether male and female participants would view the same moral dilemmas as involving the same or different issues, I asked university students to (a) read general, composite descriptions of two prosocial types of real-life moral dilemma, two antisocial types of real-life moral dilemma, and two social pressure types of real-life dilemma--dilemmas representative of the types of dilemma experienced by people of their age--and to indicate what issues they perceived the dilemmas as involving and (b) rate each type of dilemma on the extent to which they viewed it as involving care-based versus justice-based issues.

If Gilligan's (1982, 1988) position is correct, I would expect females to list more care-based issues than males across dilemmas; I would expect males to list more justice-based issues than females would across dilemmas. According to Gilligan, females should rate the dilemmas higher in care and lower in justice than males rate them. Gilligan (1982, 1988) asserted that people tend to construct moral problems in terms of their moral orientation. However, the results of the Wark and Krebs (1996) study led to the expectation that different dilemmas will be perceived in different ways. In particular, I

expected the prosocial dilemmas to be perceived as involving primarily care-based issues and the antisocial dilemmas to be perceived as involving primarily justice-based issues.

In the Wark and Krebs (1996) study, there were no overall gender differences found in moral orientation. However, there was a noteworthy trend for females to make more care-based moral judgments than males made to social pressure types of dilemma. Based on these results, I investigated the possibility that females would perceive the social pressure dilemmas as involving more care-based issues than males would perceive them as involving.

If, as Wark and Krebs (1996) suggested, gender differences in real-life moral orientation stem primarily from gender differences in the types of moral dilemma reported by males and females--that is to say, if there are no gender differences in moral orientation when type of dilemma is controlled--it becomes important to understand why males and females report different types of real-life dilemma. I investigated four possible sources of such a difference: males and females may differ in (a) the extent to which they experience different types of moral dilemma, (b) the significance they attach to them, (c) their willingness to disclose different types of dilemma, and (d) the extent to which they view different types of dilemma as moral in nature.

Method

Participants

The sample was composed of 30 male and 30 female undergraduate students who volunteered for the study to fulfill a psychology course requirement. The average age of males was 22 (range = 17 - 29) and of females was 21 (range = 17 - 40), $t(58) = 1.06$, *ns*. The mean SES ratings (based on parents' occupation and income) on a scale ranging from one to five for males was 4.00 and for females was 3.78, $t(43) = 0.77$, *ns*. The mean grade point average of males was 3.03 and of females was 3.02, $t(45) = 0.10$, *ns*. All but two of the participants (one male and one female) were unmarried³.

Procedure

Participants were tested individually in one session. After giving informed consent, participants were asked to complete a questionnaire containing (a) a request for demographic information and (b) instructions to read general descriptions of six real-life moral dilemmas (two of each of the three main types reported by participants in the Wark and Krebs, 1996, study)⁴. Two dilemmas were prosocial in nature; two were antisocial in nature, and two involved social pressure. The types of dilemma are described in detail in the Wark and Krebs study. The first type of prosocial dilemma involved deciding whether to engage in proactive behavior by intervening in another's business or personal affairs for his or her sake, such as telling a friend that he or she is in a bad relationship, expressing concern about a friend's health, or assuming responsibility to support a friend. The second type of prosocial dilemma involved deciding whether or to whom to be loyal when, for example, one is torn between two friends or relatives who dislike each other but with whom one wants to maintain relationships, or one is deciding whether or not to see a friend's ex-partner. The first type of antisocial dilemma involved deciding whether to be honest with others by resisting the temptation to commit a transgression such as cheating at a game, receiving more pay than one deserves, or stealing money. The second type of antisocial dilemma involved reacting to a transgression, such as shoplifting or cheating on an exam, committed by a friend. The first type of social pressure dilemma involved deciding what to do when a friend pressures one to engage in identity-inconsistent behaviors that violate one's values, such as using drugs or alcohol, theft, or premarital sex. The second type of social pressure dilemma involved deciding whether to do what one's parents want and expect of one, or to do what one wants when, for example, choosing an academic career, religion, or lifestyle. The dilemmas were presented in random order⁵.

After participants read the description of the first dilemma, they were instructed to respond to the following open-ended question, "What did you see to be the main issue(s)

involved in the decision; that is, what do you see to be at stake?" Next, participants were asked to respond to five questions, set up in a 7-point Likert format, which asked to what extent they (a) had experienced this type of dilemma in their own lives, (b) deemed it significant, (c) would be willing to discuss it on an anonymous questionnaire if they were experiencing it, and (d) viewed it as constituting a moral dilemma. In addition, after participants read definitions of care and justice (see Lyons, 1983), they were asked to what extent they viewed the dilemma as involving issues of care versus justice. Participants were then asked the same questions about the second type of dilemma, then about the third, and so on. The questionnaire took approximately 1 hour to complete.

Classification of Issues

Together with a research assistant, I read all the issues each participant listed for each of the six types of dilemma. I explored several ways of classifying the issues, ultimately settling on a system that distinguished between four main categories: issues concerned with upholding justice, self-oriented issues, other-oriented issues, and issues concerned with relationships. Within these categories, I identified several subtypes of issue (see Table 1 for descriptions and examples).

Insert Table 1 about here

All issues were classified in the categories in Table 1 by the authors and a research assistant. A second research assistant, who was not involved in the original classification and who was blind to the purposes of the study and to all other information about the participants, classified two-thirds of all the issues. Interrater reliability was 78% agreement.

Gilligan and her colleagues (Gilligan & Attanucci, 1988a; Lyons, 1983) have defined moral considerations concerned with relationships, caring, the promotion of the welfare of others or prevention of their harm, understanding others in their own terms, and

context-dependent decision as care based, and considerations concerned with conflicting claims between self and other (including society) and the maintenance of impartial rules, principles and standards, particularly those of fairness and reciprocity, as justice based. According to this definition, all but one of the other-oriented and relationship-type issues in Table 1 are care based. The issue about respecting others seems both care based, in its concern for others, and justice based, in its concern for rights and autonomy. All but one of the upholding justice and self-oriented issues are justice-based. The procedural fairness concern with contexts or circumstances seems care based.

Scoring of Issues

To determine both the frequency and the consistency with which particular issues were listed across dilemmas, participants were given one point for each issue classified in a particular category for each dilemma. To illustrate, if a participant's list of issues for one dilemma consisted of two statements about procedural fairness and one statement about upholding relationships, he or she would receive two points for procedural fairness, one point for upholding relationships, and no points for the remaining categories.

Results

Gender and Dilemma-Based Variations in Types of Issue Listed

A total of 1031 issues were classified. Across dilemmas, 19% of the issues were classified in the upholding justice category, 37% of the issues were classified in the self-oriented category, 24% of the issues were classified in the other-oriented category, and 20% of the issues were classified in the relationship category. There were no overall gender differences in types of issue listed or number of issues listed. Fifty-nine percent of the justice-based issues were listed by males, and 49% of the justice-based issues were listed by females. Forty-nine percent of the care-based issues were listed by males, and 51% of the care-based issues were listed by females. The mean number of issues listed across dilemmas was 12 (range = 7 - 33).

To determine the extent to which participants viewed particular types of dilemma in terms of particular types of issue, I examined the distribution of issues listed for each dilemma. The types of issue listed for each dilemma are reported in Table 2. The numbers in Table 2 represent the percent of issues classified within a particular issue category. For example, of all the issues listed for the antisocial, temptation dilemma, 29% of them were about consequences to self.

Insert Table 2 about here

As the data in Table 2 indicate, the most frequently listed types of issue for the antisocial, temptation dilemma were self-oriented. The most frequently listed type of issue for the antisocial, transgression dilemma was about combatting immorality. The most frequently listed types of issue for the two prosocial dilemmas were relationship-type issues and issues about caring for others. The most frequently listed types of issue for the social pressure dilemmas involved autonomy.

Both care-based and justice-based types of issue were listed for all dilemmas, but some dilemmas pulled more strongly for one or the other (see Table 2). Tests for differences within dilemmas on the proportion of care-based issues versus justice-based issues listed were conducted. Participants listed a higher proportion of care-based issues than justice-based issues for the prosocial dilemmas ($z_s = 9.2$ and 20.4 , $ps < .01$), and participants listed a higher proportion of justice-based issues than care-based issues for the antisocial dilemmas ($z_s = 4.33$ and 16.0 , $ps < .01$), and for the social pressure dilemmas ($z_s = 12.0$ and 15.3 , $ps < .01$).

Tests for differences between dilemmas on the proportion of care-based issues listed were also conducted. Participants listed a higher proportion of care-based issues for the prosocial dilemmas than for any of the other dilemmas ($z_s = 6.8 - 18.6$, $ps < .01$).

Participants listed a higher proportion of care-based issues for the prosocial, helping

dilemma than for the prosocial, loyalty dilemma ($z = 5.2, p < .01$), and participants listed a higher proportion of care-based issues for the antisocial, transgression dilemma than for the antisocial, temptation dilemma and the social pressure dilemmas ($z_s = 3.9 - 5.2, ps < .01$).

Within-Person Consistency

To determine the extent to which individuals view moral dilemmas in terms of internal, cognitive moral orientations, I examined the extent to which participants listed the same type of issue across different dilemmas, that is, the degree of within-person consistency. The criterion for high within-person consistency was set at listing the same issue for three (50%) or more of the dilemmas. According to this criterion, 34 participants (16 males and 17 females) listed one type of issue consistently. Sixteen of these participants (eight males and eight females) listed one type of issue consistently across four or more dilemmas. And five of these participants (three males and two females) listed one type of issue consistently across five dilemmas. No participants listed one issue across all six dilemmas.

Table 3 contains the types of issue participants listed consistently across dilemmas. Both care-based types of issue (e.g., caring for others, nature of relationships) and justice-based types of issue (e.g., consequences to self, self-respect) were listed consistently across dilemmas. Although there was a tendency for females to list the care-based types of issue more consistently than males, and for males to list the justice-based types of issue more consistently than females, which supports Gilligan's (1982, 1988) position, these differences were small.

Insert Table 3 about here

Another way to assess within-person consistency is to determine the average number of dilemmas across which participants listed particular issues. For example, if a

participant listed issues about caring for others for three of the six dilemmas, issues about self-consequences for two of the six dilemmas, and issues about combatting immorality for one of the six dilemmas, the participant would receive a consistency score of 33.33% (an average of two out of six dilemmas). Overall consistency scores were calculated as well as separate consistency scores for the care-based and justice-based issues. The overall mean consistency score for males was 27.80 and for females was 27.70. The mean consistency score for care-based issues was 27.97 for males and 28.83 for females. The mean consistency score for justice-based issues was 28.43 for males and 26.37 for females.

Gender and Dilemma-Based Variations in Perception of Care versus Justice

A 2 X 6 (Gender X Dilemma) analysis of variance (ANOVA), with repeated measures on the last factor, was conducted on the Likert-scale ratings of the extent to which participants perceived care versus justice issues in each of the six types of dilemma. The results did not show a significant main effect for Gender, $F(1, 58) = 2.02$. There was a highly significant main effect for Dilemma, $F(5, 290) = 23.94$, $p < .00001$. Post hoc comparisons using Tukey's procedure revealed that (a) participants rated the two prosocial types of dilemma ($M_s = 4.7$ and 4.3 , $SD_s = 1.3$ and 1.5) as significantly more care based (less justice based) than the social pressure dilemmas ($M_s = 2.8$ and 3.1 , $SD_s = 1.9$ and 1.9) and the antisocial dilemmas ($M_s = 2.1$ and 2.8 , $SD_s = 1.6$ and 1.7) and (b) participants rated the social pressure, parent dilemma as significantly more care based (less justice based) than the antisocial, transgression dilemma. As expected, females tended to rate the social pressure, parent dilemma as more care based on the Likert scale ($M = 3.7$) than males rated it ($M = 2.6$), $t(58) = 2.18$ and 2.24 , $p < .03$.

Gender and Dilemma-Based Variations in Experience, Significance, Willingness to Discuss, and Perceived Morality

A series of 2 X 6 (Gender X Dilemma) ANOVAs, with repeated measures on the last factor, were conducted on the Likert-scale ratings of participants' perceptions of the types

of dilemma. These analyses revealed a significant interaction between Gender and Dilemma on the experience question, $F(5, 290) = 3.07, p < .01$, and perceived significance, $F(5, 290) = 5.40, p < .0001$. Analyses of simple effects revealed that males reported experiencing antisocial, transgression types of dilemma ($M = 4.9, SD = 1.5$) more often than females reported experiencing them ($M = 3.6, SD = 1.3$), and females deemed prosocial, helping types of dilemma as more significant ($M = 6.2, SD = 0.9$) than males deemed them ($M = 4.6, SD = 1.7$). There were no significant effects related to Gender on the question about willingness to discuss. However, there was a significant main effect for Gender, $F(1, 58) = 4.18, p < .05$, on the question about perceived morality: Females viewed the full set of dilemmas as involving more moral concerns ($M = 5.5, SD = 1.3$) than males viewed them as involving ($M = 5.1, SD = 1.6$).

Discussion

In Kohlberg's classification system, moral judgments are classified as upholding one of 12 moral "norms" in terms of one of 17 "elements" of morality, and are matched with the prototypic "criterion judgments" in Colby and Kohlberg's (1987) scoring manual (see p. 42). The criterion judgments were derived from the moral judgments made to the hypothetical dilemmas on Kohlberg's test by the participants in his longitudinal sample. The extent to which this system pertains to the issues people believe are involved in real-life moral dilemmas is unclear.

I found that many of the types of issue I identified in participants' responses to descriptions of real-life dilemmas resembled the elements identified by Kohlberg in his scoring system. For example, Kohlberg's retributing (exonerating) element (#3) resembles the combatting immorality category, and the upholding law category resembles Kohlberg's normative order elements and law norm. The general utilitarian considerations category resembles Kohlberg's good/bad group consequences element (#9). Self-respect is an element in Kohlberg's scheme (#11) and is a category in the present scheme. Finally, the positive reciprocity and procedural fairness categories

resemble two of Kohlberg's fairness elements, namely, reciprocity or positive desert (#15), and maintaining equity and procedural fairness (#16), respectively.

Some of Kohlberg's elements, however, fail to capture the exact nature of the types of issue the participants listed. For example, although Kohlberg identified balancing perspectives as a fairness element (#14), I distinguished between perspective-taking for the sake of making fair and objective decisions (within the procedural fairness category) and perspective-taking for the sake of listening to and understanding others. I deemed the latter as more empathy based or care based than the former.

Kohlberg also did not distinguish between self and other to the extent that I did in the present classification system. For example, although Kohlberg's good/bad individual consequences (element #8) resembles the caring for others category, and although Kohlberg's upholding character element (#10) resembles the positive social influence category, my categories are defined as specifically other-oriented. The egoistic consequence elements that Kohlberg identified--good/bad reputation (#6) and seeking reward/avoiding punishment (#7)--resemble two of my issue categories, namely, reputation and self-consequences, but my categories are specifically self-oriented. As another example, Kohlberg identified one element (#13) about serving human dignity and autonomy, whereas I identified three such categories (i.e., self-autonomy, others' respect for self, and self's respect for other). In addition, some of the elements identified in Kohlberg's system were not represented in the issues the participants listed. For example, no issues were considered to be concerned with serving social ideal or harmony (element #12), or maintaining social contract or freely agreeing (element #17).

Although the caring for others category resembles Kohlberg's good individual consequence element, some aspects of the category are not captured adequately by Kohlberg's system. The caring for others category is defined not only by concern for good consequences to others, but also by altruism, compassion and love, and responsibility toward others. Gilligan (1988) argued that Kohlberg's system does not

adequately deal with issues of care, and I have found this to be the case. Unlike Gilligan and her colleagues (Gilligan & Attanucci, 1988a; Lyons, 1983), I distinguished between issues concerned with relationships and issues concerned with caring for others, and I have found this to be a meaningful distinction. In fact, I found that people's issues about relationships were divided between those that were concerned with upholding relationships, trust and honesty in relationships, and the nature or quality of relationships. Neither Gilligan's definition of care nor Kohlberg's affiliation norm seemed to capture this distinction.

To summarize, many of the issues the participants listed as involved in (descriptions of) real-life dilemmas resembled the types of issue or moral elements identified by Kohlberg in responses to the hypothetical dilemmas on his test. Kohlberg's content classification system handles the justice-based issues raised in real-life dilemmas relatively well. However, the participants placed greater emphasis on self-oriented issues, issues about caring for others, and relationship-type issues than is evident in Kohlberg's scheme. Indeed, 37% of the issues classified were self-oriented in nature. In Wark and Krebs (1996), we scored the 708 criterion judgments in Colby and Kohlberg's (1987) scoring manual for moral orientation and found that only 12% of them were care based, compared to 44% of the issues I identified in the present study. It is important to note that I did not ask the participants to list the moral issues they saw to be involved in the dilemmas. Kohlberg's classification scheme is based on people's moral judgments. It is possible, for example, that the care-based issues that participants listed were not all considered moral in nature by the participants.

Contrary to Gilligan's (1977, 1982, 1986, 1988) position, females did not view the set of moral dilemmas as involving more care-based issues than males did, and males did not view the dilemmas as involving more justice-based issues than females did. These results differ from those of other studies reporting gender differences in real-life moral judgment (e.g., Gilligan & Attanucci, 1988a; Lyons, 1983). With dilemma held constant,

males and females did not differ in the types of issue they perceived to be involved. Consistent with the results reported in Wark and Krebs (1996), both males and females listed more care-based issues for the prosocial types of dilemma than for other types of dilemma, and both males and females listed more justice-based issues for the antisocial types of dilemma than for the prosocial types of dilemma.

On average, one third of the issues listed for each dilemma were classified in one predominant issue category, and half of the issues were classified in two categories, and these categories differed across the different types of dilemma. In every case, the type of issue listed most frequently was a direct reflection of the nature of the dilemma. For example, the prosocial, helping dilemma was described as involving a decision about whether to engage in proactive behavior, such as assuming responsibility to help another. Over a third of the issues participants listed as involved in the helping dilemma were classified as concerned with caring for others. In essence, participants restated the nature of the dilemma, or, in other words, participants viewed the dilemma in terms of the information it supplied. This is an important finding when we consider the gender differences in moral orientation reported by studies in which dilemma content was not held constant (e.g., Gilligan & Attanucci, 1988a; Lyons, 1983). If, for example, the female participants in these studies reported relationship-type conflicts more often than the male participants reported them, and if the male participants reported transgression-type conflicts more often than the female participants did, we would expect females to make judgments centered around relationship-type issues more often than males would, and for males to make judgments centered around combatting immorality more often than females would. The answer given is to a large extent determined by the question asked. An orientation to care or justice is as much in the moral dilemmas a person may face as it is within the person. Some dilemmas involve mainly care-based issues, other dilemmas involve mainly justice-based issues, and a substantial number of people see them in these terms.

The dilemmas differed, however, in the extent to which they were viewed in terms of the information or issues they supplied versus the issues participants projected or read into them. Some dilemmas were interpreted more strongly in terms of the intrinsic types of issue than other dilemmas. For example, 27% of the issues participants listed for the antisocial, transgression dilemma were considered intrinsic in the description of the dilemma (combatting immorality), whereas, 44% of the issues participants listed for the social pressure, parent dilemma were considered intrinsic in the nature of the dilemma (self-autonomy). These findings demonstrate that moral dilemmas may differ in the strength of the structure they impose on those who view them. Like objects in the physical world, some moral dilemmas are simple; others are complex. Some clearly assume one form; others, like ambiguous or reversible figures, may be viewed in different ways.

It should be noted, however, that no dilemma was viewed entirely in its own terms. Even the social pressure, parent dilemma, which was interpreted strongly as oriented around the issue of self-autonomy, was not viewed in entirely this way by all participants; it was viewed in other terms by a substantial number of participants. Some viewed it in terms of consequences to themselves. Others viewed it in terms of its implications for relationships. And others viewed it in terms of procedural fairness.

It seems important when investigating people's interpretations of moral dilemmas to attend to the determining power of the dilemmas. The data in Table 2 suggest that most moral dilemmas are viewed predominantly in their own terms, at least in terms of the major issues they raise, but that different people read different secondary issues into them. It is as though people say they see a tree when you show them a picture of a tree, but different people focus on different aspects of it when considering its significance.

It also seems the range of secondary issues people see in dilemmas is influenced by the structure of the dilemma. For example, when participants read a description of a conflict about whether or not to resist the temptation to gain something for the self

through immoral behavior such as dishonesty, the salience of self-consequences may have increased the salience of related secondary issues such as self-respect and fairness (reciprocity). The emphasis on combatting the immorality of a friend in the antisocial, transgression dilemma may have directed participants to think of issues centered around procedural fairness and positive social influence. The emphasis on self-autonomy in the social pressure dilemmas, particularly the peer dilemma, may have made other self-oriented issues salient. Lastly, the emphasis on caring for others in the prosocial, helping dilemma and the emphasis on relationship-type issues in the prosocial, loyalty dilemma, may have made other associated care-based issues more salient than for other dilemmas.

The results of the content analysis of the types of issue participants listed are supported by the participants' Likert-scale ratings of each dilemma. Participants listed more care-based types of issue for the prosocial dilemmas than for the other dilemmas, and participants rated the prosocial dilemmas as involving more care-based issues than the other dilemmas. Participants listed more justice-based types of issue for the antisocial and social pressure dilemmas than for the prosocial dilemmas, and participants rated the antisocial and social pressure dilemmas as involving more justice-based issues than the prosocial dilemmas.

Although there was no gender difference in the types of issue listed for the social pressure, parent dilemma, females tended to rate the dilemma as involving more care-based issues than males rated it as involving. Although the statistically nonsignificant interaction between gender and dilemma on the Likert ratings precludes any conclusions, this is an intriguing finding given the trend noted in Wark and Krebs (1996), namely, that females tended to make more care-based moral judgments than males made about the social pressure dilemmas.

If the ways people viewed moral dilemmas were determined only by the type of dilemma, we would expect everyone to view particular dilemmas as involving the same issues. Conversely, if the ways people viewed moral dilemmas were determined only by

internal, moral orientations, we would expect differences between people with different moral orientations, and we would expect within-person consistency across different dilemmas. The data do not support either of these scenarios. Rather, the data support a more interactional model: The ways people view moral problems is determined to a large extent by the type of dilemma, but different people also construct the same dilemmas in different ways.

The examination of within-person consistency revealed that approximately half of the participants--an equal number of males and females--consistently listed one type of issue across three or more dilemmas. There were individual differences in the tendency to construct dilemmas, but the differences were not gender-related (although there was a nonsignificant tendency for males to see more justice-based issues across dilemmas, and for females to see more care-based issues across dilemmas). Both justice-based (upholding justice and self-oriented) and care-based (other-oriented and relationship) types of issue were consistently listed across dilemmas.

The results of the analysis of within-person consistency in issues seen in different dilemmas suggest there are individual differences in the extent to which people interpret moral dilemmas in their own terms, as opposed to the terms of the dilemmas. Some participants viewed most of the dilemmas primarily in terms of the issues most salient in the dilemmas. Other participants viewed most of the dilemmas in terms of a particular issue that seemed to be salient to them--some participants in terms of the implications for themselves; other participants in terms of the implications for others; and still other participants in terms of the implications for relationships. It seems important in the study of moral judgment to determine the extent to which people accommodate to the stimuli they encounter and the extent to which they assimilate stimuli into issues that are important to or salient to them.

A related issue concerns the extent to which participants viewed all dilemmas in terms of one or two, versus many, issues. Some participants listed only one or two issues for

each dilemma. Other participants listed many. It is possible such differences may be related to characteristics such as cognitive complexity and flexibility.

Finally, I attempted to understand why females in the Wark and Krebs (1996) study reported more prosocial real-life dilemmas than males did, and why males reported more antisocial real-life dilemmas than females did. The data suggest that males reported more antisocial dilemmas for different reasons from why females reported more prosocial dilemmas. Males reported experiencing more antisocial (transgression) types of dilemma in real life than females reported experiencing. Thus, the notion that males make more justice-based moral judgments in real life may be correct, but it may not be because males construct moral problems in more justice-based ways than females. Rather, it may be because males experience more justice-based dilemmas than females experience. The data did not, however, support the converse possibility that females experience more prosocial types of dilemma than males experience. Rather, the findings suggest that females attribute more significance to prosocial (helping) types of dilemma than males do. The tendency for males and females to report different types of real-life dilemma, in part reflecting their experience and in part reflecting the significance they attach to different dilemmas, may be a more powerful source of gender differences in moral judgment than internal moral orientations.

To conclude, it would seem that Gilligan's (1982, 1988) model of moral judgment has underestimated the strength of the determining influence of different types of dilemma. The results suggest a more complex, interactional model than that offered by Gilligan. Dilemmas may influence the issues people see in them, but dilemmas may differ in their determining power. Some dilemmas may be weak, ambiguous, multidimensional, or complex. Others may be strong, clear, unidimensional, or simple. Individuals differ in their tendency to view dilemmas in terms of features of the dilemma, versus in terms of the issues that are personally salient or important to them. Some individuals are more "projective" or "constructive" than others. Individuals differ in both how many issues

they see in dilemmas and what issues they see in dilemmas. These differences, however, are not gender-related. They may stem from personal experience, or from internal cognitive phenomena like values or cognitive complexity. This remains to be discovered.

Study 2

Study 2 was designed as a follow-up to Study 1. In Study 1, participants were asked what issues were involved in different types of dilemma. The failure to find any overall gender differences in the perception of different issues, however, does not establish that the moral judgments males and females make about the same dilemmas will not differ in moral orientation. To examine this issue, a sample of males and females was asked to make moral judgments about descriptions of three specific dilemmas chosen from the real-life dilemmas provided by participants in the Wark and Krebs (1996) study. The first dilemma was chosen on the basis that it evoked primarily care-based moral judgments from the participant who reported it. It was classified in Wark and Krebs (1996) as prosocial (helping). The second dilemma was chosen on the basis that it evoked primarily justice-based moral judgments from the participant who reported it. It was classified as antisocial (transgression). The third dilemma was classified as a social (parent) pressure type of dilemma and it evoked an equal amount of care-based and justice-based moral judgments from the participant who reported it. In addition, given Gilligan's (1977, 1982) criticism of Kohlberg's theory of moral development and her contention that males and females respond differently to his test, participants were asked to make moral judgments about dilemmas on Kohlberg's test.

Gilligan (1982, 1988) asserted that people tend to construct moral problems in terms of their moral orientation. However, the results of the Wark and Krebs (1996) study have led to the expectation that people will invoke different forms of moral judgment when reasoning about different types of dilemma. In particular, I expected the prosocial dilemma to evoke primarily care-based moral judgments, the antisocial dilemma and the dilemmas on Kohlberg's test to evoke primarily justice-based moral judgments, and the

social pressure dilemma to evoke an equal proportion of care-based and justice-based moral judgments.

In Wark and Krebs (1996), females tended to make more care-based moral judgments than males to the social pressure dilemmas. Although this tendency was statistically nonsignificant, I investigated the possibility that females would make more care-based moral judgments than males to the social pressure dilemma.

As in Study 1, participants were asked to rate each dilemma on a number of dimensions. Participants were asked to what extent they viewed each dilemma as involving care-based versus justice-based issues. Based on the results of Study 1, I expected (a) males and females to rate the prosocial dilemmas as involving more care-based (less justice-based) issues than the antisocial and social pressure dilemmas and (b) males and females to rate the social pressure dilemma as involving more care-based (less justice-based) issues than the antisocial dilemma. Based on the trend reported in Study 1, I investigated the possibility that females would rate the social pressure dilemma as involving more care-based issues than males would rate it as involving. Participants were also asked about their experience with similar dilemmas, how significant and how moral they deemed each type of dilemma, and how willing they would be to discuss each type of dilemma. Based on the results of Study 1, males were expected to report experiencing dilemmas similar to the antisocial dilemma more often than females, and females were expected to deem the prosocial dilemma as more significant than males deemed it.

Method

Participants

The sample was composed of 30 undergraduate students (15 males and 15 females) who volunteered for the study to fulfill a psychology course requirement. The average age of males was 21.9 (range = 18-28) and of females was 20.1 (range = 18-24), $t(28) = 1.83$, *ns*. The mean SES ratings (based on parents' occupation and income) on a scale

ranging from one to five for males was 3.88 and for females was 3.77, $t(22) = 0.30$, *ns*. The mean grade point average of males was 2.94 and of females was 2.81, $t(24) = 0.89$, *ns*. All participants were unmarried.

Procedure

Participants were tested individually in one session. After giving consent, participants were asked to complete a package of questionnaires containing (a) a request for demographic information, (b) a short form of Kohlberg's test, (c) a prosocial dilemma, (d) a antisocial dilemma, and (e) a social pressure dilemma, in random order. The questionnaires took approximately 1 1/2-2 hours to complete.

Kohlberg's test. Kohlberg's test was administered in accordance with the instructions outlined by Colby and Kohlberg (1987). Kohlberg's test consists of a set of hypothetical dilemmas, such as whether a man named Heinz should steal an overpriced drug to save his dying wife, followed by a set of probing questions. As in other research (see Krebs, Vermeulen, Carpendale, & Denton, 1991; Wark & Krebs, 1996), participants in the present study were given the two most frequently employed dilemmas--Dilemmas III (Heinz dilemma) and III' (Officer Brown dilemma) from Form A of Kohlberg's test.

Real-life-like dilemmas. For the antisocial dilemma, participants were asked to imagine themselves dealing with a professor who, during lectures, tends to impose prejudiced points of view on the class and, in turn, expects students to endorse his or her views on exams. While opposition to the professor's views maintains self-respect, it also raises the possibility of receiving a poor grade. For the prosocial dilemma, participants were asked to imagine themselves in the position of deciding how to help a younger sister, who has admitted to being sexually abused by a cousin, without causing a disruption in the family. For the social pressure dilemma, participants were asked to imagine themselves deciding whether to obey house rules and fulfill their parents' expectations, or to assert their independence and move out.

For each real-life-like dilemma, participants were first asked to describe the nature of the dilemma and to state the points of view of all parties involved. On ensuing pages, participants were asked to respond to a set of probe questions such as, "What would you see to be the issues involved? What is at stake?", "What makes it a moral conflict?", "What options would you consider?", "What do you think you would think about if you were faced with this conflict; how would you feel?", "How would you resolve the conflict? Would that be the right decision? If not, what would be the right decision?".

As in Study 1, participants were asked to rate each dilemma on the following five dimensions set up in a 7-point Likert format: the extent to which they (a) viewed the dilemma as involving issues of care versus justice, (b) had experienced this type of dilemma in their own lives, (c) deemed it significant, (d) would be willing to discuss it on an anonymous questionnaire if they were experiencing it, and (e) viewed it as constituting a moral dilemma.

Scoring Moral Orientation

In Lyons' (1983) scoring system, moral considerations concerned with relationships, caring, and the promotion of the welfare of others or prevention of their harm are defined as care based, whereas considerations concerned with conflicting claims between self and other (including society) and the maintenance of impartial rules, principles and standards, particularly those of fairness and reciprocity, are defined as justice based. Lyons (1983) classified moral considerations as either care based or justice based, calculated the relative number of care-based and justice-based considerations made to a dilemma, then classified the moral orientation as either primarily care based or primarily justice based. Investigators such as Gilligan and Attanucci (1988a) and Brown and her colleagues (1987) added a third category (care/justice) in order to classify moral orientations considered to be both care and justice. I have found Lyons' (1983) scoring system to be the most reliable. Although Lyons' procedure has been used primarily to classify

judgments about real-life dilemmas (Gilligan, 1982; Gilligan & Attanucci, 1988a; Lyons, 1983), the procedure may also be used to score hypothetical dilemmas.

As in previous research (e.g., Krebs, Vermeulen, Denton, & Carpendale, 1994; Wark & Krebs, 1996), I employed a refined version of Lyons' system--a 5-point scale--to assess moral orientation. Judgments that were exclusively care based received a percentage score of 100, and judgments that were exclusively justice based received a percentage score of 0. Judgments that involved equal concerns about care and justice received a percentage score of 50. Judgments that were predominantly care based, but that harbored an element of justice, received a percentage score of 75; judgments that were predominantly justice based, with some care, received a percentage score of 25 (see Table 4 for examples of scored judgments). The average number of scorable judgments in the dilemmas was approximately 7. A mean percent care score was calculated for each participant on each dilemma by averaging all scored judgments.

Insert Table 4 about here

Scoring for moral orientation was conducted blind to all other information about the participant, and one quarter of the Kohlberg, antisocial, prosocial, and social pressure dilemmas was scored by a second scorer. There was 88% agreement within 12.5 percentage points (half way between adjacent categories) for the social pressure dilemma, and 100% agreement for the Kohlbergian, antisocial, and prosocial dilemmas, $r(6) = .80, .96, .22, \text{ and } .96$, respectively. (The relatively low interrater correlation for the antisocial dilemma was due to small variability between the raters' scores.)

A number of different procedures for determining consistency in moral orientation have been employed (e.g., Lyons, 1983; Pratt et al., 1988; Walker et al., 1987). As in previous research (Wark & Krebs, 1996), I assessed the consistency of moral orientation in terms of the same or an adjacent score on a 5-point scale (i.e., J, J(C), C/J, C(J), C).

Results

Consistency in Moral Orientation Across Dilemmas

Contrary to Gilligan's (1982, 1988) expectations, none of the participants obtained the same moral orientation score on all four dilemmas, and only two of the participants (males) obtained the same or an adjacent moral orientation score on all four dilemmas.

Gender and Dilemma-Based Variations in Moral Orientation

A 2 X 4 (Gender X Dilemma) ANOVA, with repeated measures on the last factor, was conducted on the moral orientation scores. This analysis revealed a significant main effect for Gender, $F(1, 28) = 6.98, p < .02$: Females made more care-based moral judgments than males. There was a highly significant main effect for Dilemma, $F(3, 84) = 123.62, p < .00001$. As expected, post hoc comparisons revealed that the prosocial dilemma evoked primarily care-based moral judgments, the social pressure dilemma evoked equal proportions of care-based and justice-based moral judgments, and the antisocial dilemma and the dilemmas on Kohlberg's test evoked primarily justice-based moral judgments (see Table 5).

Insert Table 5 about here

These effects, however, were qualified by a significant Gender X Dilemma interaction, $F(3, 84) = 2.74, p < .05$. Analyses of simple effects revealed that females made significantly more care-based moral judgments than males on only one type of dilemma--the social pressure dilemma (see Table 5).

Gender and Dilemma-Based Variations in Perception of Care versus Justice

A 2 X 4 (Gender X Dilemma) ANOVA, with repeated measures on the last factor, was conducted on the Likert-scale ratings of the extent to which participants perceived care versus justice issues in each dilemma. There was no main effect for Gender, $F(1, 28) = 0.23, ns$. However, there was a significant main effect for Dilemma, $F(3, 84) = 14.82, p$

< .00001, unqualified by any interactions. Post hoc comparisons revealed only one systematic difference: The antisocial dilemma was rated as involving more issues of justice than the other three dilemmas (see Table 5). Contrary to expectation, the prosocial dilemma was not rated as more care based than the social pressure dilemma.

Gender and Dilemma-Based Variations in Experience, Significance, Willingness to Discuss, and Perceived Morality

A series of 2 X 4 (Gender X Dilemma) ANOVAs, with repeated measures on the last factor, were conducted on the Likert-scale ratings of the extent to which participants have experienced similar dilemmas, viewed them as significant, would be willing to discuss them, and perceived them as moral. Consistent with the results reported in Study 1, males reported experiencing dilemmas similar to the antisocial dilemma more often than females ($M_s = 4.00$ vs. 3.20), and females rated the prosocial dilemma as more significant than males rated it ($M_s = 5.87$ vs. 5.46); but, these differences were not statistically significant. It is plausible that the absence of a statistically significant difference is due to the small number of participants and the specificity of the dilemmas: the antisocial dilemma involved school-related issues--issues with which both the male and female undergraduate participants would be expected to be familiar--and the prosocial dilemma involved a particularly significant issue, namely, sexual abuse. The Stouffer method of aggregating probabilities⁶ across studies, however, produced highly significant gender differences: (a) males reported experiencing dilemmas similar to the antisocial (transgression) types of dilemma more often than females, $Z = 2.46$, $p < .007$, and (b) females deemed the prosocial (helping) types of dilemma more significant than males deemed them, $Z = 2.37$, $p < .009$.

Discussion

Contrary to Gilligan's (1977, 1982, 1986, 1988) position, I found no evidence that females possess an unqualified tendency to think about moral dilemmas from a more care-based perspective than males. Consistent with the results reported in Wark and

Krebs (1996), participants made primarily care-based moral judgments about the prosocial dilemma, an equal proportion of care-based and justice-based moral judgments about the social pressure dilemma, and primarily justice-based moral judgments about the antisocial dilemma and the dilemmas on Kohlberg's test. Moral orientation was primarily a function of the type of dilemma about which people reasoned. Only two participants displayed within-person consistency in moral orientation across the different dilemmas.

Consistent with the results reported in Study 1, there were no gender differences found in the care-justice Likert ratings of the dilemmas. Also consistent with Study 1, participants rated the antisocial dilemma as involving primarily justice-based issues. This finding supports the results reported above, namely, that participants made primarily justice-based moral judgments in response to the antisocial dilemma.

Contrary to expectation, the prosocial dilemma was not rated as involving primarily care-based issues, nor was it rated as involving more care-based issues than the social pressure dilemma. Interestingly, the prosocial dilemma was also not rated as involving more care-based issues than the dilemmas on Kohlberg's test. In fact, the dilemmas on Kohlberg's test were rated as involving slightly more care-based issues than the prosocial dilemma. However, participants made primarily justice-based moral judgments in response to the dilemmas on Kohlberg's test. It is possible that the care-based issues perceived in the dilemmas on Kohlberg's test were not considered moral issues, and when participants made judgments about the dilemmas, they oriented their judgments around the moral issues (i.e., the justice-based issues). This explanation, however, is limited in so far as it does not explain why participants made primarily care-based moral judgments about the prosocial dilemma.

The results suggest that although people may identify or perceive different types of issue as involved in a dilemma, they may tend to orient to certain types of issue when they make moral judgments about the dilemma. It is possible that participants oriented their judgments around the issues inherent or most salient in the dilemmas. In Study 1,

participants listed care-based and justice-based issues for all dilemmas, but they focussed on those that were most salient in the dilemmas. When the participants in Study 2 made judgments about the prosocial dilemma, they may have oriented their judgments primarily around the most salient issue (caring for others), and when participants made judgments about the social pressure dilemma, they may have oriented their judgments around the justice-based issue of self-autonomy. I did not ask participants to list the types of issue they saw to be involved in the dilemmas on Kohlberg's test. Although the Kohlbergian dilemmas were rated on a Likert scale as involving both care-based and justice-based issues, participants made primarily justice-based moral judgments. It is possible that justice-based issues would be listed as those that are most salient in the Kohlbergian dilemmas.

According to Gilligan (1982), real-life moral dilemmas are like ambiguous figures in that they are weak in terms of their care-justice structure. That is, they may be perceived either in terms of care or in terms of justice. In support of Gilligan's contention about gender differences in moral orientation, and in support of the trends noted in Wark and Krebs (1996) and in Study 1, females made more care-based moral judgments in response to the social pressure dilemma than males did. However, females did not rate the social pressure dilemma as involving more care-based issues than males rated it.

Interestingly, the social pressure, parent type of dilemma resembles the real-life experiences Gilligan (1982) claimed are the source of gender differences in moral orientation. According to Gilligan (1982), the early processes of identity formation and socialization in boys involve identification with their fathers and, thus, separation and autonomy from their mothers. Whereas girls identify with their mothers and, thus, maintain the connection and attachment to their mothers. Justice-based issues such as autonomy, then, become more salient for boys, and care-based issues such as relationship maintenance become more salient for girls (Gilligan, 1982). It is also possible that the gender difference may have stemmed from a gender difference in the nature of the

relationships the university-aged participants have with their parents or guardians:

Females may feel closer or value the relationship they have with their parents to a greater degree than males do, and, consequently, females may orient to relationship issues when reasoning about social pressure, parent types of dilemma more than males do. Further investigation of this issue is encouraged.

The results of Study 1 and Study 2 suggest that the tendency for males and females to report different types of real-life moral dilemma may be related to gender differences in experience with and significance attached to different types of moral dilemma. This is a novel and interesting finding. For instance, it is plausible that because males report having more experience than females with antisocial dilemmas involving transgressions--dilemmas which pull for justice-based moral judgments--males may be more familiar with justice-based ways of thinking and may be more inclined to view moral problems from a justice-oriented perspective, which is consistent with Gilligan's position. However, the results suggest that males do not, in general, interpret moral problems in more justice-based ways than females do. Similarly, it is possible that females are socialized to attach more significance to dilemmas about relationships and caring for others than males are, but the present results suggest that females are not, in general, inclined to interpret all moral problems in more care-based ways than males do.

To summarize, the results of this research suggest a systematic tendency for people to make more care-based moral judgments about prosocial types of dilemma than about antisocial types of dilemma, social pressure types of dilemma, and the dilemmas on Kohlberg's test, even though they do not rate the prosocial dilemmas as involving more care-based issues than the social pressure and Kohlbergian dilemmas. I have found this pattern of moral judgment to hold regardless of whether people actually experienced the dilemmas, as in Wark and Krebs (1996), or responded to descriptions of dilemmas, as in Study 2. Study 1 and Study 2 also revealed a systematic tendency for people to (a) rate the prosocial dilemmas as involving more care-based issues than they rated the antisocial

dilemmas as involving, and (b) rate the antisocial dilemmas as involving more justice-based issues than they rated all other types of dilemma as involving.

It is important to note that the participants in Study 1 and Study 2 did not make moral judgments about dilemmas they had actually experienced in everyday life. And although the participants in the Wark and Krebs (1996) study reported dilemmas they had actually experienced, different participants made judgments about different specific dilemmas (classified for type by the experimenters). To establish whether males and females view the same types of real-life moral dilemma in the same or in different ways, an investigation of the forms of moral judgment invoked across a broad range of different types of real-life dilemma is required. The need for a larger sample of real-life moral reasoning has been acknowledged by others (e.g., Pratt et al., 1988), but, to date, no study has undertaken this task. Particularly challenging is the need to hold dilemma content constant. This may be approximated by asking people to report and to make moral judgments about a number of particular types of real-life moral dilemma they have experienced. I employed this type of design in the third study of this research, which enabled me to test for gender-based and dilemma-based variations in moral maturity and moral orientation within and between the particular types of real-life moral dilemma.

Study 3

The first goal of this study was to examine the structural consistency of participants' moral judgment across several types of moral dilemma they had experienced in their everyday lives. Past research has found that people tend to display (a) similar forms of moral judgment across similar types of dilemma (e.g., across the nine dilemmas on Kohlberg's test, and across Kohlberg's dilemmas and philosophical types of impersonal dilemma), but (b) dissimilar forms of moral judgment across dissimilar types of dilemma (e.g., across hypothetical dilemmas and real-life dilemmas). To date, no study has investigated the forms of moral judgment people invoke across similar and dissimilar types of real-life moral dilemma. In the present study, I investigated the structural

consistency of people's moral judgment across hypothetical and real-life moral dilemmas and across different types of real-life dilemma. In particular, I asked males and females to make moral judgments about dilemmas on Kohlberg's test and to report and to make judgments about two antisocial types of real-life dilemma and two prosocial types of real-life dilemma⁷.

If Kohlberg's (1984) assumption of structural consistency in moral judgment is correct, people should base their moral judgments on the same or an adjacent stage across all types of moral dilemma. However, the results of the Wark and Krebs (1996) study have led to the expectation that people will invoke different forms of moral judgment when reasoning about different types of dilemma. In particular, I expected the dilemmas on Kohlberg's test to evoke predominantly Stage 3/4 moral judgments, prosocial types of real-life dilemma to evoke predominantly Stage 3 moral judgments, and antisocial types of real-life dilemma to evoke predominantly Stage 2 and Stage 2/3 moral judgments.

A second goal was to further examine Gilligan's (1982, 1988) assertions about gender-related differences in moral judgment. In particular, I set out to determine whether males and females invoke the same or different forms of moral judgment to the same types of real-life dilemma. Following Gilligan's (1982; see also Lyons' 1983) theorizing about self-concept, I also investigated the relation between gender, self-concept, and moral judgment. A number of studies have tested Gilligan's idea that the moral judgments of persons with a feminine gender-role identity (especially females) are primarily care based and scored at Stage 3 in Kohlberg's system, and that the moral judgments of persons with a masculine gender-role identity (especially males) are primarily justice based and scored at Stage 4 (or higher). However, the results of past research on the relation between gender-role, as assessed by a variety of instruments, and moral judgment have been mixed (e.g., Ford & Lowery, 1986; Leahey & Eiter, 1980; Lifton, 1985; Lyons, 1983; Pratt & Royer, 1982; Pratt, Golding, & Hunter, 1984; Pratt et al., 1988). In an earlier study (Wark & Krebs, 1996), we failed to find a relation between gender-role, as

measured by the Personal Attributes Questionnaire (Spence & Helmreich, 1978), and moral judgment.

Gilligan (1982) not only theorized that females and feminine people value the care aspect of morality more than males and masculine people value it, she also argued that self-concept is particularly bound to a sense of morality in women. Thus, for the purposes of the present study, I decided to ask participants to indicate how important being a caring and/or a just person is to their identity or self-concept, and I used this measure, instead of a standard gender-role measure, to investigate possible self-concept-related differences in moral judgment.

If Gilligan's (1982, 1988) assumption of consistency in moral orientation is correct, people should consistently make either primarily justice-based moral judgments or primarily care-based moral judgments across different types of moral dilemma. The results of Study 2 and the results of Wark and Krebs (1996), however, have led to the expectation that prosocial dilemmas will evoke more care-based moral judgments than antisocial dilemmas and the dilemmas on Kohlberg's test. Although the dilemmas on Kohlberg's test and the antisocial dilemmas constitute different types of dilemma, I expected both types to evoke primarily justice-based moral judgments. Thus, I expected relatively high levels of consistency in moral orientation across the dilemmas on Kohlberg's test and the antisocial dilemmas.

If Gilligan's (1982) theorizing about gender differences is correct, females should make more care-based, Stage 3 moral judgments than males make, and males should make more justice-based, Stage 4 judgments than females make. Following Gilligan's (1982; also Lyons, 1983) theorizing about self-concept, females should regard being a caring person as more important to their identity or self-concept than males should, and males should regard being a just person as more important to their identity or self-concept than females should. To the extent that moral judgment is tied to self-concept in the way that Gilligan and her colleagues (Gilligan, 1982; Lyons, 1983) theorize, identity ratings of care should be negatively related to moral maturity and positively related to the

proportion of care-based moral judgments made, whereas identity ratings of justice should be positively related to moral maturity and negatively related to the proportion of care-based moral judgments made.

As in the first two studies, participants were asked to rate each dilemma on a number of dimensions. Participants were asked to rate each dilemma on the extent to which they viewed it as involving issues of care and issues of justice. Given the results of Study 1 and Study 2, I expected that the prosocial dilemmas will be rated as involving more care-based issues and less justice-based issues than the antisocial dilemmas will be rated as involving. Given the results of Study 2, I expected that the dilemmas on Kohlberg's test will be rated as involving more care-based issues than the antisocial dilemmas will be rated as involving.

Participants were also asked about their experience with similar dilemmas, how significant and how moral they deemed each type of dilemma, how easy it was to discuss each type of dilemma on a questionnaire, and how difficult they found each moral decision to be. Based on the results of Study 1 and Study 2, males were expected to report experiencing dilemmas similar to the antisocial, transgression dilemma more often than females, and females were expected to deem the prosocial, helping dilemma as more significant than males deem it.

Method

Participants

The sample was composed of 40 male and 40 female undergraduate students who volunteered for the study to fulfill a psychology course requirement⁸. The mean age of the male participants (23.3 years; range = 17-50) did not differ significantly from the mean age of the female participants (23.3 years; range = 17-40), $t(78) = 0.00$. The mean SES ratings for males (2.73) did not differ significantly from the mean SES ratings for females (2.72), $t(77) = 0.00$. The mean grade point average (GPA) of males (3.03) did not differ significantly from the mean GPA of females (2.97), $t(72) = 0.51$. Of the 77 participants who reported ethnicity, 41 were Caucasian, 22 were Asian, two were Middle

Eastern, one was Hispanic, one was Spanish, one was Portuguese, and one was Metis/Venezuelan (eight of the participants stated they were Canadian). Five participants (three males and two females) were married, two were engaged (one male and one female), and two were divorced (one male and one female).

Procedure

Participants were tested individually in one session. After giving informed consent, participants were asked to complete a package of questionnaires containing (a) a request for demographic information, (b) a self-concept questionnaire, (c) a request to supply a list of moral dilemmas encountered in the past few years, (d) a short form of Kohlberg's test, and (e) instructions for four real-life moral dilemmas (two antisocial and two prosocial). The dilemmas were given in random order, but the dilemmas on Kohlberg's test were always presented together. The questionnaires typically took 2-2 1/2 hours to complete.

Self-concept. Participants were provided with a definition of a caring person--a person who is concerned with and values relationships, caring, and the promotion of the welfare of others or prevention of their harm--and a definition of a just person--a person who is concerned with individual rights, autonomy, reciprocity, and fairness, with reference to moral principles. Participants were then asked to respond to three questions, set up in a 7-point Likert format, which asked (a) how important to their identification, or self-concept, it is to be a caring person, (b) how important to their identification, or self-concept, it is to be a just person, and (c) if they had to choose between being a person who values care more than justice and being a person who values justice more than care, which would they choose? (For the first two questions, zero represented not important and seven represented very important. For the last question, zero represented care and seven represented justice.)

Kohlberg's test. Kohlberg's test was administered in accordance with the instructions outlined by Colby and Kohlberg (1987). Kohlberg's test consists of a set of hypothetical dilemmas, such as whether a man named Heinz should steal an overpriced drug to save his dying wife, followed by a set of probing questions. According to Colby and Kohlberg (1987), the structure of people's moral judgments to the dilemmas on Kohlberg's test is

highly consistent. It is therefore possible to assess stage of moral development with a small number of moral dilemmas (Kohlberg's "short form"). As in other research (see Krebs, Vermeulen, Carpendale, & Denton, 1991, for a review and justification; Wark & Krebs, 1996), participants in the present study were given the two most frequently employed dilemmas--Dilemmas III (Heinz dilemma) and III' (Officer Brown dilemma) from Form A of Kohlberg's test.

Real-life dilemmas. Employing the distinctions made post hoc by Wark and Krebs (1996), I asked participants to recall and to describe two antisocial types of dilemma and two prosocial types of dilemma they had experienced. For the antisocial, "transgression" dilemma, participants were asked to describe an experience in which they had to make a decision about how to react to a transgression (for example, involving violations of rules, laws, or fairness) committed by someone important to them. For the antisocial, "temptation" dilemma, participants were asked to describe an experience in which they were faced with the temptation to meet their own needs or desires, acquire resources, or advance their own gain by, for example, violating rules or laws, behaving dishonestly, immorally, or unfairly. For the prosocial, "loyalty" dilemma, participants were asked to describe an experience in which they were faced with two or more people making inconsistent demands on them, with implications for their relationship with each person. And for the prosocial, "helping" dilemma, participants were asked to describe an experience in which they had to make a decision about whether or not to take responsibility for helping someone important to them. In order to insure that participants responded appropriately to the real-life dilemmas--that is, that they discussed the types of dilemma requested--a research assistant who was blind to the purpose of the study and blind to information about the participants and to the identity of each real-life dilemma, read all real-life dilemmas and classified them into the four categories. All 80 participants employed in this study supplied complete and appropriate data.

For each real-life dilemma, participants were first asked to describe the dilemma and to state the points of view of all parties involved. On ensuing pages, participants were asked to respond to a set of probe questions such as, "What did you see to be the issues involved at the time?", "What made it a moral conflict?", "What options did you

consider?", "Do you think you did the right thing? Why or why not?", and "Is there another way to see the problem?". Approximately a quarter of a page was supplied for participants to answer each probe question.

As in Study 1 and in Study 2, participants were asked to rate each dilemma on the following dimensions set up in a 7-point Likert format: the extent to which they (a) had experienced similar dilemmas in real life, (b) deemed each type of dilemma significant, (c) found it easy to discuss this type of dilemma on this questionnaire, and (d) deemed each type of dilemma as moral. Participants in this study also were asked to rate each dilemma on the extent to which they (a) found the decision in each dilemma difficult, (b) viewed each type of dilemma as involving issues of care, and (c) viewed each type of dilemma as involving issues of justice.

Scoring

Moral maturity. The dilemmas on Kohlberg's test were scored by according to the procedure outlined by Colby and Kohlberg (1987). Kohlberg's scoring system produces two equivalent scores: (a) global stage scores that range from Stage 1 to Stage 5, including transitional stages such as 1/2, 2/3, etc., and (b) moral maturity scores (called weighted average scores by Colby and Kohlberg, 1987) ranging from 100 to 500 (see Colby & Kohlberg, 1987, pp. 158-188). One-quarter of the dilemmas were scored blindly by a second expert scorer. Interrater reliability was 100% agreement, $r(18) = .97$, within 25 weighted average points (one-quarter stage).

Real-life dilemmas involve moral issues similar to the moral issues in the dilemmas on Kohlberg's test. For example, real-life dilemmas involve decisions about how to react to (a) the transgression of another (Officer Brown), (b) the temptation to engage in antisocial behavior (Heinz), (c) issues of loyalty (Judy), and (d) prosocial responsibility (Dr. Jefferson). Thus, as in other research (Wark & Krebs, 1996), it was possible to calculate stage scores and moral maturity scores for the real-life dilemmas by matching them structurally with the criterion judgments in Colby and Kohlberg's (1987) manual (see also Krebs, Vermeulen, Carpendale, & Denton, 1991, for a review and justification). Table 6 contains examples of real-life moral judgments with their corresponding criterion

judgment matches. Interrater reliability on one-quarter of the transgression, temptation, loyalty, and helping real-life dilemmas, scored blindly, was 90%, 100%, 100%, and 100% agreement, respectively, within 25 weighted average points (one-quarter stage), $r_s(18) = .86, .99, .97, \text{ and } .92$.

Insert Table 6 about here

As in other research, the criterion for structural consistency in moral judgment I employed was the same or an adjacent stage on a 9-point scale (e.g., Stage 1, Stage 1/2, Stage 2, etc.) (see Krebs, Denton, Vermeulen, Carpendale, & Bush, 1991, for justification).

Moral orientation. As in Study 2, I employed a refined version of Lyons' (1983) scoring system--a 5-point scale--to assess moral orientation. Table 6 contains examples of real-life moral judgments with their corresponding moral orientation scores. The average number of scorable judgments in the real-life dilemmas was approximately 7. A mean percent care score was calculated for each participant on each dilemma by averaging all scored judgments.

Like scoring for stage of moral judgment, scoring for moral orientation was conducted blind to all other information about the participant, and one quarter of each type of dilemma were scored by a second scorer. Interrater reliability was 90%, 95%, 90%, 100%, and 90% agreement, for the Kohlberg, transgression, temptation, loyalty, and helping dilemmas, within 12.5 percentage points (half way between adjacent categories), $r_s(18) = .88, .92, .94, .98, \text{ and } .93$, respectively.

As in Study 2, the consistency of moral orientation was assessed in terms of the same or an adjacent score on a 5-point scale (i.e., J, J(C), C/J, C(J), C).

Results

Self-Concept

A series of t-tests was conducted on the Likert-scale ratings of the extent to which participants (a) viewed being a caring person as important to their identity or self-concept, (b) viewed being a just person as important to their identity or self-concept, and

(c) would choose between being a person who values care more than justice and being a person who values justice more than care. These analyses revealed that there were no significant differences between males and females in (a) their identity ratings of care ($M_s = 5.93$ and 6.23 , respectively), $t(78) = 1.38$, *ns*, (b) their identity ratings of justice ($M_s = 5.73$ and 5.50 , respectively), $t(78) = 0.85$, *ns*, and (c) the extent to which they would choose being a person who values care versus being a person who values justice, ($M_s = 3.03$ and 2.78 , respectively), $t(78) = 0.57$, *ns*. Matched t-tests revealed that (a) females rated being a caring person higher in importance than being a just person, $t(39) = 4.13$, $p < .0002$, and (b) males did not rate being a caring person significantly higher in importance than being a just person, $t(39) = 1.16$, *ns*. The correlations between self-concept ratings of care and self-concept ratings of justice for males and females did not differ ($r_s(38) = .49$ and $.49$, $p_s < .01$, respectively).

Relations between Moral Judgment and Self-Concept

There was one weak but systematic relation across dilemmas between moral judgment and self-concept: moral maturity was significantly positively related to identity ratings of justice, $r_s(78) = .21$ to $.28$, $p_s < .05^9$. (The relation was positive, but not significant on Dilemma III', $r(78) = .15$.) However, when examined by gender, the positive relation between identity ratings of justice and moral maturity was significant for males only, $r_s(38) = .32$ to $.39$, $p_s < .05$. Further, the relation for males was positive, but not significant on the antisocial, temptation dilemma, $r(38) = .27$, and on the prosocial, helping dilemma, $r(38) = .28$.

Variations in Moral Judgment Across Dilemmas

Consistency in Moral Stage Across Dilemmas

Contrary to Kohlberg's (1984) expectation, the majority of participants did not base their moral judgments on one stage-structure. None of the participants obtained the same global stage score (on a 9-point scale) on all six dilemmas, and only 23% of the participants (18% of the males, 28% of the females) obtained scores at the same or an adjacent substage. When the stage-score of each specific judgment was considered, I found that no participant based all his or her judgments on the same stage across all six

dilemmas, and only one participant based all his judgments on the same or an adjacent substage across all six dilemmas; 99% of the participants made judgments that spanned between three and six substages.

To compare the level of consistency of moral stage across similar versus dissimilar dilemmas, the percentages of participants who obtained the same or an adjacent substage across all pairs of dilemmas were calculated (see Table 7). Across the two dilemmas on Kohlberg's test, 44% of the participants (38% of the males, 50% of the females) obtained the same global stage score, with 95% of the participants (95% of the males, 95% of the females) obtaining scores at the same or an adjacent substage. Across the two antisocial dilemmas, 49% of the participants (48% of the males, 50% of the females) obtained the same global stage score, with 85% of the participants (85% of the males, 85% of the females) obtaining scores at the same or an adjacent substage. Across the two prosocial dilemmas, 51% of the participants (40% of the males, 63% of the females) obtained the same global stage score, with 89% of the participants (83% of the males, 95% of the females) obtaining scores at the same or an adjacent substage. (The latter percentage of females meeting the consistency criteria was greater than the percentage of males, $z = 2.08$, $p < .05$.) As the pattern of percentages displayed in Table 7 suggests, the structural consistency of people's moral judgments was higher across similar types of dilemma (e.g., across the two dilemmas on Kohlberg's test) than across dissimilar types of dilemma (e.g., across the two dilemmas on Kohlberg's test and the antisocial dilemmas).

Insert Table 7 about here

Gender and Dilemma-Based Variations in Moral Maturity

A 2 X 6 (Gender X Dilemma) ANOVA on moral maturity scores, with repeated measures on the last factor, revealed a highly significant main effect for Dilemma, $F(5, 390) = 96.69$, $p < .00001$, unqualified by any interactions¹⁰. The main effect for Gender was not significant, $F(1, 78) = 0.22$. There were no statistically significant differences

between the two dilemmas on Kohlberg's test, nor were there any differences between the two prosocial dilemmas, or between the two antisocial dilemmas. As suggested by Table 8 and Table 9, the dilemmas on Kohlberg's test tended to evoke Stage 3 and 3/4 moral judgments, the two prosocial dilemmas evoked Stage 3 moral judgments, and the two antisocial dilemmas evoked Stage 2 and Stage 2/3 moral judgments.

Insert Tables 8 and 9 about here

In all, 99%, 94%, 83%, and 81% of the participants scored higher on Kohlberg's dilemmas than on the antisocial, temptation real-life dilemmas, the antisocial, transgression real-life dilemmas, the prosocial, loyalty real-life dilemmas, and the prosocial, helping real-life dilemmas, respectively.

Consistency in Moral Orientation Across Dilemmas

Contrary to Gilligan's (1982, 1988) expectations, none of the participants obtained the same moral orientation score on all six dilemmas, and only two of the participants (females) obtained the same or an adjacent moral orientation score on all six dilemmas.

The percentages of participants who obtained the same or an adjacent moral orientation score across all pairs of dilemmas were calculated (see Table 10). Across the two dilemmas on Kohlberg's test, 43% of the participants (38% of the males, 48% of the females) obtained the same moral orientation score, with 90% of the participants (83% of the males, 98% of the females) obtaining the same or an adjacent moral orientation score. (The latter percentage of females meeting the consistency criteria was greater than the percentage of males, $z = 2.49$, $p < .05$.) Across the two antisocial dilemmas, 38% of the participants (33% of the males, 43% of the females) obtained the same moral orientation score, with 75% of the participants (73% of the males, 78% of the females) obtaining the same or an adjacent moral orientation score. Across the two prosocial dilemmas, 28% of the participants (20% of the males, 35% of the females) obtained the same moral orientation score, with 68% of the participants (63% of the males, 73% of the females) obtaining the same or an adjacent moral orientation score.

Insert Table 10 about here

Gender and Dilemma-Based Variations in Moral Orientation

A 2 X 6 (Gender X Dilemma) ANOVA on moral orientation scores, with repeated measures on the last factor, produced a significant main effect for Dilemma, $F(5, 390) = 81.04$, $p < .00001$, unqualified by any interactions. The main effect for Gender was not significant, $F(1, 78) = 0.50$. As shown in Table 8, the prosocial dilemmas evoked more care-based moral judgments than the other dilemmas. The dilemmas on Kohlberg's test and the antisocial dilemmas evoked more justice-based moral judgments than the prosocial dilemmas. Although there were no statistical differences between the two dilemmas on Kohlberg's test or between the two prosocial dilemmas, the antisocial, temptation dilemma evoked more justice-based (less care-based) moral judgments than the antisocial, transgression dilemma.

To summarize, there was considerable inconsistency in moral stage and moral orientation across dilemmas. Less than one-quarter of the participants scored at the same or an adjacent moral stage across all six dilemmas, and only two of the participants obtained the same or an adjacent moral orientation score across all six dilemmas. Although it is easier to obtain consistency by chance on a 5-point scale than on a 9-point scale, there tended to be less consistency in moral orientation than in moral maturity across dilemmas. The level of consistency improved across similar types of dilemma, namely, across the two dilemmas on Kohlberg's test, across the two antisocial dilemmas, and across the two prosocial dilemmas. The proportion of females who scored at the same or an adjacent moral stage across the two prosocial dilemmas was significantly higher statistically than the proportion of males, and the proportion of females who were consistent in moral orientation across the two dilemmas on Kohlberg's test was significantly higher statistically than the proportion of males. However, in terms of numbers, these differences were small (38 females vs. 33 males in the former case, and 39 females vs. 33 males in the latter case). Both males and females made lower-stage moral judgments in response to the real-life dilemmas, particularly the antisocial

dilemmas, than in response to the hypothetical dilemmas on Kohlberg's test. And both males and females made more care-based moral judgments about the prosocial real-life dilemmas than about the dilemmas on Kohlberg's test and the antisocial dilemmas.

Gender and Dilemma-Based Variations in Perception of Care and Justice

A 2 X 5 (Gender X Dilemma) ANOVA, with repeated measures on the last factor, conducted on the Likert-scale ratings of the extent to which participants perceived care issues in each dilemma failed to reveal a significant main effect for Gender, $F(1, 76) = 0.12$, but revealed a highly significant main effect for Dilemma, $F(4, 304) = 46.97$, $p < .00001$. The mean care ratings are displayed in Table 11. Consistent with the results of Study 1 and Study 2, post hoc comparisons revealed that the two prosocial dilemmas were rated as involving more issues of care than the two antisocial dilemmas. The antisocial, transgression dilemma was rated as involving more issues of care than the antisocial, temptation dilemma. Consistent with the results of Study 2, the dilemmas on Kohlberg's test were rated as involving more issues of care than the antisocial dilemmas, and there was no significant difference between the dilemmas on Kohlberg's test and the prosocial, helping dilemma in ratings of care (see Table 11).

Insert Table 11 about here

A 2 X 5 (Gender X Dilemma) ANOVA, with repeated measures on the last factor, conducted on the Likert-scale ratings of the extent to which participants perceived justice issues in each dilemma revealed significant main effects for Gender, $F(1, 76) = 5.36$, $p < .023$, and for Dilemma, $F(4, 304) = 26.55$, $p < .00001$. The mean justice ratings are displayed in Table 11. Contrary to Gilligan's (1982, 1988) expectation, females rated all dilemmas as involving more justice issues ($M = 4.37$) than males rated the dilemmas as involving ($M = 3.78$). Post hoc comparisons revealed that the two prosocial dilemmas

were rated as involving fewer issues of justice than the two antisocial dilemmas and the dilemmas on Kohlberg's test (see Table 11).

A series of matched t-tests were conducted to test the differences between the care and justice Likert-scale ratings for each dilemma. As would be expected, the care ratings were significantly higher than the justice ratings for the prosocial dilemmas, $t_s(78) = 9.53$ and 7.97 , $p_s < .00001$, and the justice ratings were significantly higher than the care ratings for the antisocial, temptation dilemmas, $t_s(78) = 6.19$, $p_s < .00001$. However, the difference between care ratings and justice ratings for the antisocial, transgression dilemma was not significant, $t(78) = 1.19$. For the dilemmas on Kohlberg's test, the care ratings were significantly higher than the justice ratings, $t(78) = 5.09$, $p < .00001$.

Gender and Dilemma-Based Variations in Experience, Significance, Ease of Discussion, Difficulty of Decision, and Perceived Morality

A series of 2 X 5 (Gender X Dilemma) ANOVAs, with repeated measures on the last factor, was conducted on the Likert-scale ratings of the extent to which participants had experienced similar types of dilemma, viewed them as significant, found it easy to discuss them, found each decision difficult, and perceived them as moral. The mean ratings are displayed in Table 12.

Insert Table 12 about here

These analyses revealed only one significant main effect for Gender: Females rated the decisions involved in all of the dilemmas as more difficult than males rated them ($M_s = 4.23$ and 3.60 , respectively), $F(1, 78) = 8.74$, $p < .004$. There were significant main effects for Dilemma on the experience question, $F(4, 312) = 24.39$, $p < .00001$, on the significance question, $F(4, 312) = 3.03$, $p < .018$, on the difficulty question, $F(4, 312) = 8.34$, $p < .00001$, and on the moral question, $F(4, 312) = 10.67$, $p < .00001$. Differences between the means are indicated in Table 12.

As expected, females rated the prosocial, helping dilemma as more significant than males rated it ($M_s = 4.91$ vs. 4.25). However, as in Study 2, this difference was not statistically significant. The Stouffer method of aggregating probabilities across studies, however, produced a highly significant gender difference: Females deemed the prosocial, helping dilemmas as more significant than males deemed them, $Z = 2.62$, $p < .004$. Contrary to expectation, females reported experiencing dilemmas similar to the antisocial, transgression dilemma more often than males ($M_s = 3.40$ vs. 2.93); but this difference was not statistically significant¹¹.

Discussion

The results of this study did not support Kohlberg's (1984) assumption that the structure of moral judgment is consistent within people across moral dilemmas. Consistent with the position advanced by Krebs, Denton, Vermeulen, Carpendale, and Bush (1991), stage of moral judgment varied across types of dilemma. None of the participants scored at the same global stage (on a 9-point scale), and less than one-quarter of the participants scored at the same or an adjacent substage across the six moral dilemmas. When all the specific moral judgments made by each participant in this study about all six moral dilemmas were considered, virtually all of the participants (99%) made moral judgments that spanned between three and six substages.

Colby and Kohlberg (1987) demonstrated that the structure (moral stage) of people's moral judgments about the dilemmas on Kohlberg's test is highly consistent. In support of this, 95% of the participants obtained scores at the same or an adjacent substage across the two dilemmas. Other research has shown that people's moral judgments also tend to be structurally consistent across Kohlbergian and similar types of philosophical dilemma (Bush, Krebs, & Carpendale, 1993; Wark & Krebs, 1996). The results, however, suggest this level of consistency is not exclusively due to within-person cognitive structures, as Kohlberg argued, but also to structural similarities in the dilemmas. Kohlberg's (1984; Colby & Kohlberg, 1987) claim that moral judgment is structurally consistent was based

on his examination of people's judgments about hypothetical, philosophical dilemmas. The findings suggest that people's moral judgments also may be structurally consistent across real-life dilemmas involving similar issues. Indeed, 85% and 89% of the participants obtained scores at the same or an adjacent substage across the two antisocial dilemmas and across the two prosocial dilemmas, respectively.

Participants invoked different forms of moral judgment in response to different types of moral dilemma: The dilemmas on Kohlberg's test evoked Stage 3 and 3/4 moral judgments, the prosocial real-life dilemmas evoked Stage 3 moral judgments, and the antisocial real-life dilemmas evoked Stage 2 and 2/3 moral judgments (cf. Wark & Krebs, 1996). Clearly, contrary to Kohlberg's position, the structure of moral judgment was influenced by the type of dilemma. However, the pull of the dilemmas for particular stage-structures was not so strong as to cancel out all individual differences. For example, a small number of participants made Stage 2/3 moral judgments in response to the dilemmas on Kohlberg's test, some made Stage 4 moral judgments in response to the prosocial dilemmas, and some made Stage 3/4 moral judgments in response to the antisocial dilemmas.

It is important to emphasize that consistency was highest across similar types of dilemma because similar types of dilemma evoked similar forms of moral judgment (stage-structures). It is interesting to note, however, that the forms of moral judgment evoked by the dilemmas on Kohlberg's test were more similar to--that is, there was more overlap with--the forms of moral judgment evoked by the prosocial dilemmas than by the antisocial dilemmas and, consequently, the level of consistency between the dilemmas on Kohlberg's test and the prosocial dilemmas was higher than between the dilemmas on Kohlberg's test and the antisocial dilemmas. Similarly, there was higher consistency across the prosocial and antisocial dilemmas than across the Kohlbergian antisocial dilemmas. In general, pairs of similar dilemma related to other pairs of dilemma in similar ways. For example, the moral stage scores of both the dilemmas on Kohlberg's

test were more consistent with the moral stage scores of the prosocial dilemmas than of the antisocial dilemmas.

Why did participants invoke lower stage structures when they responded to the real-life dilemmas than when they responded to the dilemmas on Kohlberg's test?

Participants performed well below their level of moral competence in the real-life moral situations, but why? Some insight is gained by examining the participants' particular responses to each dilemma. In contrast to the hypothetical dilemmas on Kohlberg's test, the real-life dilemmas involved making choices involving actual and immediate consequences to self and other, and typically took place in circumstances constrained by limited time. The decisions involved in the antisocial dilemmas centered around self-satisfaction and reward, concern for punishment, fairness, and consequences to self and others. The dilemmas on Kohlberg's test raise questions about societal and institutional issues. In response to the Kohlbergian Dilemma III, one participant, for example, expressed a Stage 4 judgment that laws should be followed "because laws are usually based on moral rules agreed to by a society, and to follow them is to contribute to a healthy society" (CJ #26, Form A, Law). However, when this participant was faced with the decision about whether to obey the law or not in a real-life situation, what she ultimately felt she should do was influenced by the circumstances she found herself in, as illustrated in the following Stage 2 judgment: "I did what I could in the circumstances. In a perfect world, I would have reported the amount as I made it, but I made my decision ... the alternative was starvation or suffering" (CJ #7, Form A, Life). In everyday life, people typically do not have the luxury to sit back and contemplate the societal ramifications of each and every moral decision they are faced with and must make. People's moral judgment in everyday life may not typically stem from lofty ideals, but, rather, from the immediate circumstances and consequences and from emotional responses. In support of this, a significant proportion of the issues participants in Study 1

perceived to be involved in (descriptions of) real-life dilemmas were self-oriented in nature.

Kohlberg (Colby & Kohlberg, 1987) defined Stage 2 morality as individualistic and instrumental. I examined the distribution of elements across stage in Colby and Kohlberg's (1987) scoring manual and found that nearly one half (47%) of all the criterion judgments classified as upholding the seeking reward/avoiding punishment element were at Stage 2. Only one criterion judgment was at Stage 2/3, and only one was at Stage 3. There were no criterion judgments upholding this element found at higher stages. The reason why the moral judgments participants made about the antisocial dilemmas were scored at relatively low stages may be because most of their judgments were based on self (individual) consequences and self-oriented reward and punishment issues.

The structure of the prosocial dilemmas differed from the structure of the antisocial dilemmas in that they involved mostly caring and relationship-type issues. In the prosocial dilemmas, participants were drawn to attend mostly to role obligations (e.g., whether or not to assume responsibility to help someone in need) and immediate relationship issues. Kohlberg (Colby & Kohlberg, 1987) defined Stage 3 morality as involving concerns about living up to others' expectations, fulfilling role obligations, being a good person, showing concern for others, and the maintenance of trust, loyalty, and respect in relationships. When the same participant who is quoted above was faced with the decision about whether to help a friend, she focussed not on herself as she did in the antisocial dilemma, but on the bond that she and her friend shared, as illustrated in the following Stage 3 judgment: "I decided against this because I wanted to help my friend and felt that refusing would jeopardize our relationship ... loyalty is more important than self-righteousness" (CJ #18, Form A, Contract). In contrast to the antisocial dilemmas, which tended to invoke self-oriented interpretations, the circumstances in the prosocial dilemmas encouraged participants to adopt others' (their friends') perspectives. In

support of this, the participants in Study 1 listed issues about listening to and understanding others and about anticipating others' reactions more often for the prosocial dilemmas than for the antisocial dilemmas.

The results of this study also did not support Gilligan's (1982, 1988) assertions about gender-related differences in moral judgment: Females did not score at lower stages than males on Kohlberg's test or on any of the real-life dilemmas, and females did not make more care-based moral judgments than males. In fact, when participants were asked to rate the dilemmas on the extent to which they viewed them as involving issues of care and issues of justice, females rated all dilemmas as involving more issues of justice than males rated them. Although gender differences in real-life moral orientation have been found in some studies (e.g., Gilligan & Attanucci, 1988a; Lyons, 1983; Pratt et al., 1988; Wark & Krebs, 1996), these differences may have been confounded by a tendency for males and females to report, and therefore to make judgments about, different types of dilemma. In the present research, a large sample of real-life moral reasoning was obtained; content related to type of dilemma was held constant, and males and females were found to display the same forms of moral judgment (both stage and orientation) to the same types of dilemma.

Consistent with Gilligan's theorizing, there was a positive relation between moral maturity and the extent to which being a just person was deemed as important to the self-concept, but this relation was weak and pertained only to the male participants. I did not find a negative relation between moral maturity and the extent to which participants deemed being a caring person as important to their identity or self-concept, nor did I find any relations between gender, self-concept and moral orientation (cf. Wark & Krebs, 1996). Although females rated being a caring person higher in importance to their self-concept than being a just person, which supports Gilligan's position, there were no differences between males and females in terms of the extent to which being a caring or

just person was rated as important to their identity or self-concepts. Considered together, these findings supply little support for Gilligan's position.

Contrary to Gilligan's (1982, 1986, 1988) model, moral orientation was not consistent within people across moral dilemmas. None of the participants obtained the same moral orientation score across the six dilemmas, and only two of the participants obtained the same or an adjacent moral orientation score. This low level of consistency has been reported by other investigators (e.g., Krebs et al., 1994; Pratt et al., 1988; Rothbart et al., 1986; Walker et al., 1987; Wark & Krebs, 1996) and is consistent with the results of Study 1 and Study 2. However, like the moral stage-structure of people's judgments, the level of consistency in moral orientation improved when examined across similar types of dilemma (e.g., across the two prosocial dilemmas). Consistent with the results reported in Wark and Krebs (1996) and in Study 2, participants made more care-based moral judgments about the prosocial dilemmas than about the antisocial dilemmas and the dilemmas on Kohlberg's test.

The moral orientation of the judgments evoked by the dilemmas on Kohlberg's test were more similar to and more consistent with the moral orientation of the judgments evoked by the antisocial dilemmas than the judgments evoked by the prosocial dilemmas. This is because the dilemmas on Kohlberg's test and the antisocial dilemmas evoked primarily justice-based moral judgments, whereas the prosocial dilemmas evoked primarily care-based moral judgments. Once again, pairs of similar dilemma related to other pairs of dilemma in similar ways. It is interesting to note that for moral stage, the Kohlbergian dilemmas were more similar to the prosocial dilemmas than to the antisocial dilemmas, but for moral orientation, the Kohlbergian dilemmas were more similar to the antisocial dilemmas than the prosocial dilemmas.

Consistent with Study 1 and Study 2, males and females rated (a) the prosocial dilemmas as involving more care-based issues than they rated the antisocial dilemmas as involving, and (b) the prosocial dilemmas as involving fewer justice-based issues than

they rated the antisocial dilemmas as involving. Consistent with Study 2, the dilemmas on Kohlberg's test were rated as involving more care-based issues than the antisocial dilemmas were rated as involving. Participants also rated the dilemmas on Kohlberg's test as involving more issues of care than the prosocial, loyalty dilemma and, consistent with the results reported in Study 2, there was no difference in ratings of care between the dilemmas on Kohlberg's test and the prosocial, helping dilemma.

In Study 1 and Study 2, participants were asked to rate each dilemma on the extent to which each involved issues of care versus issues of justice. In this study, participants were asked to rate each dilemma on the extent to which each involved issues of care and issues of justice; the participants were provided with two scales, one for care and one for justice. The findings show that care and justice are not mirror images of each other--the relationship is not one of either/or. (The correlations between the care and justice ratings for each dilemma were nonsignificant, $r_s(78) = -.007$ to $-.21$, ns.) Although people may emphasize one type of issue over the other in a dilemma, they always view dilemmas (at least the ones in this study) as involving both.

Inspection of the care ratings and justice ratings for each dilemma revealed that males and females rated the prosocial dilemmas significantly higher on the care scale than they did on the justice scale, and males and females rated one of the antisocial dilemmas (temptation) significantly higher on the justice scale than they did on the care scale. This pattern is consistent with the finding that males and females made (a) more care-based moral judgments than justice-based moral judgments about the prosocial dilemmas and (b) more justice-based moral judgments than care-based moral judgments about the antisocial dilemmas. In other words, males and females made care-based or justice-based moral judgments about dilemmas they viewed as involving predominantly care-based issues or predominantly justice-based issues, respectively. The prosocial dilemmas emphasized or clearly involved issues of care (e.g., concern with relationships and with helping another) and the antisocial, temptation dilemma emphasized or clearly involved

issues of justice (e.g., violations of law, rules, and fairness)¹². However, it is important to note that none of the dilemmas in this study was perceived in entirely care or justice terms. Participants "read" care-based issues into the antisocial dilemmas (particularly the antisocial, transgression dilemma) and participants read justice-based issues into the prosocial dilemmas.

Interestingly, although participants made more justice-based moral judgments than care-based moral judgments in response to the antisocial, transgression dilemma and in response to the dilemmas on Kohlberg's test (see Table 8), they did not rate these dilemmas higher on the justice scale than on the care scale (see Table 11). Participants rated the antisocial, transgression dilemma as involving both care and justice issues, and, contrary to Gilligan's (1982, 1988) assertions, participants rated the dilemmas on Kohlberg's test significantly higher on the care scale than they rated them on the justice scale. Although the antisocial, transgression dilemma was rated as involving fewer care-based issues than the prosocial dilemmas, the dilemmas on Kohlberg's test were rated as involving at least equal amounts of care-based issues as the prosocial dilemmas. Note, however, that the dilemmas on Kohlberg's test were rated as involving as many justice-based issues as the antisocial dilemmas.

Why did participants make more justice-based moral judgments than care-based moral judgments about the dilemmas on Kohlberg's test and about the antisocial, transgression dilemma when they rated them as involving high levels of both care-based and justice-based issues? What was it about these types of dilemma that evoked justice-based rather than care-based moral judgments? Are all dilemmas involving care and justice resolved in favor of justice (cf. Kohlberg, 1984)?

It is possible that people deem justice-based issues as more moral in nature than care-based issues and, consequently, when people are confronted with moral dilemmas involving both types of issue, they resolve the dilemmas from the most moral perspective (i.e., justice). Given this possibility, I would expect people to deem prosocial dilemmas,

which involve predominantly care-based issues, as less moral in nature than antisocial dilemmas, which involve predominantly justice-based issues. The participants in the present study, however, did not rate the prosocial dilemmas as any less moral than they rated the antisocial dilemmas. Further, even though participants rated the dilemmas on Kohlberg's test as involving relatively high levels of care-based issues, the dilemmas on Kohlberg's test were rated as involving significantly more moral concerns than any of the other types of dilemma. There was a significant positive correlation between care ratings and moral ratings for the Kohlbergian dilemmas ($r(78) = .32, p < .01$). In contrast, there was a significant positive correlation between justice ratings and moral ratings for the antisocial, transgression dilemma ($r(78) = .29, p < .01$).

It is possible that the tendency for both males and females to make more justice-based than care-based moral judgments about the Kohlbergian and antisocial, transgression types of dilemma may have stemmed from an association between particular moral stage structures and particular moral orientations. In the Wark and Krebs (1996) study, we found that the criterion judgments in Kohlberg's scoring manual were predominantly justice based at all stages, but that the proportion of justice-based to care-based criterion judgments varied across stages¹³. In particular, the proportion of care-based criterion judgments is relatively high at Stages 2/3 and 3 (25% and 23%, respectively), and relatively low at Stages 2 and 4 (15% and 4%, respectively). It is possible that in invoking high proportions of Stage 3 moral judgments (70%), the prosocial dilemma evoked high proportions of care-based moral judgments. Even though the dilemmas on Kohlberg's test and the antisocial, transgression dilemma were rated as involving high levels of both care-based and justice-based issues, these dilemmas evoked fewer care-based moral judgments because they evoked more Stage 4 and Stage 2 moral judgments, respectively (see Table 9). This explanation, however, is limited because not all of the considerations scored for moral orientation were scored for moral stage.

As suggested in Study 2, it is possible that participants oriented their judgments around the issues inherent or most salient in the dilemmas. Although participants in Study 1 listed both care-based and justice-based issues for all dilemmas, they focussed on those that were most salient in the dilemmas. It is possible that when the participants in Study 3 made judgments about the antisocial dilemma, they may have oriented their judgments around the most salient issue, namely, the justice-based issue about combatting immorality. As discussed in Study 2, I did not ask participants to list the types of issue they saw to be involved in the dilemmas on Kohlberg's test. It is possible that participants made primarily justice-based moral judgments because the most salient issues they perceived the dilemmas to involve were justice based.

According to Nunner-Winkler (1984), Kohlberg's dilemmas involve conflicts between care and justice and, therefore, should evoke both care and justice moral reasoning. However, it is possible that participants made more justice-based moral judgments than care-based moral judgments to the dilemmas on Kohlberg's test because of the nature of the probe questions. That is, the questions on Kohlberg's test may have pulled more for justice-based moral judgments than care-based moral judgments.

As in Study 1 and Study 2, I investigated possible reasons why females in the Wark and Krebs (1996) study reported more prosocial real-life dilemmas than males, and why males reported more antisocial real-life dilemmas than females. Across the three studies, I found a tendency for females to attach more significance to prosocial dilemmas than males attach to them. Across Study 1 and Study 2, I found a tendency for males to report experiencing antisocial moral dilemmas to a greater extent than females reported experiencing them, but this pattern was not replicated in Study 3. It is conceivable that the difference in results may have stemmed from a difference in the format of the dilemmas employed: The participants in Study 1 and Study 2 responded to descriptions of dilemmas, whereas the participants in Study 3 reported and made judgments about dilemmas they had actually experienced. It is possible that the process of actively

recalling and reasoning about different types of actually-experienced dilemmas alters the perceived salience or experience of each type.

One interesting gender difference did emerge: Females rated the decisions involved in the dilemmas as more difficult than males rated them. Although this finding does not supply an explanation for the observed gender differences in type of real-life dilemma reported, it is an intriguing result. Although females rated the decisions in all types of real-life dilemma as relatively difficult, when asked to discuss a real-life dilemma, females tend to report the type (prosocial) involving the most difficult decisions. A similar finding was reported by Ford and Lowery (1986). They found that females rated their real-life dilemmas as more important in their lives and as involving more difficult decisions than males rated their real-life dilemmas, and, for both males and females, the more important the conflicts rated, the more care-based reasoning was used.

Why males and females tend to report different types of real-life dilemma remains to be answered. It is possible that females experience more guilt with regard to prosocial types of real-life dilemma than males experience, and males may experience more guilt with regard to antisocial types of real-life dilemma than females experience. Although I did not explore this possibility in the present study, some research (Williams & Bybee, 1994) suggests that females feel guilty about different types of transgression from those males feel guilty about. Further investigation of this possible gender difference should prove fruitful, particularly if it can be established that males and females report the types of real-life dilemma they feel most guilty about. How we interpret and experience moral situations is at least as important as how we resolve them (Walker, Pitts, Hennig, & Matsuba, 1995).

General Discussion

According to Kohlberg, people interpret and understand moral information in terms of general, highly organized cognitive "structures of the whole" that define their current stage of moral development (Colby & Kohlberg, 1987). According to Kohlberg (1984),

moral development involves the transformation and displacement of old stages by new, more advanced stages of moral reasoning. People, therefore, cannot invoke lower stage forms of moral reasoning, because they are no longer available. The results of this study, however, support an "additive-inclusive" model of stage change (see Levine, 1979) in which lower stage-structures are retained and invoked in response to certain types of moral dilemma.

I found that moral judgment was structurally consistent across similar types of moral dilemma (e.g., across the two Kohlbergian dilemmas, across the two prosocial dilemmas, and across the two antisocial dilemmas). However, different types of moral dilemma evoked different stages of moral judgment. In support of the results reported in Wark and Krebs (1996), I found that the prosocial dilemmas tended to evoke Stage 3 moral judgments and the antisocial dilemmas tended to evoke Stage 2 and 2/3 moral judgments. In support of Kohlberg's (Colby & Kohlberg, 1987) contention that his test assesses moral competence, the dilemmas on Kohlberg's test tended to evoke Stage 3, 3/4, and higher moral judgments from the participants.

The results suggest that the almost exclusive use by researchers of the moral dilemmas on Kohlberg's test may have led to a fundamentally flawed model of how people make moral judgments to dilemmas of varying content and, especially, to dilemmas encountered in everyday life. Although the examination of people's moral judgments to real-life dilemmas suggests that many of the issues people raise are similar to the moral elements identified by Kohlberg in people's responses to the hypothetical dilemmas on his test, some of the issues the participants raised are not well represented in Kohlberg's scheme. Kohlberg (1984) believed that morality is based in a concern with justice. In this research, it was learned that people do indeed attend to justice-based issues in real-life dilemmas, but it was also learned that people perceive and attend to issues about caring for others, relationships, and about themselves.

Although Kohlberg's model of moral judgment has taught us a great deal about the development of people's most advanced ability to understand philosophical moral problems--the development of moral competence--it has not provided us with an adequate model of how people actually make moral decisions in their everyday lives. Surely, our ultimate goal is to understand how people make moral decisions in everyday life. Although people may make high-stage moral judgments in response to the hypothetical dilemmas on Kohlberg's test, the constraints inherent in real-life dilemmas may lead people to invoke lower stage structures. It therefore becomes important to investigate further the structures and constraints inherent in real-life moral situations. In this study, it was learned that the structure of the antisocial dilemmas pulled participants to attend to issues about themselves and about fairness, whereas the structure of the prosocial dilemmas pulled participants to attend to issues about caring for others and about relationships. Clearly, we need more than Kohlberg's test to predict people's moral judgment and moral behavior in everyday life. If we want to understand how people make real-life moral judgments, we must attend to how people understand and respond to real-life moral situations.

According to Gilligan (1982, 1988), moral orientation stems from gender-related differences in cognitive orientation; thus, moral orientation should be consistent within people across varying content or types of moral dilemma. The results of this research, however, suggest that moral orientation is largely a function of the type of moral dilemma about which people reason. I found no unqualified gender differences in moral orientation. Both males and females viewed prosocial dilemmas in more care-based ways than they viewed antisocial dilemmas. The prosocial dilemmas involved issues about caring for others and relationships--issues that Gilligan considers care based. The antisocial dilemmas involved issues centered around self consequences, combatting immorality, and fairness--issues that Gilligan considers justice based.

There was a tendency toward a gender difference in moral orientation: Females tended to view social pressure, parent types of dilemma from a more care-based perspective than males viewed them. The female participants in Study 3 rated being a caring person higher in importance to their self-concept than being a just person, but there were no relations between self-concept ratings and moral orientation. I also found a gender difference that contradicted Gilligan's expectations: The female participants in Study 3 rated all the dilemmas as involving more justice-based issues than males rated them as involving.

To conclude, it would seem that Gilligan's (1982, 1988) model of moral judgment, like Kohlberg's model, has underestimated the strength of the determining influence of different types of dilemma. The results of this research suggest a more complex, interactional model. Dilemmas may influence the issues people see in them, and dilemmas may influence the type of moral judgments people make about them. Although I found individual differences in the ways dilemmas are viewed, these differences were not gender-related, as Gilligan contended.

I encourage further investigation of the types of real-life dilemma people encounter, especially across different life stages, and examination of the relation between moral experience and moral performance across the lifespan. Research on developmental differences in real-life moral reasoning has begun (e.g., Walker et al., 1995), and promises to contribute significantly to the field of moral development, especially in light of the implications for moral education and moral experience. Researchers such as Damon (1977) and Grusec and Goodnow (1994) have shown that children's moral judgments may be context and domain specific, and, as discussed elsewhere (Wark & Krebs, 1996), moral judgment may become more context and domain specific with development.

References

- Bartek, S., Krebs, D. L., & Taylor, M. (1993). Coping, defending, and the relations between moral judgment and moral behavior in prostitutes and other female juvenile delinquents. Journal of Abnormal Psychology, *102*, 65-73.
- Brown, L. M. (Ed.). (1987). A guide to reading narratives of moral conflict and choice for self and moral voice. Unpublished manuscript, Harvard University, Gehd Study Center, Boston.
- Bush, A., Krebs, D. L., & Carpendale, J. (1993). The structural consistency of moral judgments about AIDS. Journal of Genetic Psychology, *154*, 167-176.
- Carpendale, J. I. M., & Krebs, D. L. (1992). Situational variation in moral judgment: In a stage or on a stage? Journal of Youth and Adolescence, *21*(2), 203-224.
- Carpendale, J. I. M., & Krebs, D. L. (in press). Variations in moral judgment as a function of type of dilemma and moral choice. Journal of Personality.
- Colby, A., & Kohlberg, L. (1987). The measurement of moral judgment. (Vol. 1). New York: Cambridge University Press.
- Damon, W. (1977). The social world of the child. San Francisco: Jossey-Bass.
- Denton, K., & Krebs, D. L. (1990). From the scene to the crime: The effect of alcohol and social context on moral judgment. Journal of Personality and Social Psychology, *59*, 242-248.
- Donenberg, G., & Hoffman, L. (1988). Gender differences in moral development. Sex Roles, *18*, 701-717.
- Ford, M., & Lowery, C. (1986). Gender differences in moral reasoning: A comparison of the use of justice and care orientations. Journal of Personality and Social Psychology, *50*(4), 777-783.
- Gilligan, C. (1977). In a different voice: Women's conceptions of self and of morality. Harvard Educational Review, *47*, 481-517.
- Gilligan, C. (1982). In a different voice. Cambridge: Harvard University Press.

Gilligan, C. (1986). Reply (to critics). Signs: Journal of Women in Culture and Society, 11, 324-333.

Gilligan, C. (1988). Adolescent development reconsidered. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), Mapping the moral domain: A contribution of women's thinking to psychological theory and education (pp. vii-xxxix). Cambridge, MA: Harvard University Press.

Gilligan, C., & Attanucci, J. (1988a). Two moral orientations: Gender differences and similarities. Merrill-Palmer Quarterly, 34, 223-237.

Gilligan, C., & Attanucci, J. (1988b). Much ado about ... knowing? Noting? Nothing? A reply to Vasudev concerning sex differences and moral development. Merrill-Palmer Quarterly, 34, 451-456.

Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. Developmental Psychology, 30, 4-19.

Higgins, A., Power, C., & Kohlberg, L. (1984). The relationship of moral atmosphere to judgments of responsibility. In W. M. Kurtines & J. L. Gewirtz (Eds.), Morality, moral behavior, and moral development (pp. 74-106). New York: Wiley.

Kohlberg, L. (1984). Synopses and detailed replies to critics. In L. Kohlberg (Ed.), Essays on moral development (Vol. 2): The psychology of moral development (pp. 338-361). New York: Harper & Row.

Krebs, D. L., Denton, K., Vermeulen, S. C., Carpendale, J. I., & Bush, A. (1991). The structural flexibility of moral judgment. Journal of Personality and Social Psychology, 61, 1012-1023.

Krebs, D. L., Vermeulen, S. C., Carpendale, J. I., & Denton K. (1991). Structural and situational influences on moral judgment: The interaction between stage and dilemma. In W. Kurtines & J. Gewirtz (Eds.), Handbook of moral behavior and development: Theory, research, and application (pp. 139-169). Lawrence Erlbaum Publishers.

Krebs, D. L., Vermeulen, S. C., & Denton, K. L. (1991). Competence and performance in moral judgment: From the ideal to the real. Moral Education Forum, *16*, 7-22.

Krebs, D. L., Vermeulen, S. C., Denton, K. L., & Carpendale, J. I. (1994). Gender and perspective differences in moral judgment and moral orientation. Journal of Moral Education, *23*, 17-26.

Leahey, R., & Eiter, M. (1980). Moral judgment and the development of real and ideal androgynous self-image during adolescence and young adulthood. Developmental Psychology, *16*, 362-370.

Levine, C. G. (1979). Stage acquisition and stage use: An appraisal of stage displacement explanations of variation in moral reasoning. Human Development, *22*, 145-164.

Lifton, P. (1985). Individual differences in moral development: The relation of sex, gender, and personality to morality. Journal of Personality, *53*(2), 306-334.

Linn, R. (1984). Practising moral judgment within the day care center: A look at the educator's moral decision under stress. Early Child Development and Care, *15*, 117-132.

Linn, R. (1987a). Moral disobedience during the Lebanon war: What can the cognitive-developmental approach learn from the experience of the Israeli soldiers? Social Cognition, *5*, 383-402.

Linn, R. (1987b). Moral reasoning and behavior of striking physicians in Israel. Psychological Reports, *60*, 443-453.

Lyons, N. (1983). Two perspectives: On self, relationships, and morality. Harvard Educational Review, *53*, 125-145.

Nunner-Winkler, G. (1984). Two moralities? A critical dimension of an ethic of care and responsibility versus an ethic of rights and justice. In W. M. Kurtines & J. L. Gewirtz (Eds.), Morality, moral behavior, and moral development (pp. 348-361). New York: Wiley.

Pratt, M., Golding, G., & Hunter, W. (1984). Does morality have a gender? Merrill Palmer Quarterly, *30*, 321-340.

Pratt, M., Golding, G., Hunter, W., & Sampson, R. (1988). Sex differences in adult moral orientations. Journal of Personality, *56*(2), 373-391.

Pratt, M. W., Golding, G., & Kerig, P. (1987). Lifespan differences in adult thinking about hypothetical and personal moral issues: Reflection or regression? International Journal of Behavioral Development, *10*(3), 359-375.

Pratt, M., & Royer, J. (1982). When rights and responsibilities don't mix: Sex and sex-role patterns in moral judgment orientation. Canadian Journal of Behavioral Science, *14*(3), 190-204.

Rothbart, M., Hanley, D., & Albert, M. (1986). Gender differences in moral reasoning. Sex Roles, *15*, 645-653.

Spence, J. T., & Helmreich, R. L. (1978). Masculinity and femininity: Their psychological dimensions, correlates, and antecedents. Austin, TX: University of Texas Press.

Vasudev, J. (1988). Sex differences in morality and moral orientation: A discussion of the Gilligan and Attanucci study. Merrill-Palmer Quarterly, *34*, 239-244.

Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. Child Development, *55*, 677-691.

Walker, L. J. (1991). Sexism in moral psychology. In W. M. Kurtines & J. L. Gewirtz (Eds.), Moral behavior and development: An introduction. New York: McGraw-Hill.

Walker, L. J., de Vries, B., & Trevethan, S. (1987). Moral stages and moral orientations. Child Development, *58*, 842-858.

Walker, L. J., Pitts, R. C., Hennig, K. H., & Matsuba, M. K. (1995). Reasoning about morality and real-life moral problems. In M. Killen & D. Hart (Eds.), Morality in everyday life: Developmental perspectives. New York: Cambridge University Press.

Wark, G. R., & Krebs, D. L. (1996). Gender and dilemma differences in real-life moral judgment. Developmental Psychology, *32*, 220-230.

Williams, C., & Bybee, J. (1994). What do children feel guilty about? Developmental and gender differences. Developmental Psychology, *30*, 617-623.

Yussen, S. (1977). Characteristics of moral dilemmas written by adolescents. Developmental Psychology, *13*, 162-163.

Author's notes

The research described in this article was supported by Doctoral Fellowship #752-94-1253 and by Grant #410-94-0345 from the Social Sciences and Humanities Research Council of Canada. The assistance of Danielle Krebs, Sophie Bartek, and all the members of the Simon Fraser University social development lab is gratefully acknowledged. Support from John Wark and Rosemary Wark is considered invaluable. Correspondence regarding this article should be sent to Gillian Wark, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada, V5A-1S6, (604) 291-3354, Fax #291-3427.

Footnotes

¹Wark and Krebs (1996) employed the distinction made post hoc by Walker et al. (1987) and asked participants to recall and to describe two significant real-life moral conflicts, one that did not directly involve them (impersonal real-life dilemma) and one that directly involved them and another person or group of people with whom they had a significant relationship (personal real-life dilemma).

²Of those dilemmas classified in Wark and Krebs (1996) as antisocial, 73% were about transgressions, and 76% of those classified as prosocial were about helping.

³Participants in the Wark and Krebs (1996) study were drawn from the same population as participants in the present research. Samples from each study did not differ in terms of age, SES, grade point average, or marital status.

⁴All material (e.g., questionnaires) employed in the present research is available on request of the author.

⁵Two additional types of dilemma were provided, but only the six representing the three most frequent types were included in statistical analyses.

⁶As described in Walker (1984; p. 687), the Stouffer method involves (a) computation of the exact p value of the test statistic, (b) computation of the Z scores associated with each p value, (c) division of the sum of the Z scores by the square root of the number of test results being combined, and (d) computation of the appropriate p value for this overall Z score. This latter p value is the probability level for the observed pattern of results.

⁷Due to the time and effort demands placed on the participants, I did not ask them to report and to make moral judgments about social pressure types of real-life dilemma.

⁸Although a total of over 100 participants volunteered for the present study, only those participants who supplied complete sets of data were employed. Some participants did not complete the set of questionnaires due to time and effort constraints, and some failed to follow the instructions. An examination of the uncompleted sets of questionnaires did not reveal any systematic patterns in the types of questionnaires or

dilemmas left unanswered. The participants who did not supply complete and appropriate sets of data did not differ from those who did in terms of age, SES, GPA, or scores on Kohlberg's test.

⁹A 2 X 5 ANCOVA on moral maturity, controlling for self-concept (on the justice scale), was conducted. The analyses revealed a significant covariate effect, $F(1, 77) = 11.11, p < .0013$. However, the pattern of effects were identical to the original ANOVAs.

¹⁰A 2 X 6 ANCOVA on moral maturity, controlling for moral orientation, and a 2 X 6 ANCOVA on moral orientation, controlling for moral maturity, were conducted. The analyses revealed significant covariate effects, $F_s(1, 389) = 3.92, p < .05$. However, the pattern of effects were identical to the original ANOVAs.

¹¹To determine whether the pattern of results changed across studies, a 2 X 3 (Gender X Study) ANOVA was conducted on the Likert-scale ratings of the extent to which participants have experienced dilemmas similar to the antisocial, transgression dilemma. The ANOVA failed to produce a significant main effect for Gender, $F(1, 164) = 2.10$, or a significant main effect for Study, $F(2, 164) = 0.56$. There was a marginally significant interaction effect, $F(2, 164) = 3.02, p < .051$. As would be expected, analyses of simple effects revealed that the pattern of differences between the genders was the same only across Study 1 and Study 2.

A 2 X 3 (Gender X Study) ANOVA was conducted on the Likert-scale ratings of the extent to which participants viewed the prosocial, helping dilemma as significant. The ANOVA revealed a significant main effect for Gender, $F(1, 164) = 10.51, p < .001$: Across studies, females rated the helping dilemmas as more significant than males rated them. The main effect for Study was significant, $F(2, 164) = 3.85, p < .023$. The interaction between Gender and Study was not significant, $F(2, 164) = 2.60$, indicating that the pattern of differences between the genders did not change across studies.

¹²Demonstrating that males and females view prosocial types of dilemma as involving primarily care-based issues and antisocial types of dilemma as involving primarily justice-based issues is not tautological because participants' responses were not simple restatements of the descriptions they read. For example, participants raised

several care-based issues (e.g., trust, empathy, understanding others, loyalty within the context of relationships, caring, concern for others and others' feelings, etc.) when they were reasoning about the prosocial dilemmas that were not referred to in the instructions they were provided with for those types of dilemma, and participants raised several justice-based issues (e.g., concern with upholding principles for self and society, punishment and retribution, interfering in others' business, respecting the rights of self and others, etc.) when they were reasoning about the antisocial dilemmas that were not referred to in the instructions they were provided with for those types of dilemma. Although the perceived presence of certain issues (e.g., justice-based issues) may have made associated issues (e.g., other justice-based issues) more salient, I do not believe this is tautological. Rather, I believe that these results demonstrate the determining power of the type of real-life dilemma on how people perceive different situations and, therefore, on the forms of moral judgment people invoke in different situations.

¹³Two scorers independently scored all 708 criterion judgments in Colby and Kohlberg's (1987) scoring manual for care and justice (see Wark and Krebs, 1996). There was 91.8% agreement ($Kappa = .89$) between the two scorers. The following percentages represent the percent of all the judgments at a particular stage that were scored as care based (i.e., C or C(J)): Stage 1: 0%; Stage 1/2: 0%; Stage 2: 15%; Stage 2/3: 25%; Stage 3: 23%; Stage 3/4: 20%; Stage 4: 4%; Stage 4/5: 10%; and Stage 5: 3%.

Table 1

Descriptions of Issue CategoriesI. Upholding Justice

Normative order: Concern with upholding laws, codes of conduct, rules, religion, and concern with legality of events/occurrences (1%).

Combatting immorality: Concern with the righting of wrongs or with getting even (e.g., evening things up by transgressing; justifying transgressions in terms of principles of equity or reciprocity; retribution; reporting (repeat) offenders (6%).

Positive reciprocity: Concern with upholding positive reciprocity through (a) upholding deservingness (e.g., earning reward through hard and honest work), (b) upholding fairness (e.g., how fair one is to others, upholding fairness in general), (c) upholding reciprocity (e.g., pay back favors) (2%).

Procedural fairness: Concern with making fair decisions through (a) balancing perspectives (e.g., considering perspectives of all parties involved, including self and other), (b) reaching agreement or compromise, through communication and diplomacy, (c) analyzing situations and problems with objectivity and/or neutrality (e.g., assessing accuracy of evidence, considering who has the right or objective assessment of situation, having an open mind), and (d) consideration of circumstances, factors involved, and degree of seriousness in situation (8%).

General utilitarian considerations: Concern with consequences (and their severity) of actions or decisions to either (a) unspecified other (e.g., consequences) or (b) society (e.g., concern for general good, avoidance of chaos in society) (1%).

table 1 continues

II. Self-Oriented Consequences

Consequences to self: Concern with gains/losses to self (e.g., avoiding situations that pose difficulty or bother, concern with self's happiness and other benefits, concern with avoidance of losses or harm to self, concern with relevance of situations to self) (12%).

Consequences to self-reputation: Concern with (a) maintenance of good reputation and avoidance of bad reputation and (b) approval by others (e.g., what would others think, would they approve?) (2%).

Consequences to self-respect: Concern with how one view's oneself in terms of (a) being true to oneself and/or abiding by one's standards/morals (e.g., honesty with oneself, integrity, concern with personal values), (b) how one feels about oneself (e.g., ability to live with oneself or one's decision, self-image, self-worth), and (c) guilt or sense of conscience (8%).

Self-autonomy: Concern with making one's own decisions without being influenced by others, and concern with independence, self's rights, and self choice (e.g., doing what self thinks is right, going by self' judgment) (13%).

Other's respect for and trust in self: Concern with (a) maintaining others respect for self, self's right to decisions, self-autonomy (e.g., friends should respect self's decision, will self lose other's respect) and (b) maintaining other's trust in self (e.g., will self lose other's trust?) (2%).

III. Other-Oriented Consequences

Listening to, considering, and understanding perspectives: Considering and identifying others' thoughts, feelings, and opinions (e.g., considering whether someone knows that something is wrong or that they need help, considering the importance or value of a position another is holding, how another person feels), and understanding others' motives and intentions (e.g., why someone is requesting something of the self, why someone has a certain point of view or has done something) (3%).

table 1 continues

Putting self in other's shoes: Understanding a situation or person by placing oneself in the position, to figure out how the self would feel if faced with the same situation (1%).

Adapting self's response to anticipated reactions of others: Consideration of other's reaction to the self's action, inaction, decision, or intervention (e.g., will the other respond or listen to help/advice, how others will react) (2%).

Positive social influence: Encouraging moral behavior by confronting or by offering advice or support to other in order to (a) have them take responsibility for their wrongdoing, (b) correct the wrongdoing, or (c) have other realize that he/she has done wrong (1%).

Respect for others and their rights and autonomy: Concern with upholding (a) respect for and value of others, (b) respect for other's opinions and their value, (c) others' rights, and (d) others' right to decide and their ability to make sound judgments without influence from self (3%).

Caring for others: Concern with helping others, prevention of harm to others, consequences to others, compassion and love, and the responsibility of the self to help others or to please others and make them happy (14%).

IV. Relationship Issues

Upholding relationship: Concern for the maintenance of relationships and relations with others (e.g., connection with others, loss of relationship or support from other, avoidance of conflict or aggravation in relationship, how decisions or actions will affect relationships) (10%).

Upholding trust and honesty in relationships: Concern with the maintenance of trust and honesty in relationships (e.g., other's loss of trust, honesty between friends) (2%).

Loyalty: Concern with conflicting loyalties (e.g., loyalty to duty or to friend) (0.2%).

table 1 continues

Nature of relationship: Concern with and consideration of the nature or quality of relationships (e.g., the degree of closeness to others, how much one cares for other and how much other cares for self, how important the other is to self, how long the self has known the other, history of relationship, openness of relationship) (9%).

Note. The percent of all issues classified (total = 1031) in each category are in brackets. Percentages do not add up to 100 due to rounding.

Table 2

Types of Issue Listed as a Function of Type of Dilemma

Issue	Type of Dilemma							
	Antisocial		Prosocial		Social Pressure			
	Temptation	Transgress	Help	Loyal	Peer	Parent		
<u>Upholding Justice:</u>								
Normative order	2	2	0	0	3	0		
Combatting immorality	1	27	0	0	1	0		
Positive reciprocity	7	3	0	2	0	1		
Procedural fairness	5	9	9	10	7	8		
General Utilitarian	2	1	0	1	3	0		
<u>Self-oriented:</u>								
Self consequences	29	8	5	8	14	9		
Self reputation	6	1	0	2	2	1		

table 2 continues

Type of Dilemma

Issue	Antisocial		Prosocial		Social Pressure	
	Temptation	Transgress	Help	Loyal	Peer	Parent
<u>Self-oriented:</u>						
Self respect	21	6	2	5	11	4
Self autonomy	0	1	1	7	33	44
Others' respect for self	4	0	0	1	5	4
<u>Other-oriented:</u>						
Listening to perspectives	0	2	2	3	0	9
Self in other's shoes	2	1	1	1	0	0
Anticipated reactions	1	2	8	1	0	1
Positive social influence	0	7	0	0	0	0
Respect for others	3	4	4	1	2	4
Caring for others	7	10	37	16	8	6

table 2 continues

Type of Dilemma

Issue	Antisocial		Prosocial		Social Pressure	
	Temptation	Transgress	Help	Loyal	Peer	Parent
<u>Relationship:</u>						
Upholding relationships	3	6	12	28	6	5
Uphold trust in relationships	2	1	3	2	1	0
Loyalty	0	1	0	0	0	0
Nature of relationship	7	8	12	14	6	4
Total	102	100	96	102	102	100

Note. The percentages do not always add up to 100 due to rounding.

Table 3

Types of Issue Listed Consistently Across Dilemmas

Type of Issue	Number of participants who listed issue for at least 3 dilemmas
<u>Broad Categories:</u>	
Upholding Justice	22 (13 males and 9 females)
Self-oriented	30 (16 males and 14 females)
Other-oriented	29 (13 males and 16 females)
Relationship	28 (12 males and 16 females)
<u>Subtypes of Issue:</u>	
Caring for others	15 (7 males and 8 females)
Nature of relationships	11 (7 males and 4 females)
Consequences to self	11 (8 males and 3 females)
Upholding relationships	8 (2 male and 6 females)
Self-respect	7 (1 male and 6 females)
Procedural fairness	5 (4 males and 1 female)
Respect for others	2 (1 male and 1 female)
Anticipating others' reactions	1 (female)
Listening to/understanding perspectives	1 (male)
Positive reciprocity	1 (male)
Upholding the law	1 (female)

Table 4

Examples of Classified Moral Judgments

Care:

"I would have to determine which is more important; the health and welfare of my sister or being afraid of what effects this will have on my family."

"... love has no boundaries. If he is passionately in love with her he would give his own life to save hers."

Care(Justice):

"His wife is dying, she should have the company of her husband in her dying days, Heinz is also suffering enough, that should be punishment enough."

"... has an obligation to wife. Love involves commitment and putting the other person's interests ahead of one's own ... marriage is a special relationship."

Care/Justice:

"...by telling everybody what happened I'm protecting others from my cousin and I believe everybody has the right to know."

"I don't consider hurting anyone ... as a right decision ... as far as standing up for and only thinking of my sister, I believe that's right."

Justice(Care):

"... what he is doing is very wrong just because some people are slightly different. He must be stopped ... many people may be suffering ... the needs of the many outweigh the needs of the few."

"We are not the only ones in this world. We must respect others and their opinions/beliefs. We must do our utmost to ensure safety, happiness and stability."

Justice:

"... the professor is discriminating ... he should be treating everyone equally".

"If they are not punished there will not be any order in our society."

Table 5
Moral Orientation as a Function of Gender and Type of Dilemma

Type of Dilemma	Females		Males		Both Genders	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
	Moral Orientation Scores ^a					
Kohlberg	24.1	9.5	22.7	12.5	23.4 _a	10.9
Antisocial	5.7	10.1	5.3	7.4	5.5 _b	8.7
Prosocial	72.2	19.1	61.3	17.3	66.7 _c	18.8
Social Pressure	58.3 _e	16.2	41.0 _f	15.1	49.7 _d	15.7
Mean	40.1	12.9	32.6	12.4	36.3	12.8
	Care/Justice Likert-Scale ^b					
Kohlberg	4.3	2.4	4.2	2.6	4.2 _a	2.5
Antisocial	1.4	1.1	2.0	1.9	1.7 _b	1.6
Prosocial	4.3	2.1	2.9	2.6	3.6 _a	2.3
Social Pressure	4.9	1.7	4.7	2.5	4.8 _a	2.1
Mean	3.7	1.8	3.5	2.4	3.6	2.1

^aMoral orientation scores represent the mean percentage of care-based moral judgments.

^bThe scale ranged from 0-7; the higher the score, the more the care.

Note. Means from the same column with different subscripts were different at the $p < .008$ level or lower. Means for males and females with different subscripts were different at $p < .01$. M = Mean; SD = Standard deviation.

Table 6
Examples of Classified Real-Life Moral Judgments

Real-Life Moral Judgment	Moral Orientation Score	Kohlberg Criterion Judgment Match
I feel I did the right thing because ... I protected myself against getting a possible speeding ticket.	Justice J	[Heinz should not steal] because he would be taking too great a risk (Form A, Law, CJ #8, Stage 2, p. 70)
there was only one option really to consider which was obeying my father since I lived under his roof and he fed and clothed me.	J	[Joe should give his father the money] because his father has done a lot of things for him, fed him, bought his clothes ... (Form A, Authority, CJ #7, Stage 2, p. 249)
to report him is wrong ... his crime didn't take any lives, so it's not too serious.	Care/Justice C/J	[Heinz should not be reported] because he wasn't really hurting anyone ... (Form A, M&C, CJ #10, Stage 2/3, p. 113)
I think that the best or the "right" thing to do is to tell my family ... because its safer for my cousin and for others whom she is selling drugs to.	Care C	[The judge should punish Heinz] in order to make things safer for people (Form A, Punishment, CJ #11, Stage 2/3, p. 156)

table 6 continues

Real-Life Moral Judgment	Moral Orientation Score	Kohlberg Criterion Judgment Match
the other aspect of morality meant I care for my mother and do not want to have her arrested, embarrassed or feel hurt & betrayed by me.	C	[It is important to keep a promise] because if you don't, you hurt the other person's feelings (Form A, Contract, CJ #17, Stage 3, p. 208)
I couldn't stand by while this guy was doing something wrong, [I] decided to tell her because I felt it was my responsibility as a friend to do so.	Care(Justice) C(J)	[Heinz should steal the drug] because he shouldn't just sit back and watch her die ... he is still her husband. (Form A, Life, CJ #13, Stage 3, p. 26)
The police had their job--to maintain that rules & laws protecting my rights and the rights of others are obeyed.	J	[The judge should punish Heinz] because a judge's function is to uphold the law (Form A, Punishment, CJ #20, Stage 3/4, p. 164)
[Did you do the right thing?] Yes, because you have a greater responsibility to your friends.	C/J	[Heinz should steal the drug] because he would feel a responsibility to care or provide for his wife (Form A, Life, CJ #21, Stage 3/4, p. 34)

table 6 continues

Real-Life Moral Judgment	Moral Orientation Score	Kohlberg Criterion Judgment Match
I thought it was wrong because not lying is important in maintaining a cooperative system for a greater benefit to all.	Justice(Care) J(C)	[It is important to keep a promise] for the sake of the orderly or smooth functioning of society (Form A, Contract, CJ #30, Stage 4, p. 223)
At stake was my initial committment and honor to work ... What should have been done was that I continued my job ... as I was obligated to do.	J	[It is important to keep a promise] for the sake of personal honor, integrity, or self-respect (Form A, Contract, CJ #32, Stage 4, p. 225)

Note. The decontextualized moral judgments were interpreted by scorers in the context of participants' general statements.

Table 7

Percent of Participants Obtaining the Same or Adjacent Stage Across Types of Dilemma

	Kohlberg		Prosocial		Antisocial	
	III	III'	Helping	Loyalty	Transgression	Temptation
Kohlberg:						
III	---	95	69	79	50	46
III'		---	75	81	56	49
Prosocial:						
Helping			---	89	75	51
Loyalty				---	79	71
Antisocial:						
Transgression					---	85
Temptation						---

Note. The criterion of structural consistency was same or adjacent stage on a 9-point scale.

Table 8

Mean Moral Maturity and Moral Orientation Scores as a Function of Gender and Type of Dilemma

Dilemma	Females	Males	Both Genders
Moral Maturity			
Kohlberg:			
III	330.8 (48.6)	333.8 (44.4)	332.3 _a (46.3)
III'	319.3 (39.4)	336.8 (44.9)	328.0 _a (42.9)
Prosocial:			
Helping	286.3 (26.6)	289.3 (29.9)	287.8 _b (28.2)
Loyalty	290.0 (28.2)	292.2 (41.9)	291.1 _b (35.5)
Antisocial:			
Transgression	256.8 (37.5)	250.1 (44.0)	253.4 _c (40.8)
Temptation	243.9 (45.7)	240.5 (37.5)	242.2 _c (41.6)
Mean	280.3 (35.9)	281.5 (38.9)	280.9 (37.5)
Moral Orientation			
Kohlberg:			
III	28.8 (13.7)	28.5 (13.7)	28.6 _b (13.7)
III'	28.8 (14.3)	24.4 (18.8)	26.6 _b (16.7)
Prosocial:			
Helping	63.7 (27.3)	64.6 (24.9)	64.2 _a (26.0)
Loyalty	67.5 (24.3)	57.4 (28.6)	62.5 _a (26.9)

table 8 continues

Dilemma	Females	Males	Both Genders
Moral Orientation			
Antisocial:			
Transgression	27.4 (23.2)	26.6 (21.9)	27.0 _b (22.4)
Temptation	13.3 (17.8)	17.6 (25.5)	15.4 _c (21.9)
Mean	40.2 (21.0)	38.5 (22.9)	39.3 (22.0)

Note. Means from the same column that do not share a similar subscript are different at the $p < .003$ level or lower. Standard deviations are in brackets. Moral orientation scores represent the mean percentage of care-based moral judgments.

Table 9

Percent of Participants at Each Moral Stage as a Function of Type of Dilemma

	Moral Stage						
	2	2/3	3	3/4	4	4/5	5
Kohlberg:							
III	0	6	44	30	16	3	1
III'	0	5	49	30	14	3	0
Prosocial:							
Helping	5	19	71	5	0	0	0
Loyalty	1	23	69	3	5	0	0
Antisocial:							
Transgression	24	43	29	4	1	0	0
Temptation	28	55	15	0	3	0	0

Note. Percentages added across stages do not always add up to 100 due to rounding.

Table 10

Percent of Participants Obtaining the Same or Adjacent Moral Orientation Score Across Types of Dilemma

	Kohlberg		Prosocial		Antisocial	
	III	III'	Helping	Loyalty	Transgression	Temptation
Kohlberg:						
III	---	90	38	46	81	74
III'		---	41	44	78	75
Prosocial:						
Helping			---	68	41	30
Loyalty				---	38	30
Antisocial:						
Transgression					---	75
Temptation						---

Table 11
Mean Care and Justice Likert Scores^a

Type of Dilemma	Care	Justice
Kohlberg	6.2 _a (1.3)	4.7 _b (2.0)
Prosocial:		
Helping	5.8 _{ac} (1.7)	3.0 _a (2.2)
Loyalty	5.6 _{bc} (1.6)	2.7 _a (2.2)
Antisocial:		
Transgression	4.5 _d (2.0)	5.0 _b (1.9)
Temptation	2.8 _e (2.4)	5.0 _b (2.1)

^aThe scale ranged from 0-7; the higher the score, the more the care or justice.

Note. Means from the same column that do not share a similar subscript are different at the $p < .005$ level or lower. Standard deviations are in brackets.

Table 12

Mean Ratings as a Function of Type of Dilemma and Question

Type of Dilemma	Question									
	Exp- erience		Signif- icance		Easy to discuss		Difficulty of decision		Moral Concern	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Kohlberg	1.6 _a	1.7	4.6 _a	2.0	2.7	1.8	3.2 _a	1.8	5.9 _a	1.4
Prosocial:										
Helping	3.6 _{bc}	2.3	4.6 _{ab}	1.9	3.2	1.9	4.4 _b	1.9	4.8 _b	1.6
Loyalty	4.1 _b	2.0	4.7 _{ab}	1.9	2.7	1.9	4.7 _b	1.9	4.6 _b	1.5
Antisocial:										
Transgression	3.2 _c	2.0	4.4 _{ab}	1.9	2.8	1.9	4.0 _{ab}	1.9	5.0 _b	1.5
Temptation	3.7 _{bc}	2.0	3.9 _b	2.1	2.7	2.0	3.3 _a	2.0	4.9 _b	1.8

Note. The scale ranged from 0-7. Means from the same column that do not share a similar subscript are different at the $p < .005$ level or lower. M = Mean; SD = Standard deviation.