

Gender Differences in Moral Orientation: A Meta-Analysis

Sara Jaffee and Janet Shibley Hyde
University of Wisconsin—Madison

C. Gilligan's (1982) critique of L. Kohlberg's theory of moral reasoning and her assertion that two modes of moral reasoning (justice and care) exist have been the subject of debate within the field of psychology for more than 15 years. This meta-analysis was conducted to review quantitatively the work on gender differences in moral orientation. The meta-analysis revealed small differences in the care orientation favoring females ($d = -.28$) and small differences in the justice orientation favoring males ($d = .19$). Together, the moderator variables accounted for 16% of the variance in the effect sizes for care reasoning and 17% of the variance in the effect sizes for justice reasoning. These findings do not offer strong support for the claim that the care orientation is used predominantly by women and that the justice orientation is used predominantly by men.

The 1982 publication of Carol Gilligan's *In a Different Voice* marked one of those rare moments when social science research breaches the ivied walls of academia and captures the public's imagination. Gilligan's assertion that females and males speak in different moral voices—a care voice characterized by the desire to maintain relationships and to respond to others' needs and a justice voice characterized by considerations of fairness and equity—resonated with readers' experience (Mednick, 1989). *Ms.* magazine named Gilligan its 1984 Woman of the Year, lauding her for research that “created a new appreciation for a previously uncatalogued female sensibility, as well as possibilities for new understanding between the genders” (Van Gelder, 1984, p. 37). In 1996, *Time* magazine included Gilligan among its “Time 25”—25 innovative Americans with the ability “to show us the world anew, to educate and entertain us, to change the way we think about ourselves and others” (“Time 25,” 1996, p. 54). Wrote the editors,

How likely is it that a single book could change the rules of psychology, change the assumptions of medical research, change the conversation among parents and teachers and developmental professionals about the distinctions between men and women, boys and girls? (p. 66)

More than 15 years after the publication of this influential work, are its theses supported by available scientific data?

Gilligan's (1982) research on female moral reasoning challenged cognitive-developmental stage theories of moral development on two fronts. First, she argued for a broader conceptualization of moral reasoning that encompassed a care orientation as well as a justice orientation. The care orientation is characterized by a focus on maintaining relationships, responding to the needs of others, and a responsibility not to cause hurt. The justice orienta-

tion is concerned with principles of fairness and equity such as those assessed in conventional measures of moral reasoning (Gilligan, 1982). Although acknowledging that males and females could use either a justice or a care perspective, Gilligan asserted that care reasoning was used predominantly by females and justice reasoning was used predominantly by males (Gilligan, 1982; Gilligan & Attanucci, 1988). Second, she argued that Kohlberg's Moral Judgment Interview (MJI; Colby et al., 1987)—the most widely used measure of moral reasoning—was gender biased because it was validated on an all-male sample and because its scoring scheme characterized considerations of care and response as less sophisticated than considerations of justice and fairness.

Over the last 15 years, Gilligan's (1982) work has inspired a wealth of empirical research and criticism, some of which has supported her claims and some of which has not. At the same time, researchers have seldom agreed on how Gilligan's care and justice orientations should be defined or how they should be measured. Consequently, it has been difficult to compare the results of these studies. The goal of this article is to evaluate Gilligan's assertions that (a) there are gender differences in Kohlbergian moral stage and (b) the care and justice orientations are gender related. The former assertion has been reviewed meta-analytically by several investigators, and those results are summarized here. Meta-analysis was used in the present paper to evaluate the latter assertion. In addition, a number of complexities surrounding the study of gender and moral reasoning are reviewed. The following sections present Gilligan's theory of moral reasoning, as well as the criticisms of this work.

Gilligan's Theory of Moral Reasoning

Gilligan's (1977, 1982) theory of moral reasoning stemmed from Kohlberg's cognitive-developmental stage theory of moral development. Kohlberg (1969, 1976, 1984) proposed that individuals progress through a sequence of invariant and universal stages of moral reasoning. These six stages are grouped into three levels, each of which represents a qualitative advance in the individual's ability to understand and integrate diverse points of view (Kohlberg, 1976).

Sara Jaffee and Janet Shibley Hyde, Department of Psychology, University of Wisconsin—Madison.

We would like to acknowledge the contributions of Rose Jadack and MaryBeth Nolan to this project. The Graduate School of the University of Wisconsin provided financial support.

Correspondence concerning this article should be addressed to Janet Shibley Hyde, Department of Psychology, University of Wisconsin—Madison, 1202 W. Johnson Street, Madison, Wisconsin 53706.

Early research on Kohlbergian moral development reported that Stage 3 was the modal stage for females and Stage 4 was the modal stage for males (Fishkin, Keniston, & MacKinnon, 1973; Haan, Smith, & Block, 1968; Holstein, 1969; Kohlberg & Kramer, 1969; Poppen, 1974). Stage 3 reasoning is characterized by the desire to maintain relationships and to meet others' expectations, and Stage 4 is characterized by a law-and-order mentality in which laws are upheld so as to maintain the social order. These findings led some researchers to accuse Kohlberg's theory of gender bias (Gilligan, 1977, 1982; Haan, 1978). Specifically, Gilligan (1977, 1982) argued that because Kohlberg derived his theory of moral development from an all-male sample, he neglected to recognize a distinctively female mode of moral reasoning—one that is characterized by a desire to maintain relationships and a responsibility not to cause hurt. This care orientation contrasts with a distinctively male mode of moral reasoning—the justice orientation—that is based on the abstract principles of justice, fairness, and individualism captured by Kohlberg's MJJ. Unlike Kohlberg's moral stages, the care and justice orientations do not represent cognitive structures that develop in a stagelike sequence. Instead, they represent frameworks that can be modified by experience and through which individuals interpret and resolve moral problems. However, Gilligan argued that the care orientation's emphasis on maintaining relationships led to its classification as a Stage 3 response in Kohlberg's scoring scheme.

Origins and Development of Moral Orientations

Gilligan and colleagues asserted that the care and justice orientations are rooted in early childhood experiences of attachment and inequality that foster, respectively, a relational and an individualistic self-concept (Gilligan & Attanucci, 1988; Gilligan & Wiggins, 1987). Because children are born into a position of inequality and cannot survive without an attachment figure, all children are exposed to the conditions that form the basis of both moral orientations. Thus, all individuals have the capacity to understand, experience, or implement either moral orientation. The reason that males and females come to use one mode of moral reasoning over the other is that these experiences of attachment and inequality are differentially reinforced in a society dichotomized by gender. Working from Chodorow's (1978) neopsychoanalytic account of gender identity development, Gilligan proposed that because women are the primary caretakers in most cultures, girls' self-concepts are based on a definition of themselves as similar to and connected with their mothers whereas boys' self-concepts are rooted in their difference and separation from their mothers. Moreover, boys identify with their fathers, who may be perceived as authority figures who hold power over them. Thus, the experience of attachment and connection is more salient to girls, whereas the experience of inequality and separation is more salient to boys. These initial experiences of attachment and inequality may be confirmed in later childhood and adolescence, resulting in an association between gender and moral orientation (Gilligan & Wiggins, 1987). Gilligan and Wiggins (1987) concluded:

The sex difference question, when framed in this way, does not carry the implication that one sex is morally superior, nor does it imply that moral behavior is biologically determined. Instead, it draws attention to two perspectives on morality. To the extent that biological sex, the psychology of gender, and the cultural norms and values that define

masculine and feminine behavior affect the experience of equality and attachment, these factors presumably will influence moral development. (p. 282)

Empirical Tests of Gilligan's Developmental Theory

These claims about the origins and development of moral voice are largely untested. Benenson, Morash, and Petrakos (1998) observed 41 mother-child dyads (children were 4 and 5 years old) in a play setting and found that, compared with boys, girls were physically closer to their mothers, engaged in more mutual eye contact with their mothers, and were rated higher on global enjoyment. Although the authors cited these data as supportive of Chodorow's (1978) theory, it is not clear whether emotional closeness, as measured in this study, matches Chodorow's concept of parental identification. Moreover, because emotional closeness between fathers and their children was not studied, it is impossible to determine whether the nature of children's relationships with their fathers differs from their relationships with their mothers.

Other researchers have interpreted Gilligan to mean that gender differences in moral orientation stem from parental socialization practices that foster a relational self-concept in females and an individualistic self-concept in males (Lollis, Ross, & Leroux, 1996; Walker, 1997). Lollis and colleagues (1996) observed parents' interventions in property conflicts among 2-4-year-old siblings. They found that mothers intervened more often and used more care reasoning than fathers, although parents did not differ in their use of justice reasoning. However, there was no evidence that girls received more care-oriented interventions nor that boys received more justice-oriented interventions from either parent. Thus, the authors concluded that boys and girls were not being socialized differently in this respect. This argument, however, cited a social learning explanation for the development of gender differences in moral reasoning, whereas Gilligan posited a neopsychoanalytic account in which the origin of gender differences in moral reasoning lies in the child's sense of identification with each parent.

Various researchers have called into question the testability of Gilligan's claims (Walker, 1995), citing the difficulties inherent in measuring constructs like parental identification, equality/inequality, and attachment/detachment. Empirical tests of Gilligan's claims would indeed prove challenging. Such studies might describe how the connected and individuated self-concepts are differentially reinforced for girls and boys and how the self-concept changes over time. They also might elucidate the mechanisms by which parental identification leads to gender differences in the self-concept and, ultimately, to gender differences in moral orientation.

Levels of Care Reasoning

Originally, Gilligan (1982) proposed that the care orientation developed in three phases, the first characterized by an exclusive focus on one's own needs (i.e., caring for oneself only), the second characterized by self-sacrifice and a focus on others' needs, and the third characterized by the ability to balance others' needs and one's own needs. This sequence was derived from her interviews with 29 women, diverse in age, ethnic background, and socioeconomic status (SES), who were considering whether to undergo an

abortion. Gilligan (1982) conducted a follow-up study 1 year after the initial abortion interview. Although many of the women had traversed at least two of the levels, none of the 21 women had progressed across all three, and many showed regression. Gilligan seems to have dropped this developmental sequence in her more recent work (Gilligan, Brown, & Rogers, 1990).

Skoe and Marcia (1991) developed and validated the Ethic of Care Interview (ECI) to assess Gilligan's levels of care reasoning. The ECI is negatively related to measures of authoritarianism and positively related to measures of ego identity, cognitive complexity, role-taking, and MJI scores (Skoe & Diessner, 1994; Skoe, Pratt, Matthew, & Currer, 1996). Although some evidence exists that females score higher than males on the ECI (Skoe & Gooden, 1993; Skoe et al., 1996), this finding has not been replicated consistently (Skoe & Diessner, 1994; Söchting, Skoe, & Marcia, 1994). Unfortunately, the ECI has not been used to answer important questions about the developmental progression of care reasoning or mechanisms of change. Although Skoe and Diessner (1994) reported a positive association between ECI scores and age, this finding was based on cross-sectional data. Skoe et al. (1996) reported moderately stable levels of care reasoning across a 4-year period in mid- to late adulthood, but it would be informative to explore longitudinally the development of care reasoning at earlier ages.

There has been little discussion of how or why individuals move from one level of care to the next. Although Gilligan (1982) suggested that individuals progress through these levels in times of crisis, it is not clear if one reaches the highest level and remains there or if each new crisis causes the cycle to start anew (Sichel, 1985). The latter would suggest that the levels of care reasoning represent a process by which individuals resolve moral problems rather than a developmental sequence in which each level represents an advance in moral maturity. Skoe and Marcia (1991) suggested that advances in moral thought are based on questioning previously held beliefs and formulating new, more inclusive positions, and they compared this to cognitive-developmental processes of disequilibrium and accommodation. However, such a cognitive-developmental account of care reasoning has not been assessed empirically.

Measures of Moral Orientation and Methods of Coding

A number of measures have been developed to assess moral orientation. These generally fall into two categories: interview measures and objective measures. The majority of these measures were designed to assess Gilligan's conceptualization of the care and justice orientations.

Interview Measures

Interview measures usually ask participants to describe a moral dilemma from their own lives. Responses are coded for the presence of justice and care reasoning. A number of coding schemes have been developed. For instance, Lyons's (1982) coding scheme calculates the number of care and justice considerations (each idea the participant presents in discussing the dilemma) and classifies moral orientation as care or justice on the basis of the modal response. Moral orientation is recorded as "split" if the participant voices as many care considerations as justice considerations. Other

researchers have modified this coding scheme to better differentiate care and justice reasoning (Gilligan & Attanucci, 1988; Krebs, Vermeulen, Denton, & Carpendale, 1994). Responses are classified as care only or justice only if all the considerations can be categorized as care- or justice-based. If at least 75% of the considerations can be categorized as care- or justice-based, the response is classified as care focus or justice focus, respectively. If less than 75% of the considerations can be categorized as care- or justice-based, the response is categorized as mixed. Finally, Brown and colleagues (1988) developed a coding system in which a response narrative is read four separate times, each time from a different standpoint. In the first reading, attention is focused on the story being told. In the second reading, all references to the self are underlined. In the third and fourth readings, attention is focused on moral voice: first the care perspective, then the justice perspective. Finally, the reader completes a summary worksheet in which his/her interpretation and summary of the text are substantiated by quotes from the interviews. These summaries allow the reader to code for the presence of care and justice reasoning and the predominance of one or the other moral orientation. Brown et al. maintained that by reading the narrative from many different standpoints, the reader remains open to the possibility that the same statement can be interpreted in different ways.

The Fable Interview (Johnston, 1988) is another interview measure that taps Gilligan's moral orientation construct. However, instead of asking participants to discuss dilemmas from their own lives, participants respond to dilemmas embedded within two of Aesop's fables. The first solution to each fable is considered the spontaneous solution. Respondents are then probed to determine if they can provide an alternative solution to the dilemma. If, for instance, the spontaneous solution is a care solution, respondents are encouraged to offer a justice solution. If they do so, they are then asked to decide which is the better solution. A modified version of Lyons's (1982) coding scheme is used to classify respondents as care, justice, or split reasoners. The unit of analysis in the fable coding is the entire solution offered by the respondent (as opposed to the individual considerations in each response). One advantage of the Fable Interview is that comparisons can be made between individuals because all are responding to the same dilemmas.

Eisenberg and colleagues (Eisenberg, Lennon, & Roth, 1983; Eisenberg-Berg, 1979) and Kohlberg and his colleagues (Colby et al., 1987) have developed coding schemes for individuals' responses to prosocial moral dilemmas and MJI dilemmas, respectively. Eisenberg's scheme calculates the frequency with which an individual uses the various types of prosocial reasoning in responding to a prosocial dilemma. In Kohlberg's scheme, orientation scores are generated on the basis of the content of an individual's response to a dilemma. These interview measures are not meant to tap Gilligan's conceptualization of care or justice reasoning.

Self-Administered Objective Measures

Objective tests of moral reasoning and moral orientation have also been developed. Some of these, including the Care/Response Orientation Scale (CROS; Atunzu, 1986), the Measure of Moral Orientation (MMO; Liddell, 1990), and the Moral Orientation Scale (Yacker & Weinberg, 1990), were designed to assess the

care and justice orientations as conceptualized by Gilligan. Others, such as the Prosocial Moral Reasoning Objective Measure (Carlo, Eisenberg, & Knight, 1992), were designed for other purposes, although they have been used to assess care and/or justice reasoning. In these tests, respondents are asked to generate their own moral dilemma or to respond to a hypothetical one. They are then presented with multiple responses representing examples of care and justice reasoning and asked to pick the one they would use or prefer in resolving the dilemma. These tests produce continuously distributed care and/or justice orientation scores.

Finally, as discussed above, the ECI (Skoe & Marcia, 1991) assesses Gilligan's levels of care reasoning. The ECI consists of four dilemmas administered in a structured interview format. One of these dilemmas is generated by the respondent, and the other three are standard interpersonal dilemmas. The ECI can be scored according to level or according to the total score across the four dilemmas.

Criticisms of the Moral Orientation Construct

Gilligan has been accused of oversimplifying Kohlbergian moral reasoning in her description of the justice orientation (Puka, 1991; Walker, 1989; Walker, de Vries, & Trevethan, 1987). Critics have argued that the justice orientation comes closest to the rigid law-and-order reasoning of Kohlberg's Stage 4, and Gilligan has been criticized for ignoring the contention that justice and rights exist in the context of social responsibilities and obligations. Thus, at the higher stages of moral reasoning, individuals follow rules only if those rules benefit the common good (Kohlberg, Levine, & Hewer, 1983).

Critics have also argued that too little work has been done to characterize or validate the existence of Gilligan's care and justice orientations other than to demonstrate that these modes of moral reasoning are present in individuals' responses to moral dilemmas (Walker, 1989; Walker et al., 1987). Consequently, it is not clear whether moral orientations are used consistently across situations and over time or how moral orientations relate to each other or to moral behavior.

Intraindividual Consistency

Gilligan's assertion that most individuals prefer one mode of reasoning over the other implies a high level of intraindividual consistency in their use (Gilligan & Attanucci, 1988; Langdale, 1986). Indeed, Gilligan and Attanucci (1988) reported that two thirds of their participants focused on only one orientation in their discussion of a self-generated moral dilemma. Langdale (1986) demonstrated that, within a single dilemma, approximately 87% of participants used a single orientation and that the number of individuals with a predominant justice or care orientation (as opposed to mixed orientation) was significantly greater than chance. The number of people who showed a predominant justice or care orientation across different dilemmas was also significantly greater than chance.

Other researchers, however, have been unable to replicate these results (Krebs et al., 1994; Pratt, Golding, & Hunter, 1983; Pratt, Golding, Hunter, & Sampson, 1988; Walker et al., 1987; Wark & Krebs, 1996). For example, Wark and Krebs (1996) reported that only 9% of their college-age participants obtained the same moral

orientation score across three types of dilemmas and only 29% obtained the same or an adjacent score. Walker and colleagues (1987) asserted that consistency would be demonstrated by individuals who used the same orientation 75% of the time or more. Less than 20% of participants met this consistency criterion across self-generated and hypothetical dilemmas. Within a single dilemma, only about half met the criterion. Similarly, Pratt and colleagues (1988) reported that only 60% of their participants used the same orientation across two self-generated dilemmas. In summary, there is little evidence that moral orientations are used consistently within or between dilemmas.

There is mixed evidence regarding the extent to which moral orientations are used consistently over time. Walker (1989) followed children and adults over a 2-year interval and found that half the respondents evidenced a different orientation at the follow-up than at the initial interview. However, Skoe and colleagues (1996) collected two waves of data, 4 years apart, from a sample of middle-aged and elderly adults and reported that care reasoning levels were moderately stable within mid- to late adulthood. It should be noted that these longitudinal analyses addressed different issues. Whereas the Walker analysis addressed the question of whether individuals use different moral orientations over time, the Skoe et al. analysis assessed changes in level of care reasoning over time.

Validity

There is qualified evidence for the construct validity of Gilligan's moral orientations. Pratt, Diessner, Hunsberger, Pancer, and Savoy (1991) reported that the justice and care orientations were associated with variations in the self-concept of the sort described by Gilligan (1982) and Lyons (1983), such that those who had more individuated self-concepts tended to discuss justice-oriented personal dilemmas. However, their results did not clarify whether males were more likely to report an individuated self and females were more likely to report a connected self. Lyons, however, reported that women were more likely than men to report a connected self, men were more likely than women to report an individuated self, and that men and women also differed in predicted ways in their use of the care and justice orientations. Finally, Liddell and colleagues (Liddell, 1998; Liddell, Halpin, & Halpin, 1992) reported that scores on a measure of connected self-concept were positively related to a standardized measure of care reasoning and scores on a measure of a rights-oriented self-concept were positively related to a standardized measure of justice reasoning.

There is limited evidence for the convergent validity of the moral orientation constructs. Liddell (1998) compared the MMO (Liddell, 1990) with semistructured interviews coded according to Lyons's (1982) protocol. Care and justice scores, as measured by the MMO, showed positive but nonsignificant correlations with, respectively, care and justice as measured by the semistructured interview.

Other evidence does not support predictions derived from Gilligan's work. Because Gilligan argued that justice reasoning is favored at the highest stages of Kohlberg's framework, it is expected that the care orientation would be negatively associated and the justice orientation would be positively associated with Kohlbergian moral stage. Although Krebs and colleagues (1994) re-

ported that moral stage scores were negatively correlated with the care orientation for men, Pratt and colleagues (1988) reported a positive relationship between care reasoning and moral stage for women and no association between the two for men. Walker and colleagues (1987) found no relationship between moral stage and moral orientation when the latter was assessed with hypothetical dilemmas but found a positive relationship between the care orientation and moral stage when orientation was assessed with self-generated dilemmas. Moreover, individuals at the highest stages of moral reasoning were more likely to use both care and justice reasoning in their response to dilemmas.

In summary, there is little evidence to support the notion that individuals use a particular moral orientation consistently over time and situations. Although there is some evidence that moral orientations are associated with related constructs and measures, these predicted relationships are not obtained consistently. Finally, there is virtually no evidence regarding the extent to which moral orientations develop over time. This pattern of evidence contrasts sharply with Kohlberg's account of moral reasoning and development, for which there is ample evidence of reliability, validity, and stage sequence. Arguably then, these are not comparable phenomena and, if so, accounts of one need not encompass the other, nor can they be criticized for disregarding the other (Puka, 1991).

What Constitutes Moral Maturity?

As various critics have pointed out, it is not always clear what Gilligan believes is the relationship between the justice and care orientations or what constitutes moral maturity (Auerbach, Blum, Smith, & Williams, 1985; Flanagan & Jackson, 1987; Mason, 1990; Puka, 1991; Sichel, 1985; Walker, 1995). Most commonly, Gilligan has drawn on the metaphor of the ambiguous figure to illustrate her discussion of moral reasoning. The care orientation and the justice orientation are framed as two ways of seeing a moral problem. Choosing to see the problem from one perspective may lead one to neglect the ways in which the problem might be solved from the other perspective (Gilligan & Wiggins, 1987). At times, Gilligan has asserted that these perspectives are incompatible alternatives to one another but are both adequate from a normative point of view (Gilligan, 1986b, 1987). In other places, she has argued that they complement one another and that each is deficient without the other (Gilligan et al., 1990). Thus, it is not clear whether moral maturity is characterized by the ability to integrate and balance the justice and care perspectives, to maintain them in a complementary tension, or whether there is a morally mature care orientation in the absence of a morally mature justice perspective. Notably, her more recent discussions favor the notion that the justice and care perspectives are maintained in complementary tension (Gilligan et al., 1990).

Even granting that moral maturity is characterized by the ability to maintain both the justice and the care perspective in some sort of tension, this formulation does not resolve the question of how individuals should solve moral problems—the question of whether and when either moral voice should take precedence (Sichel, 1985). Although it is possible and, Gilligan might suggest, necessary to perceive moral problems from multiple perspectives, moral action requires choice, and Gilligan's framework offers few guidelines as to which moral voice is more adequate in which situations.

This, however, is an ethical issue and not one that empirical research can address.

Evidence for and Against Gender Differences in Moral Reasoning

Gender Differences in Kohlbergian Moral Stage

Gilligan's claim that the MJI is gender biased has been convincingly debunked. Critical reviews of the moral development literature have failed to find evidence that Kohlberg-based measures yield gender differences in moral reasoning scores. Instead, these reviews have found that gender differences in moral reasoning are small to nonexistent (Rest, 1979; Thoma, 1986; Walker, 1984, 1991, 1995). Rest (1979) conducted a critical review of 17 studies with 20 independent comparisons of male and female participants on the Defining Issues Test (DIT; Rest, 1979) and found only two significant gender differences, both favoring females.

Walker's (1984) critical review and meta-analysis of 79 studies that measured moral reasoning with the MJI found small gender differences favoring males only in adulthood. Sex accounted for only one twentieth of one percent of the variability in moral reasoning development. Walker noted that gender differences appeared most frequently in studies that confounded gender with education or occupational status and in studies that used the earlier, less reliable versions of the Kohlberg scoring manual (see Baumrind, 1986, for a critique of Walker and Walker, 1986b, for a reply; see also Walker, 1991, 1995). Walker (1991) updated this review and obtained similar results.

Thoma (1986) conducted a meta-analysis of gender differences in moral reasoning in 56 samples that were administered the DIT (Rest, 1979) and found a small effect favoring females. Gender differences accounted for less than 0.5% of the variance in DIT scores, whereas age and education accounted for nearly 53% of the variance in DIT scores. Thoma concluded that females are not at a disadvantage when measured by the DIT.

Finally, Colby and Damon (1983) pointed out that Kohlberg's model has now been validated on a sample of males and females. They found that females passed through the same stages in the same order as males (Colby, Kohlberg, Gibbs, & Lieberman, 1983). When occupation and education were controlled, gender differences in moral reasoning level disappeared. Thus, they concluded that Gilligan's allegations of gender bias in Kohlberg's theory are unwarranted.

Gilligan's (1986a) response to critics has been to suggest that a lack of gender differences on the MJI may simply demonstrate that females learn to use the justice orientation as effectively as males. Because the MJI codes only those data that fit within its stage definitions, considerations of care may still be ignored. Although this is a valid point, it begs the question of what sort of data would disconfirm Gilligan's hypothesis.

It is important to point out that the studies cited above do not test Gilligan's assertion that males are more likely than females to use the justice orientation. The research cited above demonstrates that there are minimal gender differences in justice reasoning stage, but does not speak to the question of whether there are gender differences in the use of the justice orientation.

Gender Differences in Moral Orientation

Although the extant literature fails to find gender differences in stage scores on Kohlberg-based measures, the research on gender differences in moral orientation is less conclusive. Most researchers have now acknowledged that more than one mode of moral reasoning exists (Kohlberg et al., 1983), but there is considerable controversy as to whether those moral orientations can be reliably associated with gender. Whereas some researchers have found evidence for Gilligan's claim that care reasoning is used predominantly by females and justice reasoning predominantly by males (Gilligan & Attanucci, 1988; Johnston, 1988; Yacker & Weinberg, 1990), other researchers have found gender differences in care reasoning only (Galotti, Kozberg, & Farmer, 1991; Garmon, Basinger, Gregg, & Gibbs, 1996; Gibbs, Arnold, & Burkhart, 1984; Liddell, Halpin, & Halpin, 1993; Wark & Krebs, 1996), and still others have failed to find gender differences in the use of either care or justice reasoning (Beal, Garrod, Ruben, & Stewart, 1997; Friedman, Robinson, & Friedman, 1987; Walker et al., 1987). Moreover, many researchers have found evidence that gender differences in moral orientation are moderated by other variables, such as dilemma content (Walker et al., 1987; Wark & Krebs, 1996, 1997) and social class (Beal et al., 1997; Puka, 1989; Tronto, 1987).

Given the widely disparate findings regarding gender differences in the use of justice and care reasoning, there is a need for a review of this body of research. Meta-analysis provides an appropriate tool for synthesizing the research on gender differences in moral orientation and for determining the direction and the actual magnitude of the effect. Importantly, meta-analysis also allows the researcher to examine how the size of the effect might be moderated by other variables. We hypothesized that the effect size for gender differences in moral orientation might be moderated by the following variables, each of which is addressed in turn: (a) age, (b) SES, (c) moral orientation construct, (d) type of dilemma, (e) coding scheme, (f) scale, (g) gender of the protagonist, and (h) publication status. Although Walker (1995) provided an excellent narrative review of some of these moderators, he did not subject the studies to quantitative meta-analysis.

Moderator Variables

Age

Gilligan and colleagues (Gilligan, 1982; Gilligan & Wiggins, 1987) proposed that gender differences in moral orientation emerge in early childhood and persist across the life course. However, some cross-sectional research has suggested that gender differences in moral orientation are moderated by age. Walker and colleagues (Walker, 1989; Walker et al., 1987) reported gender differences in care reasoning favoring women among adults who discussed self-generated dilemmas but failed to find such gender differences among 1st-, 4th-, 7th-, and 10th-grade children. Similarly, in a study of young, middle, and older adults, Pratt and colleagues (1988) reported that gender differences in moral orientation were significant only for the middle-age group. Pratt and colleagues proposed that family organization during the period of active parenting (i.e., middle age) leads to increased gender-role polarization (Guttman, 1985) and thus heightens gender differences in moral orientation. Consistent with this hypothesis, they

found that mothers were significantly less justice-oriented than fathers, but this gender difference failed to emerge among nonparents. However, other cross-sectional studies have failed to find evidence that gender differences in moral orientation are moderated by age (Craft, 1992; Galotti et al., 1991; Garrod, Beal, & Shin, 1990; Langdale, 1986; Pratt et al., 1991).

In a longitudinal study of prosocial moral reasoning that followed a cohort of children from age 4.5 to age 20, Eisenberg and colleagues (1987) reported that gender differences in sympathetic and role-taking reasoning, both of which have been likened to Gilligan's care orientation, emerged at approximately age 11 or 12. However, Walker (1989) failed to find evidence among children of gender differences in moral orientation at either of two assessments in a short-term longitudinal study. Gender differences in moral orientation emerged at both time points for adults who described dilemmas from their own lives. Meta-analysis is needed to sort out the inconsistencies in these cross-sectional and longitudinal findings.

Socioeconomic Status

Some researchers have proposed that gender and subordinate or minority status are confounded in studies of moral orientation (Puka, 1989; Tronto, 1987). They have argued that subordinate status and the condition of powerlessness promote an inherent concern with others because those others, in part, determine one's outcomes (Tronto, 1987). However, other researchers have argued that low social status promotes a concern with fairness and rights because these rectify the social inequalities experienced by minority or subordinate groups (Beal, et al., 1997). Garrod and Beal (1993) found that children from a rural, working-class community were somewhat more likely to emphasize rights considerations than were children from a more affluent university suburb. Similarly, Gilligan and Attanucci (1988) reported that minority students were more likely than nonminority students to adopt a rights perspective.

These hypotheses might be interpreted to mean that SES exerts a main effect on moral reasoning, in which case gender differences in both care and justice reasoning should be smallest for lower class groups. Alternatively, they might be interpreted to mean that gender and social class interact, such that gender differences in care reasoning (but not justice reasoning) should be greatest among lower class groups because low-SES females will be more likely to use the care orientation as a function of both their gender and their social class.

Moral Orientation Construct

Most measures of moral orientation base their definition of the care and justice orientations on Gilligan's description of these constructs. However, some researchers have taken a different approach to measuring moral orientation. For instance, Eisenberg and her colleagues (Eisenberg et al., 1983; Eisenberg-Berg, 1979) have focused on prosocial moral reasoning, which is defined as "reasoning about conflicts in which the individual must choose between satisfying his or her own wants or needs and those of others in contexts in which laws, punishments, authorities, formal obligations, and other external criteria are irrelevant or deemphasized" (Eisenberg-Berg, 1979, p. 128). Eisenberg and her col-

leagues have characterized types of prosocial moral reasoning, several of which (e.g., concern with other's physical, material, or psychological needs; role-taking; sympathetic orientation) load on a single factor that represents other-oriented reasoning and are conceptually similar to Gilligan's description of the care orientation (Eisenberg, Fabes, & Shea, 1989).

Other researchers claim that moral orientation can be measured at the aspect level or the orientation level in Kohlberg's scoring scheme (de Vries & Walker, 1986; Garmon et al., 1996). Unlike Kohlberg's stages, these moral orientations are more closely related to the content rather than the structure of moral reasoning. It has been argued that the perfectionism orientation (emphasizing dignity and autonomy, good conscience and motives, and harmony with self and others) and the utilitarian orientation (emphasizing welfare or happiness consequences for oneself and others) are consistent with the care orientation, whereas the normative orientation (emphasizing duties and rights) and the fairness orientation (emphasizing justice) are consistent with the justice orientation (Walker, 1986a, 1995). Other researchers have focused on the aspects that characterize each stage of moral reasoning (Garmon et al., 1996). Specifically, they have pointed to similarities between empathic role-taking, which is a Stage 3 aspect that is characterized by empathic references to another's psychological or emotional welfare, and the care orientation.

There is some disagreement as to whether Kohlberg's orientations are adequate measures of the care orientation. Smetana (1984) pointed out that stage and aspect scores are confounded in Kohlberg's scoring system. For instance, the perfectionist orientation can be scored only at the higher levels of moral reasoning. Nevertheless, because a number of studies have assessed gender differences in care and justice reasoning using Kohlberg's moral orientations, these studies were included in the meta-analysis.

Type of Dilemma

Some researchers have suggested that gender differences in the use of care and justice reasoning might be accounted for by the kinds of dilemmas used to elicit moral reasoning (Clopton & Sorell, 1993; Ford & Lowery, 1986; Gilligan, 1982; Langdale, 1986; Mednick, 1989; Walker, 1989; Walker et al., 1987; Wark & Krebs, 1996, 1997). One possibility is that when dilemmas are self-generated, females discuss dilemmas that deal with personal concerns (e.g., dilemmas dealing with relationships between the participant and close others) and males discuss impersonal dilemmas (e.g., conflicts between others or involving generalized others, such as clients or students). Because personal dilemmas are inherently concerned with relationships, they may elicit care reasoning, whereas impersonal dilemmas may elicit considerations of fairness and reciprocity (Walker, 1989; Walker et al., 1987). Indeed, several researchers have found that women generate more personal real-life dilemmas than men do and men generate more impersonal real-life dilemmas (Pratt et al., 1988, 1991; Skoe & Diessner, 1994; Skoe et al., 1996; Walker et al., 1987; Wark & Krebs, 1996, 1997). When type of dilemma is controlled, gender differences in moral orientation are eliminated, suggesting the important role of situational factors in moral orientations (Walker, 1989; Walker et al., 1987; Wark & Krebs, 1996). Consistent with this possibility, Wark and Krebs (1997) reported that prosocial dilemmas evoked more care-based moral judgments than Kohlberg dilemmas or

dilemmas dealing with transgressions or rule violations and these, in turn, evoked more justice-based moral judgments than prosocial dilemmas did. Other researchers, however, have found no evidence of gender differences in dilemma content (Ford & Lowery, 1986; Peter & Gallop, 1994). Given these conflicting findings, meta-analysis is needed to determine whether there are systematic trends in research outcomes.

Coding Scheme

As discussed in the section on measures of moral orientation, several coding schemes have been developed to code for justice and care reasoning in individuals' responses to moral dilemmas. These schemes differ on the basis of the unit of analysis (e.g., considerations vs. entire solution) and whether they are objective or interpretive schemes. Given this heterogeneity, coding scheme was included as a moderator variable.¹

Scale

Measures of moral reasoning can also be distinguished by their scale of measurement. In the coding schemes mentioned above, moral orientation can be analyzed as a categorical variable (i.e., respondents can be classified as having a care or justice orientation) or as a continuous variable (e.g., a "percent care" score is computed for each participant on the basis of the percentage of care-based considerations in each response). Some continuous measures rate respondents along a care continuum and a justice continuum, yielding independent measures of each orientation.

Because the effect sizes that form the basis of a meta-analysis are computed with means and standard deviations, moral orientation measures that yield continuously scaled scores may provide more reliable effect size estimates than measures that describe the proportion of respondents who fall into one or the other moral orientation category. We wanted to be able to estimate the magnitude of the effect size for gender differences in moral orientation when only the most reliable estimates of the effect size were included in the analysis, as well as when all effect sizes were included (this issue is elaborated in the Results section). Consequently, scale was included as a moderator variable.

Gender of the Dilemma's Protagonist

Most of the research on gender of the protagonist has dealt with its role as a potential moderator of gender differences in moral stage. These studies have addressed the possibility that women may score lower on Kohlberg's MJI (Colby et al., 1987) because they are unable to identify completely with the male protagonists of the dilemmas (Holstein, 1976; however, it is important to note

¹ It is important to note the distinction between coding scheme and moral orientation measure. The same measure (e.g., the MJI) might be scored according to any of several different coding schemes. The moderator analyses do not explore how the magnitude of the gender difference in moral orientation differs as a function of the measures themselves (e.g., the Fable Interview or the MMO) because of heterogeneity in the coding schemes used to score the same measure across studies and because there were too few of certain measures (e.g., the CROS) to allow for meaningful interpretation of the moderator analyses.

that not all the MJI dilemmas feature male protagonists). Such studies have provided inconsistent evidence that gender of the protagonist has an effect on moral judgment level (Bussey & Maughan, 1982; Freeman & Giebink, 1979; Garwood, Levine, & Ewing, 1980; Krebs et al., 1994; Lonky, Roodin, & Rybash, 1988; Orchowsky & Jenkins, 1979; Turiel, 1976).

Relatively few studies have examined the extent to which the gender of the protagonist moderates gender differences in moral orientation. Albrecht (1989) found that dilemmas with male protagonists elicited significantly more justice reasoning than dilemmas with female protagonists. However, Beal et al. (1997) reported a preference for care reasoning regardless of the protagonist's gender. Similarly, Krebs et al. (1994) failed to find differences in moral orientation depending on whether participants responded to Kohlberg's dilemmas from their own perspective or from a third-person perspective. Because of these inconsistencies, gender of the protagonist was included as a moderator variable.

Publication Status

Because of a bias toward publishing significant findings, it is possible that an overall effect size based only on published data would overestimate the population effect size for gender differences in moral orientation (Rosenthal, 1979). To address this possibility, unpublished studies were included in the meta-analysis, and publication status was included as a moderator variable.

Goals

The overall goal of this meta-analysis was to determine whether the much-heralded claims of Gilligan are supported by empirical evidence. To this end, we evaluated Gilligan's (1982) assertion that there are gender differences in the use of care and justice reasoning. Furthermore, we investigated whether potential gender differences in moral orientation might be moderated by other variables.

Method

Sample of Studies

The sample of studies came from two sources: (a) The PsycLIT, ERIC, and Dissertation Abstracts computerized databases were searched simultaneously for the years 1966–1998, using the key words (care reason*) or (care orient*) or (ethic* of care) or (prosocial moral*) or (moral orient*), which yielded 741 citations, and (b) the reference sections of those studies that were drawn from the databases and included in the meta-analysis were searched for citations that did not appear in the database search. This strategy yielded an additional 16 citations. It is important to note that the computerized database search did not use the search term (gender and care reason*) as this might have led to a selective sampling of studies that found significant gender differences.

In the case of computerized literature searches, abstracts were inspected and included if they met the following criteria: (a) The study was empirical, (b) there were at least five males and five females in each sample, (c) the study did not report data that had already been reported elsewhere, (d) the measure assessed moral orientation, and (e) age and gender were not confounded (e.g., gender differences in moral reasoning between mothers and sons).

Of the 757 citations elicited by the database and reference section searches, 180 met inclusion criteria. Although this may seem like a small proportion of the total number of citations, it is important to underscore that our search term resulted in a very large number of citations that were nonempirical commentaries on Carol Gilligan's work or citations regarding the ethic of care in nursing. Of the PsycLIT and ERIC studies, 25 were not empirical, 2 included data that were reported elsewhere, 371 did not assess moral orientation, 10 reported on same-sex samples, and 37 were duplicates (i.e., the same study appeared in both the PsycLIT and ERIC databases). Of the dissertations, 132 were excluded because they did not assess moral orientation or because they were nonempirical.

In the second phase of the search, copies of the 114 papers and 66 dissertations were obtained to ensure that they met inclusion criteria and included enough information to compute an effect size. In those cases in which an author reported that moral orientation had been measured but failed to report enough statistical information to compute an effect size, a letter or e-mail was sent to the author at the address specified for reprints or at a more recent address found in the *American Psychological Association 1997 Membership Directory* or in the *Society for Research in Child Development 1996 Directory of Members*.² There were 29 cases in which this was necessary, and we received 23 responses to these requests for additional data, 11 of which included the requested data. In the 18 cases in which no data were available (12 cases in which the author responded that data were not available plus 6 cases in which the author did not respond at all), effect sizes were estimated as zero for 10 of the studies on the basis of the fact that the author reported a nonsignificant gender difference in moral orientation.

Of the 180 studies and dissertations that were deemed eligible on the basis of the abstracts, 113 yielded enough information to compute an effect size or met inclusion criteria once a copy of the study itself had been obtained and examined. This sample comprised 70 published and 43 unpublished studies. Of the 67 studies that were excluded, 7 could not be located, 2 used same-sex samples, 9 did not assess moral orientation, 1 was not empirical, 17 were unpublished studies that did not include enough information to compute an effect size, 8 were published studies for which an effect size could not be computed or estimated even after contacting the author, 15 were studies in which the data were reported elsewhere, and in one case, there was not enough information to compute an effect size and the author's address could not be located. In addition, 7 dissertations were unavailable from UMI.

It is possible to obtain several independent effect sizes from a single article if, for example, data from several age groups are reported (e.g., in a cross-sectional design). These groups can be regarded as separate samples (L. V. Hedges, personal communication, 1987). The result was 113 usable sources, yielding 160 independent effect sizes for gender differences in care orientation and 95 independent effect sizes for gender differences in justice orientation. In the case of the care orientation analysis, this represented the testing of 5,783 males and 6,654 females. In the case of the justice orientation analysis, this represented the testing of 3,831 males and 4,307 females. This compares favorably with other reviews of gender differences in moral reasoning, namely, Walker's (1984) critical review of gender differences in moral stage in which he analyzed gender differences from 66 studies involving 6,780 participants.

Five studies that used the ECI and measured levels of care reasoning were analyzed separately because the level at which one reasons about a given dilemma is conceptually distinct from the amount of care or justice reasoning used in solving the dilemma.

² Letters were mailed only to the authors of published studies as it was too difficult to locate the authors of dissertations.

Coding the Studies

For each study, the following information was recorded: (a) all statistics on gender differences in moral orientation, including means and standard deviations or *t*, *F*, or *r* and (b) the number of male and female participants. The following potential moderator variables were also coded. It is important to note that some of the categories within each moderator variable were collapsed in the analyses because of insufficient numbers of effect sizes per category.

Age. The age(s) of the participants were recorded. If the article reported no age but reported that participants were undergraduates or students in an introductory college course, age was set equal to 19. If a grade level was reported, 5 years were added to that level to yield the age (e.g., third graders were recorded as 8-year-olds). If the sample comprised nonadults and the age range exceeded 10 years, the sample age was recorded as mixed. Otherwise, the midpoint of the age ranges was recorded.

Socioeconomic status. The SES of the participants was recorded as lower class, lower middle class, middle class, upper middle class, upper class, or unreported or mixed, on the basis of how participants were classified in the studies themselves. SES for university students and faculty was estimated as middle class.

Moral orientation construct. We recorded whether researchers measured moral orientation as conceptualized by Gilligan (1982), by Kohlberg (Colby et al., 1987), or by Eisenberg (Eisenberg, et al., 1983; Eisenberg-Berg, 1979).³

Type of dilemma. Dilemmas were coded as *standard hypothetical dilemmas* if they were hypothetical dilemmas from traditional measures of moral reasoning (e.g., the Heinz dilemma) or if they were dilemmas that were likely to be unfamiliar to the participants or not relevant to their everyday lives. Hypothetical dilemmas that involved issues of direct relevance to participants (e.g., dilemmas about cheating that were administered to university students) were coded as *real-life hypothetical dilemmas*. These were dilemmas that participants might expect to face at some point in their lives. Dilemmas that the participants generated from their own lives were coded as *self-generated dilemmas*. Finally, measures in which dilemmas were not used or participants were asked to rate how much they used care or justice reasoning to solve moral problems in the absence of any specific dilemma were coded as *no dilemma*. Hypothetical dilemmas (standard or real-life) that were designed to elicit care reasoning were coded as *hypothetical care dilemmas*, and hypothetical dilemmas (standard or real-life) that were designed to elicit justice reasoning were coded as *hypothetical justice dilemmas*. For instance, Eisenberg's prosocial moral reasoning dilemmas (Eisenberg et al., 1983; Eisenberg-Berg, 1979) were coded as hypothetical care dilemmas because the issue of caring is salient and is in conflict with self-interest or responsibility to the self. Finally, self-generated dilemmas were coded as *self-generated care* or *self-generated justice* if they were designed to elicit care or justice reasoning, respectively.

Coding scheme. We recorded whether moral orientation was scored according to (a) the coding scheme developed by Brown and colleagues (Brown et al., 1988), (b) Lyons's coding scheme (Lyons, 1982), (c) a modified version of Lyons's coding, (d) Kohlberg's coding scheme (Colby et al., 1987), (e) Eisenberg's coding scheme (Eisenberg et al., 1983; Eisenberg-Berg, 1979), (f) close-ended objective responses, or (g) any other coding scheme.

Scale. It was also noted whether these coding schemes yielded continuous care and/or justice scores or whether participants were categorized as using a care or justice orientation on the basis of their responses to the dilemmas.

Gender of the protagonist. The protagonists of the various dilemmas were male, female, or the same sex as the participant, or the dilemmas included both male and female protagonists. It was also noted if the participant was the protagonist of the dilemma (e.g., in self-generated dilemmas), if there was no protagonist (e.g., cases where participants were

asked to rate the extent to which they used care and justice reasoning to solve moral dilemmas), or if the gender of the protagonist was not indicated in the dilemma.

Publication status. It was noted whether studies were published or unpublished.

Interrater Agreement

All studies were coded by Sara Jaffee. A random sample of 30% of the studies was rated by Janet Hyde to obtain estimates of interrater agreement. Interrater agreement was computed for ratings of the eight moderator variables (age, SES, moral orientation construct, type of dilemma, coding scheme, scale, gender of protagonist, and publication status) and ranged from 85% to 100%. Discrepancies were resolved by discussion and consensus.

Computation of Effect Size

The effect size computed was *d*, defined as the mean for males minus the mean for females, divided by the mean within-sex standard deviation (Hedges & Becker, 1986). Thus, positive values of *d* indicate that males used more justice or care reasoning than females, and negative values of *d* indicate that females used more justice or care reasoning than males.

Depending on the statistics available for a given study, formulas provided by Hedges and Becker (1986) and Morris and DeShon (1997) were used for the computation of *d* and the homogeneity statistics. The exception was studies that reported the proportion of males and females who were classified as care reasoners, justice reasoners, or some combination of the two (e.g., studies that used Lyons coding or followed the coding scheme of Brown et al., 1988). In these cases, chi-square statistics were used to test for the association of gender and moral orientation. However, chi-square statistics are not easily translated into effect sizes. So that we could include these studies in the meta-analysis, the moral orientation classifications were converted into percent care scores. Care reasoners were assigned a percent care score of 75%, justice reasoners were assigned a percent care score of 25%, and mixed justice and care reasoners were assigned a percent care score of 50%. This strategy allowed us to derive an effect size from the mean percent care scores. These percent care cutoffs roughly match those used by researchers who have modified Lyons's (1982) coding scheme (Gilligan & Attanucci, 1988; Krebs et al., 1994). This strategy has also been used before by Krebs and colleagues (1994) to create a continuous measure of care reasoning. Because the studies were coded according to whether they yielded a continuous score or a categorical score that needed to be converted by means of the strategy described above, we were able to test explicitly whether the magnitude of the effect size for gender differences in care orientation differed as a function of our own coding strategy.

In 39 of the independent samples that measured care reasoning and in 15 of the independent samples that measured justice reasoning, the author did not include enough information to compute an effect size and did not or could not respond to written solicitations for additional data. In all these cases, the authors reported that the gender difference in moral orientation was nonsignificant. In these cases, the effect size *d* was estimated as zero. Consequently, analyses were conducted twice, first with the entire sample of effect sizes and then with the subsample of effect sizes estimated as zero excluded.

³ Prosocial moral reasoning is scored at a number of levels. Because young children generally do not use higher levels of moral reasoning, gender differences in needs-oriented reasoning were computed for the prosocial moral reasoning studies of children (11 years or younger). Gender differences in Level 4 reasoning (sympathetic, role-taking, positive or negative affect regarding the consequences to others) were computed for the prosocial moral reasoning studies of adolescents and young adults.

A random sample of 30% of the effect sizes was computed independently by Sara Jaffee and Janet Hyde. There were discrepancies in 5% of the d values; these were resolved. All values of d were corrected for bias in estimation of the population effect size, using the formula provided by Hedges (1981). The complete listing of all studies, with effect sizes and moderator variable codes, is provided in Table 1.

Results

The results are divided into three sections. The first reports the results for a meta-analysis of gender differences in the care orientation. The second reports the results for a meta-analysis of gender differences in the justice orientation. The third reports the results for a meta-analysis of gender differences in level of care reasoning as measured by the ECI (Skoe & Marcia, 1991). Whereas the first two meta-analyses assess how much males and females differ in the extent to which they used or endorsed the care and justice orientations respectively, the third analysis examines whether gender differences exist in the level of care reasoning attained by male and female participants.

Magnitude of Gender Differences in Care Orientation

For 73% of the 160 independent samples, the gender difference in care reasoning as reported in the study was not statistically significant. A power analysis revealed that a sample size of 31 would be required to detect a moderately sized effect of .50 for gender differences in care reasoning with $\alpha = .05$. Seventy-three percent of the independent samples included samples sizes of that magnitude. Thus, it is unlikely that the nonsignificant gender differences in care reasoning were the result of low power to detect differences.

The overall effect size d for gender differences in care orientation was -.28, indicating a small gender difference favoring females. A homogeneity analysis using procedures specified by Hedges and Becker (1986) indicated that the effect sizes were nonhomogeneous, $H_B = 438.43$, $p < .001$. Therefore, analyses were conducted to determine the extent to which the magnitude of the effect size was moderated by other variables. The results of the moderator analyses are presented in Table 2.

Age. For purposes of the moderator analysis, participants were categorized into five age groups. Participants who were 11 years old or less were classified as children, and participants who were older than 11 years and less than or equal to 19 years (but were not university students) were classified as adolescents. University students were coded as such. Participants who were aged 20 to 49 (but were not university students) were classified as younger adults, and participants who were 50 years or older were classified as older adults. In cases where the sample included nonadults and the reported age range exceeded 10 years, participants were classified as mixed age.

For all age groups, the effect size indicated that females used more care reasoning than males. However, the magnitude of the effect size differed among age groups, $H_B = 70.91$, $p < .001$. For children and university students, the overall effect sizes indicated a small gender difference ($d = -.08$ and $d = -.18$, respectively), but for adolescents, the effect was moderate in size ($d = -.53$). For younger adults, the overall effect size d was -.33. The small number of samples within the older adult and mixed-age groups suggests that these effect sizes should be interpreted with caution.

Socioeconomic status. As there were relatively few effect sizes at certain levels of the SES variable, SES was collapsed into four categories: lower class (comprising lower class and lower middle class), middle class (comprising middle class and upper middle class), upper class, and mixed/unreported. The homogeneity analysis revealed that the magnitude of the effect sizes differed significantly as a function of SES category, $H_B = 9.05$, $p < .05$. Although females consistently scored higher than males on care reasoning across the SES groups, the magnitude of the effect size increased from -.08 to -.42 as SES increased from the lower-class group to the upper-class group.

Moral orientation construct. The homogeneity analysis revealed that the magnitude of the effect sizes differed significantly as a function of the moral orientation construct, $H_B = 17.29$, $p < .001$. The magnitude of the effect size d was small to moderate in size for studies that defined the care orientation according to Gilligan's description or defined it as the orientation or aspect level in Kohlberg's scheme (-.32 and -.25, respectively). However, for studies that conceptualized care reasoning as prosocial moral reasoning, the magnitude of the effect size was close to zero.

Type of dilemma. As there were relatively few of certain types of dilemmas, these types were collapsed into six categories: (a) standard hypothetical dilemmas, (b) real-life hypothetical dilemmas, (c) self-generated dilemmas, (d) moral reasoning measures that did not contain a dilemma, (e) dilemmas meant to elicit care reasoning (both hypothetical and self-generated), and (f) dilemmas meant to elicit justice reasoning (both hypothetical and self-generated). The homogeneity analysis revealed that the magnitude of the effect sizes differed significantly as a function of type of dilemma, $H_B = 63.07$, $p < .001$.

For both standard and real-life hypothetical dilemmas, the magnitude of the effect sizes was small and indicated that females used more care reasoning than males (-.19 and -.20, respectively). Similarly, the effect sizes for dilemmas designed to elicit care reasoning and dilemmas designed to elicit justice reasoning were -.17 and -.18, respectively. The magnitude of the effect size for self-generated dilemmas was moderate in size ($d = -.37$), and measures that did not include a dilemma yielded the largest difference favoring females ($d = -.57$).

Coding scheme. The homogeneity analysis revealed that the magnitude of the effect sizes differed significantly as a function of coding scheme, $H_B = 61.49$, $p < .01$. The effect size was close to zero for studies that used Eisenberg's (Eisenberg et al., 1983; Eisenberg-Berg, 1979) coding scheme, and the largest effect size ($d = -.61$) was noted for studies that used a modified version of the Lyons coding scheme (Lyons, 1982). Other measures produced effect sizes that were small to moderate in magnitude and favored females.

Scale. The homogeneity analysis revealed that the magnitude of the effect sizes differed significantly as a function of scale, $H_B = 7.46$, $p < .01$. For the studies that used continuously scaled measures, the magnitude of the effect size d was -.26. For studies that used categorical measures, the magnitude of the effect d was -.38, indicating a somewhat larger difference favoring females.

Gender of the protagonist. Because there were only two studies in which the protagonists were all female and one study in which the gender of the protagonist was not specified, these were not included in the moderator analysis. Homogeneity analyses revealed that the magnitude of the effect sizes differed signifi-

cantly as a function of the protagonist's gender, $H_B = 84.88, p < .001$. When the protagonist was male or when the participant was the protagonist, the effect size was small to moderate in magnitude and favored females ($d = -.36$ and $d = -.27$, respectively). When there were both male and female protagonists, however, the gender difference in care reasoning was virtually nonexistent ($d = -.03$). Similarly, when the protagonist was the same gender as the participant, the magnitude of the effect size indicated a small difference ($d = -.11$). However, in cases where there was no protagonist, the magnitude of the effect size d was $-.57$.

Publication status. The magnitude of the effect size for gender differences in care orientation did not differ significantly as a function of whether or not the study was published, $H_B = .13, ns$.

Regression Analysis for Care Orientation

Because the overall homogeneity of variance analysis indicated that the effect sizes were nonhomogeneous, a multiple regression analysis was conducted to determine how much of the variance in effect sizes was accounted for by the moderator variables. Moreover, because many of the moderator variables are highly correlated, it is important to control for the effect of other moderators when determining the extent to which any given one accounts for variance in the magnitude of the effect sizes. The corrected effect size for gender differences in care reasoning was the criterion variable (Hedges, 1981). The eight moderator variables (age, SES, moral orientation construct, type of dilemma, coding scheme, scale, gender of protagonist, and publication status) were entered in stepwise fashion. Because of their categorical nature, all variables except age and SES were contrast coded, and the contrasts are specified in Table 3.

Overall, the moderator variables accounted for a significant 16% of the variance in the effect sizes for gender differences in care reasoning, $F = 7.25, p < .001$. The effect of neither age nor SES was significant. Two contrasts were specified to determine whether moral orientation construct was a significant predictor of the magnitude of the effect sizes for gender differences in care reasoning. The first contrast compared studies that used Eisenberg's construct of prosocial moral reasoning with those that characterized care reasoning according to Gilligan's or Kohlberg's conceptualization of the care construct. Because prosocial dilemmas are designed to elicit care reasoning, it was expected that the magnitude of the effect size for gender differences in care reasoning would be smaller for this group compared with the other two. However, the regression analysis revealed that this contrast was not a significant predictor of the magnitude of the effect size for gender differences in care reasoning. A second contrast was constructed to compare studies that defined care reasoning according to Gilligan with those that defined care reasoning at the element or aspect level in Kohlberg's scheme. It was predicted that the magnitude of the effect size for gender differences in care reasoning would be greater for the former group because, according to Gilligan (1982), Kohlberg's scoring scheme does not adequately capture the care orientation and encourages use of the justice orientation instead. The regression analysis revealed that this contrast was not a significant predictor of the magnitude of the effect size for gender differences in care reasoning.

Two contrasts were specified to determine whether type of dilemma was a significant predictor of the magnitude of the effect

sizes for gender differences in care reasoning. The first contrast compared self-generated dilemmas with those that effectively standardized the dilemma content for participants (i.e., standard hypothetical dilemmas, real-life hypothetical dilemmas, and dilemmas designed specifically to elicit care or justice reasoning). This contrast tested the hypothesis that the magnitude of the effect size for gender differences in care reasoning would be greater when participants were allowed to generate their own dilemmas. This contrast was a significant predictor of the magnitude of the effect sizes for gender differences in care reasoning, $\beta = -.21, p < .05$. A second contrast compared studies that did not include a dilemma with those that did to explore whether the presence of a dilemma accounted for variation in the magnitude of the effect size. This contrast was also significant, $\beta = -.29, p < .001$.

Two contrasts were specified to determine whether coding scheme was a significant predictor of the magnitude of the effect sizes for gender differences in care reasoning. The first contrast compared measures that are scored objectively with those that are not and hypothesized that the magnitude of the effect size would be smaller for objectively scored studies. A second contrast compared coding schemes in which responses are matched to criterion responses (see, e.g., Colby et al., 1987; Eisenberg et al., 1983; Lyons, 1983) with those in which interpretive techniques are used (see, e.g., Brown et al., 1988; modified Lyons, 1983). Neither contrast was significant.

The contrast comparing studies that yielded a continuous moral orientation score with studies that yielded a categorical moral orientation score was significant, $\beta = .18, p < .05$.

Two contrasts were specified to test the effect of gender of the protagonist on the magnitude of the effect sizes for gender differences in care reasoning. The first contrast compared studies in which the participant served as the protagonist in the dilemma with those in which there was a fictional protagonist. This contrast tested whether a difference in the perspective from which the participant responded to the dilemma influenced the magnitude of the effect size. A second contrast was specified to compare studies that used dilemmas with male protagonists with studies that used dilemmas in which the protagonists were not exclusively male or were the same gender as the respondent. Neither contrast was a significant predictor of the magnitude of the effect size for gender differences in care reasoning.

Finally, the regression analysis revealed that publication status did not have a significant effect on the magnitude of the effect size for gender differences in care reasoning.

Because the moderator variables were highly correlated, squared semipartial regression coefficients were computed to determine how much of the variance in the magnitude of the effect sizes was uniquely accounted for by each of the moderator variables when the other moderators were controlled. Table 3 indicates that the contrast comparing self-generated dilemmas with standard hypothetical dilemmas, real-life hypothetical dilemmas, or dilemmas designed to elicit care or justice reasoning accounted for 4% of the variance in the effect sizes for care reasoning. The contrast comparing studies in which there was no dilemma with those in which there was a dilemma uniquely accounted for 8% of the variance in the effect sizes for care reasoning. Finally, the contrast comparing studies that yielded categorical versus continuous outcomes uniquely accounted for 3% of the variance in the effect sizes for care reasoning.

Table 1
Effect Size Estimates and Moderator Variable Codes

Study	N males	N females	d_c	d_j	Age	SES	Construct	Type of dilemma	Coding	Scale	Protagonist gender	Published
Abaris (1990)	69	112	.14	.04	3	3	1	2	6	1	5	0
Abide (1994)	29	59	-.02	-.12	3	3	1	2	6	1	5	0
Abide (1994)	46	100	-.40	-.30	3	3	1	2	6	1	5	0
Akman (1991)	19	19	.33	-.67	3	3	1	1	1	0	2	0
Albrecht (1989)	56	68	.18	.18	3	3	1	1	2	1	3	0
Arvizu (1995)	28	67	-.33	-.24	3	3	1	2	6	1	5	1
Atunzu (1986)	92	110	-.28	.4	4	4	1	4	6	1	4	0
Barnett, Quackenbush, & Sinisi (1995)	244	273	-.29	.03	3	3	1	3	6	1	5	1
Barnett, Quackenbush, & Sinisi (1995)	81	73	-.16	.10	3	3	1	3	6	1	5	1
Beal, Garrod, Ruben, & Stewart (1997)	32	32	0 ^a		1	2	1	1	3	1	3	1
Beal, Garrod, Ruben, & Stewart (1997)	24	24	0 ^a		1	2	1	1	3	1	3	1
Bollerud (1987)	22	17	-2.05	2.05	2	4.5	1	7	2	0	5	0
Brown, Tappan, Gilligan, Miller, & Argyris (1989)	37	43	-.48	.48	2	6	1	3	1	0	5	1
Carlo, Eisenberg, & Knight (1992)	25	38	-.29		2	3	3	7	6	1	7	1
Carlo, Koller, & Eisenberg (1998)	21	17	-.67		2	2	3	7	6	1	7	1
Carlo, Koller, & Eisenberg (1998)	13	27	0		2	2	3	7	6	1	7	1
Carlo, Koller, Eisenberg, DaSilva, & Frolich (1996)	57	79	0		2	3	3	7	6	1	7	1
Carney (1992)	19	18	-1.62	.53	1	6	1	4	3	1	4	0
Carney (1992)	20	20	-.56	1.11	1	6	1	4	3	1	4	0
Carney (1992)	26	27	-2.58	.63	2	6	1	4	3	1	4	0
Cassidy, Chu, & Dahlsgaard (1997)	13	18	.59		1	3.5	1	2	7	1	3	1
Castor-Scheufler (1994)	205	270	0 ^a	0 ^a	3	3	1	1	6	1	3	0
Clopton & Sorell (1993)	20	20	.06		4	2	1	5	1	1	5	1
Clopton & Sorell (1993)	20	20	-.07		4	2	1	5	1	1	5	1
Cole (1987)	12	12	.34	-.36	2	2	1	3	2	0	5	0
Conley, Jadack, & Hyde (1997)	26	33	-.07		4	6	1	2	2	0	3	1
Craft (1992)	15	20	-.47	.46	1	6	1	2	6	1	5	0
Craft (1992)	19	29	.42	-.42	1	6	1	2	6	1	5	0
Crown & Heatherington (1989)	20	20	0 ^a		3	3	1	2	2	0	3	1
Curror (1994) ^b	28	34	-.37		2	6	1	2 & 3	7	1	7	0
Dekovic & Gerris (1994)	20	20	0 ^a		1	6	3	7	5	1	7	1
Dekovic & Gerris (1994)	19	19	0 ^a		1	6	3	7	5	1	7	1
Dekovic & Gerris (1994)	24	23	0 ^a		1	6	3	7	5	1	7	1
Derry (1989)	20	20	0 ^a		4	3	1	6	2	0	5	1
de Vries & Walker (1986)	36	36	.06 ^a	-.13 ^a	3	3	2	4	4	0	4	1
Dezolt (1992)	10	10	0 ^a	0 ^a	1	3	1	1	2	0	1	0
Dezolt (1992)	10	10	0 ^a	0 ^a	2	3	1	1	2	0	1	0
Dezolt (1992)	10	10	0 ^a	0 ^a	2	3	1	1	2	0	1	0
Diamonti (1993)	39	96	.13	-.13	2	6	1	3	2	0	5	0
Dickey, Kroll, & Jenkins (1987)	10	10	-1.26	.62	4	4	1	3	7	1	5	0
Diederichs (1993)	12	12	-.18	.18	1	1	1	1	3	1	1	0
Dohrenwend (1995)	51	28	-.34	-.04	2	3	1	4	6	1	4	0
Donenberg & Hoffman (1988)	34	35	-.89 ^a		2	3	1	7	2	1	2	1
Dossett (1989)	7	7	-.73	.73	4	7	1	3	1	1	5	0
Dossett (1989)	7	7	-.94	.94	4	7	1	3	1	1	5	0
Dossett (1989)	7	7	-.11	.11	4	3	1	3	1	1	5	0
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	14	16	0 ^a		1	2	3	7	5	1	7	1
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	16	13	0 ^a		1	2	3	7	5	1	7	1
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	16	17	0 ^a		1	2	3	7	5	1	7	1
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	14	16	0 ^a		1	6	3	7	5	1	7	1
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	18	12	0 ^a		1	6	3	7	5	1	7	1
Eisenberg, Boehnke, Schuler, & Silbereisen (1985)	17	22	0 ^a		1	6	3	7	5	1	7	1
Eisenberg, Carlo, Murphy, & Van Court (1995)	20	34	-.28		2	3	3	7	5	1	7	1
Eisenberg, Hertz-Lazarowitz, & Fuchs (1990)	15	13	.40		1	3.5	3	7	5	1	7	1
Eisenberg, Hertz-Lazarowitz, & Fuchs (1990)	10	14	-.10		1	6	3	7	5	1	7	1
Eisenberg, Lennon, & Roth (1983)	17	16	.63		1	3	3	7	5	1	7	1
Eisenberg, Miller, Shell, McNalley, & Shea (1991)	16	16	-.37		2	3	3	7	5	1	7	1
Eisenberg, Pasternack, Cameron, & Tryon (1984)	19	25	0 ^a		1	3	3	7	5	1	7	1
Eisenberg & Shell (1986)	28	28	0 ^a		1	3	3	7	5	1	7	1
Eisenberg et al. (1987)	6	7	-.50		1	3	3	7	5	1	7	1
Eisenberg-Berg & Hand (1979)	17	16	-.17		1	3.5	3	7	5	1	7	1
Eisenberg-Berg & Mussen (1978)	37	35	-.69		2	4	3	7	5	1	7	1
Eisenberg-Berg & Neal (1981)	12	10	.12		1	3	3	7	5	1	5	1

Table 1 (continued)

Study	N males	N females	d_c	d_f	Age	SES	Construct	Type of dilemma	Coding	Scale	Protagonist gender	Published
Eisenberg-Berg & Neal (1981)	12	11	-.38		1	3	3	7	5	1	7	1
Eisenberg-Berg & Neal (1981)	19	22	.39		1	3	3	7	5	1	5	1
Eisenberg-Berg & Neal (1981)	18	23	.10		1	3	3	7	5	1	7	1
Eisenberg-Berg & Roth (1980)	17	16	.31		1	3	3	7	5	1	7	1
Estes (1994)	28	28	.15	-.15	3	3	1	1	2	0	1	0
Ford & Lowery (1986)	101	101	-.18	.07	3	3	1	3	6	1	5	1
Forsyth, Nye, & Kelley (1988)	58	158	0 ^a		3	3	1	4	7	1	4	1
Friedman, Robinson, & Friedman (1987)	47	54	.17		3	3	1	1	6	1	3	1
Fuchs, Eisenberg, Hertz-Lazarowitz, & Sharabany (1986)	12	17	-.29		1	3.5	3	7	5	1	7	1
Fuchs, Eisenberg, Hertz-Lazarowitz, & Sharabany (1986)	21	15	-.56		1	3.5	3	7	5	1	7	1
Fuchs, Eisenberg, Hertz-Lazarowitz, & Sharabany (1986)	33	22	.01		1	3.5	3	7	5	1	7	1
Galotti, Kozberg, & Farmer (1991)	27	34	-.42	.31	2	6	1	4	6	1	4	1
Galotti, Kozberg, & Farmer (1991)	34	39	-.92	-.12	2	6	1	4	6	1	4	1
Galotti, Kozberg, & Farmer (1991)	27	25	.04	-.09	3	3	1	4	6	1	4	1
Garmon, Basinger, Gregg, & Gibbs (1996)	166	162	-.39		6	6	2	4	4	0	4	1
Garrod, Beal, & Shin (1990)	10	8	.59		1	4	1	1	3	1	1	1
Garrod, Beal, & Shin (1990)	5	10	-1.17		1	4	1	1	3	1	1	1
Garrod, Beal, & Shin (1990)	13	8	.17		1	4	1	1	3	1	1	1
Gibbs, Arnold, & Burkhardt (1984)	22	53	-.68		6	3	2	1	4	0	3	1
Gilligan & Attanucci (1988)	46	34	-1.04	1.04	6	6	1	3	3	0	5	1
Graham (1995)	27	26	-.17	-.15	2	4	1	4	6	1	4	0
Hagar (1990)	20	20	-.26	.26	4	3.5	1	3	2	0	5	0
Hopkins-Baker (1995)	204	312	0 ^a		3	3	1	2	6	1	5	0
Hosick (1994)	20	15	-.62	.62	4	6	1	3	2	0	5	0
Hosick (1994)	15	15	-.42	.42	4	6	1	3	2	0	5	0
Jadack, Hyde, Moore, & Keller (1995)	38	36	.03	-.03	3	3	1	2	2	0	3	1
Janoff (1991)	100	100	-1.16	1.19	4	3	1	4	6	1	4	0
Janssens & Dekovic (1997)	20	20	-.30		1	6	3	7	5	1	7	1
Janssens & Dekovic (1997)	19	19	-.50		1	6	3	7	5	1	7	1
Janssens & Dekovic (1997)	24	23	-.23		1	6	3	7	5	1	7	1
Johnston (1988)	28	30	-.56	.56	2	3	1	1	3	0	1	1
Kalkoske (1992)	93	79	-.57	-.23	2	6	1	1	3	1	1	0
Korschgen (1988)	19	19	.67		4	6	1	3	1	1	5	0
Krebs, Vermeulen, Denton, & Carpendale (1994)	10	10	.44	-.44	3	3	1	1	3	1	1	1
Krebs, Vermeulen, Denton, & Carpendale (1994)	10	10	-.15	.15	3	3	1	1	3	1	5	1
Lamm (1992)	12	12	-.69	.69	1	6	1	1	3	1	1	0
Langdale (1986)	16	16	-.83	.83	1	6	1	1	2	0	1	1
Langdale (1986)	24	22	-2.09	2.09	2	6	1	1	2	0	1	1
Langdale (1986)	30	29	-1.24	1.24	4	6	1	1	2	0	1	1
Lantsberger (1993)	95	51	-.38		3	6	1	2	6	1	5	0
Liddell (1998)	12	25	.06	-.28	3	3	1	2	6	1	5	1
Liddell & Davis (1996)	133	248	-.35	0 ^a	3	3	1	2	6	1	5	1
Liddell, Halpin, & Halpin (1993)	185	174	-.50	.16	3	3	1	2	6	1	5	1
Logan (1990)	30	60	-.60	.60	6	3	1	3	2	0	5	0
Lourenco (1991)	16	16	-.07	.07	1	3	2	8	4	0	7	1
Lyons (1983)	14	16	-1.40	1.40	6	4	1	3	2	0	5	1
Maqsd (1998)	40	40	-.41	-.33	2	6	1	2	6	1	5	1
Maqsd (1998)	40	40	.13	-.15	2	6	1	2	6	1	5	1
Maqsd (1998)	40	40	-.42	-.02	2	6	1	2	6	1	5	1
Miller, Eisenberg, Fabes, & Shell (1996)	34	40	0 ^a		1	6	3	7	5	1	7	1
Moore (1988)	35	52	-.26	-.36	3	6	1	1	6	1	3	0
Norris (1996)	126	151	-.21	-.34	4	6	1	7	6	1	5	0
Orenstein (1991)	60	60	-.36	.05	4	3.5	1	4	6	1	4	0
Perez (1994)	97	187	0 ^a		4	6	1	2	6	1	5	0
Peter & Gallop (1994)	26	25	-.63		4	3	1	2	2	1	5	1
Pratt, Diessner, Hunsberger, Pancer, & Savoy (1991)	13	15	-.50	.50	4	6	1	3	2	0	5	1
Pratt, Diessner, Hunsberger, Pancer, & Savoy (1991)	16	20	-.65	.65	5	6	1	3	2	0	5	1
Pratt, Golding, & Hunter (1983)	10	10	0 ^a	0 ^a	3	3	2	1	4	0	1	1
Pratt, Golding, & Hunter (1983)	10	10	0 ^a	0 ^a	4	6	2	1	4	0	1	1
Pratt, Golding, & Hunter (1983)	10	10	0 ^a	0 ^a	5	6	2	1	4	0	1	1
Pratt, Golding, Hunter, & Sampson (1988)	11	8	.15	-.15	3	3	1	3	2	0	5	1
Pratt, Golding, Hunter, & Sampson (1988)	12	8	-2.72	2.72	4	3	1	3	2	0	5	1
Pratt, Golding, Hunter, & Sampson (1988)	11	12	-.85	.85	5	3	1	3	2	0	5	1
Pratt & Royer (1982)	24	24	0 ^a	0 ^a	3	3	1	1	6	1	3	1
Pratt & Royer (1982)	20	20	0 ^a	0 ^a	4	6	1	1	6	1	3	1
Pratt & Royer (1982)	12	12	0 ^a	0 ^a	2	6	1	1	6	1	3	1
Pratt & Royer (1982)	12	12	0 ^a	0 ^a	3	3	1	1	6	1	3	1
Pratt & Royer (1982)	12	12	0 ^a	0 ^a	4	6	1	1	6	1	3	1

(table continues)

Table 1 (continued)

Study	N males	N females	d_c	d_j	Age	SES	Construct	Type of dilemma	Coding	Scale	Protagonist gender	Published
Roberson (1988)	20	26	-.41	.41	2	2	1	3	2	1	5	0
Rogers (1987)	196	198	-1.0	.49	2	3.5	1	4	3	1	4	0
Rothbart, Hanley, & Albert (1986)	25	.25		.70	3	6	1	3	2	1	5	1
Saloner (1988)	10	10	.06		1	1	3	7	5	1	7	0
Schiller (1991)	10	10	-.81	.81	2	4	1	1	1	0	1	0
Schiller (1991)	10	10	-.63	.63	2	4	1	7	1	0	1	0
Schiller (1991)	10	10	-.10	.10	2	4	1	8	1	0	1	0
Schlesinger (1987)	20	20	-.46	.46	4	3	1	3	2	0	5	0
Self, Skeel, & Jecker (1993)	59	28	-.22	.22	4	6	1	3	7	0	5	1
Sitzer (1990)	43	55	-.36	.36	3	3	1	1	2	1	1	0
Skoe & Diessner (1994) ^b	58	76	-.24		3	3	1	2 & 3	7	1	7	1
Skoe & Gooden (1993) ^b	23	23	-1.02		2	6	1	2 & 3	7	1	7	1
Skoe, Pratt, Matthew, & Curror (1996) ^b	30	30	-.84		5	6	1	2 & 3	7	1	7	1
Skoe, Pratt, Matthew, & Curror (1996) ^b	6	8	-1.19		4	6	1	2 & 3	7	1	7	1
Skoe, Pratt, Matthew, & Curror (1996) ^b	5	8	-.62		5	6	1	2 & 3	7	1	7	1
Söchting, Skoe, & Marcia (1994) ^b	45	45	.28		3	3	1	2 & 3	7	1	7	1
Stander & Jensen (1993)	78	75	-.76 ^a		3	3	1	4	6	1	4	1
Stiller & Forrest (1990)	32	45	-1.18	.67	3	3	1	3	3	1	5	1
Stookey (1994)	38	67	.10		4	3	1	2	6	1	5	0
Tietjen (1986)	11	11	0 ^a		1	6	3	7	5	1	7	1
Tietjen (1986)	11	7	0 ^a		1	6	3	7	5	1	7	1
Tietjen (1986)	10	8	0 ^a		2	6	3	7	5	1	7	1
Tietjen (1986)	11	13	0 ^a		4	6	3	7	5	1	7	1
Vera & Levin (1989)	40	39	-.11	.11	3	3	1	2	2	1	6	0
Walker (1986a)	24	38	-.03	.03	4	6	2	1	4	0	1	1
Walker, de Vries, & Trevethan (1987)	54	67	-.42	.42	6	6	1	5	2	0	5	1
Walker, de Vries, & Trevethan (1987)	66	53	-.30	.30	6	6	1	6	2	0	5	1
Walker & Moran (1991)	26	26	0 ^a	0 ^a	6	6	2	1	4	0	1	1
Wark (1993)	51	46	-.30	.30	3	3	1	6	3	0	5	0
Wark & Krebs (1996)	55	55	-.07	.07	3	4	1	6	3	0	5	1
Wark & Krebs (1997)	40	40	-.11	.11	3	3	1	1	3	1	1	0
Webster (1996)	19	26	-.35	.35	4	6	1	3	1	0	5	0
Webster (1996)	13	24	-.94	.94	5	6	1	3	1	0	5	0
White (1994)	124	110	-.43	.49	4	3	1	2	7	1	5	1
White & Manolis (1997)	138	120	-.40	.51	4	3	1	2	2	0	5	1
Wilson (1995)	52	68	-.38		4	3	1	4	6	1	4	1
Yacker & Weinberg (1990)	20	28	-.42		4	3	1	2	6	1	5	1
Yacker & Weinberg (1990)	29	22	-.16		4	3	1	2	6	1	5	1

Note. d_c = care reasoning effect size; d_j = justice reasoning effect size; Age: 1 = child; 2 = adolescent; 3 = university; 4 = adult; 5 = elderly; 6 = mixed age; SES (socioeconomic status): 1 = lower class; 2 = lower middle class; 3 = middle class; 3.5 = middle class to upper middle class; 4 = upper middle class; 4.5 = upper middle class to upper class; 5 = upper class; 6 = mixed/unreported; 7 = male & female participants not matched on SES; Construct: 1 = Gilligan care or justice; 2 = Kohlberg care or justice; 3 = Eisenberg prosocial moral reasoning; Type of dilemma: 1 = standard hypothetical; 2 = real-life hypothetical; 3 = self-generated; 4 = no dilemma; 5 = self-generated care; 6 = self-generated justice; 7 = hypothetical care; 8 = hypothetical justice; Coding: 1 = Brown et al. (1988); 2 = Lyons (1982); 3 = modified Lyons (1982); 4 = Colby et al. (1987); 5 = Eisenberg, Lennon, & Roth (1983); 6 = objective coding; 7 = other; Scale: 0 = categorical score; 1 = continuous score; Protagonist gender: 1 = male; 2 = female; 3 = both male & female protagonists; 4 = no protagonist; 5 = participant is protagonist; 6 = gender not specified; 7 = protagonist is same gender as participant; Published: 0 = unpublished; 1 = published.

^a Estimated effect size. ^b Ethic of Care Interview studies included in meta-analysis of gender differences in level of care reasoning.

Magnitude of Gender Differences in Justice Orientation

The gender difference in justice orientation as reported in the studies was not statistically significant for 72% of the 95 independent samples. Again, it is unlikely that the nonsignificant gender differences in justice reasoning were due to low power to detect differences. Seventy-six percent of the independent samples had sufficiently large sample sizes to detect a moderate effect size for gender differences in justice reasoning.

The overall effect size d for gender differences in justice orientation was .19, indicating a small gender difference favoring males. A homogeneity analysis using procedures specified by Hedges and Becker (1986) indicated that the effect sizes were nonhomogeneous, $H_T = 307.44$, $p < .001$. Therefore, analyses were conducted to determine the extent to which the magnitude of the effect size was moderated by other variables. The results of the moderator analyses are presented in Table 2.

Age. Homogeneity analyses revealed that the magnitude of the effect sizes differed significantly as a function of the respondent's age, $H_B = 52.56$, $p < .001$. As there were so few independent samples in the child and older adult age groups, the effect sizes for gender differences in justice reasoning within these groups must be interpreted with caution. For university students, the magnitude of the effect size d was close to zero. For adolescents, the magnitude of the effect size was small and favored males ($d = .22$), but for younger adults, the magnitude of the effect size d was .40.

Socioeconomic status. Homogeneity analyses revealed that the magnitude of the effect sizes did not differ significantly as a function of the respondent's SES, $H_B = 4.30$, ns .

Moral orientation construct. A homogeneity analysis revealed that the magnitude of these effect sizes did not differ significantly as a function of whether measures based their definition of justice

Table 2
Magnitude of Gender Differences in Moral Orientation as a Function of Moderator Variables

Moderator variables	Care reasoning				Justice reasoning			
	<i>k</i>	<i>d_c</i>	95% confidence interval	<i>H_w</i>	<i>k</i>	<i>d_j</i>	95% confidence interval	<i>H_w</i>
Age group								
Children (11 or younger)	46	-.08	-.18 to .02	51.93	9	.35	.12 to .59	15.19
Adolescents (12-19)	32	-.53	-.61 to -.44	134.36 ^a	23	.22	.12 to .31	89.70 ^a
University	36	-.18	-.24 to -.12	72.37 ^a	32	.04	-.03 to .10	35.03
Younger adults (20-49)	34	-.33	-.41 to -.25	90.54 ^a	22	.40	.30 to .50	103.44 ^a
Older adults (50 or older)	4	-.63	-1.02 to -.25	2.82	4	.63	.25 to 1.02	2.82
Mixed age	8	-.48	-.61 to -.34	15.50 ^a	5	.42	.22 to .62	8.70 ^a
Socioeconomic status group								
Lower class	13	-.08	-.26 to .10	5.82	3	.15	-.26 to .56	2.18
Middle class	76	-.27	-.32 to -.22	227.62 ^a	45	.18	.12 to .23	134.83 ^a
Upper class	13	-.42	-.58 to -.26	38.99 ^a	8	.43	.20 to .66	30.61 ^a
Mixed/unreported	58	-.32	-.39 to -.25	156.95 ^a	39	.18	.10 to .26	135.52 ^a
Moral orientation construct								
Gilligan	111	-.32	-.36 to -.28	385.81 ^a	83	.19	.14 to .23	283.80 ^a
Kohlberg	9	-.25	-.40 to -.10	8.66	12	.19	-.01 to .39	23.64 ^a
Eisenberg	40	-.09	-.19 to .01	26.73	—	—	—	—
Type of dilemma								
Standard hypothetical	36	-.19	-.28 to -.10	80.07 ^a	30	.10	.00 to .19	75.02 ^a
Real-life hypothetical	26	-.20	-.27 to -.14	50.99 ^a	16	.09	.00 to .17	38.86 ^a
Self-generated	28	-.37	-.46 to -.28	72.54 ^a	26	.42	.31 to .53	52.49 ^a
No dilemma	17	-.57	-.66 to -.48	108.07 ^a	12	.40	.28 to .51	56.12 ^a
Designed to elicit care	47	-.17	-.26 to -.09	62.16	4	.05	-.14 to .23	40.25 ^a
Designed to elicit justice	6	-.18	-.37 to .01	1.53	7	.10	-.02 to .22	2.74
Coding scheme								
Brown et al. (1988)	13	-.23	-.43 to -.04	20.23 ^a	10	.37	.14 to .60	14.84
Lyons (1982)	33	-.40	-.50 to -.31	101.39 ^a	30	.42	.32 to .51	95.57 ^a
Lyons (1982), modified	20	-.61	-.72 to -.51	98.20 ^a	12	.32	.20 to .44	33.37 ^a
Colby et al. (1987)	10	-.25	-.39 to -.10	8.68	8	.03	-.18 to .24	1.16
Eisenberg, Lennon, & Roth (1983)	36	-.08	-.19 to .03	22.90	—	—	—	—
Objective coding	44	-.23	-.28 to -.18	112.31 ^a	33	.04	-.02 to .10	102.35 ^a
Other	4	-.23	-.41 to -.04	13.23 ^a	2	.50	.25 to .75	.05
Scale								
Categorical	49	-.38	-.45 to -.30	121.78 ^a	42	.41	.32 to .50	119.15 ^a
Continuous	111	-.26	-.30 to -.22	309.19 ^a	53	.11	.06 to .16	153.37 ^a
Gender of protagonist								
Male	25	-.36	-.49 to -.23	58.43 ^a	21	.34	.19 to .48	49.95 ^a
Male and female	15	-.03	-.14 to .09	11.03	9	-.01	-.14 to .12	3.59
No protagonist	17	-.57	-.66 to -.48	108.07 ^a	12	.40	.28 to .51	56.12 ^a
Participant is protagonist	61	-.27	-.32 to -.22	151.84 ^a	49	.17	.12 to .23	157.48 ^a
Same gender as participant	39	-.11	-.21 to .00	24.18	—	—	—	—
Publication status								
Unpublished	53	-.29	-.35 to -.24	232.27 ^a	47	.17	.10 to .23	168.23 ^a
Published	107	-.28	-.33 to -.23	206.03 ^a	48	.21	.14 to .27	138.53 ^a

Note. *k* represents the number of effect sizes; *d_c* is the effect size for care reasoning; *d_j* is the effect size for justice reasoning; *H_w* is the within-group homogeneity statistic (Hedges & Becker, 1986).

^a Significant nonhomogeneity at *p* < .05, according to the chi-square test.

reasoning on Gilligan's description of the justice orientation or on Kohlberg's fairness and/or normative orientations, $H_B = 0$, *ns*.

Type of dilemma. A homogeneity analysis revealed that the magnitude of the effect size differed significantly as a function of the type of dilemma, $H_B = 41.96$, $p < .001$. For hypothetical dilemmas (standard or real-life) and dilemmas designed to elicit care or justice, the magnitude of the effect size for gender differences in justice reasoning was close to zero. The effect sizes for self-generated dilemmas and measures that did not use a dilemma

were moderate in size and favored males ($d = .42$ and $d = .40$, respectively).

Coding scheme. A homogeneity analysis revealed that the magnitude of the effect size differed significantly as a function of the coding scheme, $H_B = 60.10$, $p < .001$. The effect size was close to zero for those studies that used Kohlberg's scoring scheme or objective scoring schemes ($d = .03$ and $d = .04$, respectively) and $d = .42$ for studies that used Lyons's (1982) coding scheme.

Table 3
Regression Coefficients From Analysis Predicting Care Effect Size From Moderator Variables

Moderator variable	β	p	sr	sr^2
Age group	-.04	ns	-.03	^a
Socioeconomic status	.04	ns	.04	^a
Moral orientation construct				
Eisenberg vs. Kohlberg & Gilligan	-.01	ns	-.01	^a
Kohlberg vs. Gilligan	-.12	ns	-.13	.02
Type of dilemma				
Self-generated vs. standard hypothetical, real-life hypothetical, designed to elicit care or justice	-.21	.05	-.21	.04
No dilemma vs. dilemma	-.29	.001	-.29	.08
Coding scheme				
Objective vs. others	.12	ns	.11	.01
Brown et al. & Lyons modified vs. Lyons, Eisenberg, & Kohlberg	.06	ns	.06	^a
Scale (continuous vs. categorical)	.18	.05	.18	.03
Gender of Protagonist				
Participant vs. fictional	-.07	ns	-.06	^a
Male vs. nonmale	-.07	ns	-.07	^a
Publication status	-.05	ns	-.05	^a

Note. β = standardized beta; sr = semipartial correlation; sr^2 = squared semipartial correlation.

^a $sr^2 < .01$.

Scale. A homogeneity analysis revealed that the magnitude of the effect size differed significantly as a function of the scale of measurement, $H_B = 34.92$, $p < .005$. Although the effect size was $d = .11$ for measures that yielded continuous outcomes, it was larger in size for measures that yielded categorical outcomes ($d = .41$).

Gender of protagonist. A homogeneity analysis revealed that the magnitude of the effect size differed significantly as a function of the gender of the protagonist, $H_B = 40.30$, $p < .001$. The magnitude of the effect was small when the participant was the protagonist or when there were both male and female protagonists. The magnitude of the effect was moderate in size when the protagonist was male or when the dilemma did not contain a protagonist.

Publication status. The magnitude of the effect size did not differ significantly as a function of whether the study was published, $H_B = .68$, ns.

Regression Analysis for Justice Orientation

Because the overall test for the homogeneity of variance revealed that the effect sizes were nonhomogeneous, a multiple regression analysis was conducted to determine how much the moderator variables accounted for the variance in the effect sizes (Table 4). The corrected effect size for justice orientation was the criterion variable (Hedges, 1981), and the moderator variables were entered in stepwise fashion. The moderator variables accounted for a significant 17% of the variance in the effect sizes for gender differences in justice reasoning, $F = 9.38$, $p < .001$. Again, squared semipartial regression coefficients were computed to determine how much of the variance in the effect sizes was uniquely accounted for by each of the moderator variables when the other moderators were controlled. Table 4 indicates that the contrast

comparing self-generated dilemmas with standard hypothetical dilemmas, real-life hypothetical dilemmas, and dilemmas designed to elicit care or justice reasoning was a significant predictor of the magnitude of the effect size for gender differences in justice reasoning, $\beta = .28$, $p < .01$, and uniquely accounted for 8% of that variance. The contrast comparing studies yielding categorical outcomes versus those with continuous outcomes was also significant, $\beta = -.25$, $p < .01$, and accounted for 7% of the variance in the effect sizes for justice reasoning. None of the other moderator variables had significant effects.

Estimated Effect Sizes Excluded

When the effect sizes estimated as zero were excluded, the magnitude of the overall gender difference in care reasoning was $d = -.35$, and the magnitude of the gender difference in justice reasoning was $d = .22$. Because the effect sizes for justice and care reasoning were similar in magnitude and sign regardless of whether the estimated effect sizes were excluded, the moderator analyses were not rerun.

Levels of Care Reasoning

As discussed in the introduction, the ECI (Skoe & Marcia, 1991) measures the levels of care reasoning originally proposed by Gilligan (1982). A separate meta-analysis was conducted to estimate the magnitude of the effect size for gender differences in level of care reasoning. The overall effect size d for gender differences in level of care reasoning was $-.34$, indicating a small to moderate difference favoring females. As there were only seven independent samples included in this meta-analysis, it was not possible to conduct moderator analyses. It is worth noting, however, that with one exception, females demonstrated higher levels of care reasoning than males across age groups and that the

Table 4
Regression Coefficients From Analysis Predicting Justice Effect Size From Moderator Variables

Moderator variable	β	p	sr	sr^2
Age group	.04	ns	.04	^a
Socioeconomic status	-.01	ns	-.01	^a
Moral orientation construct				
Kohlberg vs. Gilligan	-.04	ns	-.04	^a
Type of dilemma				
Self-generated vs. standard hypothetical, real-life hypothetical, designed to elicit care or justice	.28	.01	.28	.08
No dilemma vs. dilemma	.14	ns	.15	.02
Coding scheme				
Objective vs. others	-.12	ns	-.09	^a
Brown et al. & Lyons modified vs. Lyons & Kohlberg	-.02	ns	-.01	^a
Scale (continuous vs. categorical)	-.25	.01	-.26	.07
Gender of protagonist				
Participant vs. fictional	-.03	ns	-.03	^a
Male vs. nonmale	.10	ns	.10	.01
Publication status	.05	ns	.05	^a

Note. β = standardized beta; sr = semipartial correlation; sr^2 = squared semipartial correlation.

^a $sr^2 < .01$.

magnitude of the effect size was greater among middle-aged and older adults compared with adolescents and young adults.

Discussion

This review used modern meta-analytic techniques to analyze 113 studies of moral reasoning (160 independent samples measuring care and 95 measuring justice). This meta-analysis allowed us to evaluate empirically Gilligan's assertion that the justice and care orientations are strongly gender differentiated. The magnitude of the effect size for gender differences in care reasoning was $-.28$, indicating a small difference favoring females. The magnitude of the effect size for gender differences in justice reasoning was $.19$, indicating a small difference favoring males. The small magnitude of these effects, combined with the finding that 73% of the studies that measured care reasoning and 72% of the studies that measured justice reasoning failed to find significant gender differences, leads us to conclude that, although distinct moral orientations may exist, these orientations are not strongly associated with gender.

This analysis also indicates that the magnitude of the effect size for gender differences in care and justice reasoning differs as a function of specific variables that have often been confounded with the measurement of moral orientation. An important contribution of this meta-analysis is the finding that specific moderator variables uniquely account for variance in the effect sizes for care and justice reasoning when other moderators are controlled.

What Moderates Gender Differences in Moral Orientation?

All variables except publication status significantly moderated the magnitude of the effect size for gender differences in care and justice reasoning, and the implication of these findings is discussed below. However, the results must be interpreted with caution, bearing in mind that the regression analyses revealed relatively few unique effects of these variables on the magnitude of the effect size for justice or care reasoning.

Age. According to Gilligan and colleagues (Gilligan & Wiggins, 1987), gender differences in the use of care and justice reasoning should increase as children get older and their experiences of attachment and inequality are differentially reinforced. The findings from the moderator analyses do not support this claim. Although the magnitude of the effect size for gender differences in care reasoning did increase from childhood to adolescence, it then decreased sharply amongst university students. One possibility is that as individuals get older, they become more adept at using more than one mode of moral reasoning. Gilligan (1986a) has suggested that to be successful in a male-dominated society, females are expected to be conversant in both modes of moral reasoning. Indeed, the more recent work of Gilligan and her colleagues (Brown & Gilligan, 1993) has explored the idea that, in adolescence, girls become aware that their own moral voice is neither recognized nor valued by society. Consequently, they come to question their own understanding of their experience and conform to interpretations imposed by others. In so doing, they learn to use the justice perspective instead of or along with the care perspective. Consistent with this possibility is the finding that gender differences in justice reasoning were also smallest among university students. However, neither of these findings make clear

whether gender differences in care or justice reasoning decreased among university students because young women changed their mode of moral reasoning (e.g., used less care reasoning or more justice reasoning), because young men changed their mode of moral reasoning (e.g., used more care reasoning or less justice reasoning), or because both groups changed.

Gender differences in both care and justice reasoning increased again in the younger adult group. To the extent that younger adults are more likely to be involved in family formation, this finding is consistent with the possibility that family organization during the period of active parenting leads to increased gender-role polarization and, consequently, to gender-stereotyped moral reasoning (Pratt et al., 1988).

Socioeconomic status. Although gender differences in care reasoning increased from lower-class to upper-class groups, the magnitude of the effect size for justice reasoning did not differ as a function of social class. Again, these findings do not speak to the question of whether gender differences in care reasoning are smallest among lower-class groups because low social status promotes a concern with others who, in part, determine one's outcomes (Tronto, 1987) or because low social status promotes a concern with fairness and rights because these rectify the inequalities faced by minority or subordinate groups (Beal et al., 1997). The findings do, however, suggest that social class and gender do not interact, such that lower-class women are more likely than other groups to endorse the care orientation or the justice orientation.

Moral orientation construct. Given Gilligan's (1982) claims that Kohlberg's coding scheme is insensitive to the care orientation, it is notable that the magnitudes of the effect size for gender differences in care reasoning as conceptualized by Gilligan and by Kohlberg were of approximately equal size. In the case of justice reasoning, the magnitude of the effect size was identical for studies that defined the justice orientation according to Gilligan's definition versus those studies that defined the justice orientation according to Kohlberg. It is also notable that the smallest effect size for gender differences in care reasoning was from studies that conceptualized care as prosocial moral reasoning. This might be explained by the fact that the prosocial dilemmas "pull for" care reasoning from both male and female respondents.

Type of dilemma. Moderately large gender differences were found for both care and justice reasoning among measures that did not include a dilemma. These included the Washington University Sentence Completion Test (WUSCT; Rogers, 1987), the CROS (Atunzu, 1986), the Sociomoral Reflection Measure-Short Form (SRM-SF; Gibbs, Basinger, & Fuller, 1992), and the Ford and Lowery questionnaire (Ford & Lowery, 1986). The CROS and the Ford and Lowery questionnaire present participants with descriptions of care or justice reasoning and ask them how much these characterize their own thinking about moral dilemmas or how important such considerations are in resolving moral problems. In the case of these measures, it is the participant, and not the researcher, who determines whether or not the participant is a justice or care reasoner. Participants' perceptions of their own moral reasoning may be influenced by their own self-concept and may differ from the perceptions of a researcher who is ostensibly coding a response from a more objective perspective.

In the WUSCT, respondents complete sentence stems that are then coded for the presence of justice and care reasoning. The

SRM-SF is designed to measure developmental stages of moral judgment. It elicits responses related to the evaluation and justification of the importance of sociomoral values such as contract and truth, affiliation, property and law, and legal justice. These can then be coded for various stage aspects, such as empathic role-taking, that are related to moral reasoning content. It is less clear why these measures should lead to larger gender differences in care reasoning. It is worth noting, however, that the no dilemma group included only one study that used the SRM-SF and only three studies that used the WUSCT.

Importantly, the moderator analyses also revealed that the magnitudes of the effect size for both care and justice reasoning were greater for self-generated dilemmas compared with dilemmas in which the content was standardized. Indeed, this comparison accounted for 4% of the variance in the magnitude of the effect sizes for care and 8% of the variance in the magnitude of the effect sizes for justice reasoning. This finding provides strong support for studies that have found that gender differences in moral orientation are a function of the situation to which individuals are responding (Walker, 1989; Walker et al., 1987; Wark & Krebs, 1997). On the basis of such findings, Walker (1986a) has argued that males and females do not differ in basic moral orientation but rather in the kinds of problems they encounter and report. This argument suggests that researchers who study gender differences in moral orientation must make sure that moral orientation is not inadvertently confounded with dilemma content.

Coding scheme. Although the magnitude of the effect size for gender differences in both care and justice reasoning differed significantly as a function of the coding scheme, this variable accounted for little of the variance in effect size estimates when the other moderators were controlled. This suggests that the coding schemes themselves are less important than other characteristics associated with them, such as the type of outcome (categorical vs. continuous) they yield.

Scale. The effect size for gender differences in care and justice reasoning was larger when moral orientation was a categorical instead of a continuous outcome. There are at least two explanations for this finding. One is that categorical measures cause males and females to look more different than they are. For instance, in Lyons's coding scheme, classification as a justice or care reasoner is based on modal moral orientation. Consequently, if 51% of one person's considerations are care considerations, she is considered a care reasoner and if 51% of another person's considerations are justice considerations, he is considered a justice reasoner. Despite the fact that these two individuals offer virtually the same proportion of care and justice considerations, they are categorized as displaying distinct moral orientations.

The second possibility is that because categorical outcomes were converted into continuous scores for the purpose of computing an effect size, the effect size for these transformed scores was less precise than the effect size for continuous measures. In some cases, these transformed scores may have inflated the magnitude of the effect size. If this was the case, the overall effect size for continuous measures better represents the true gender difference in care orientation.

This finding implies that it might be useful for future studies to use measures that can be scored both continuously and categorically. Such measures need not assume that justice and care represent opposite ends of a single spectrum (although such a concep-

tualization is implicit in, e.g., a percent care score). Justice and care reasoning can be measured as a two-dimensional construct as in, for example, the MMO (Liddell, 1990).

Gender of protagonist. Although the magnitude of the effect sizes for gender differences in both care and justice reasoning differed significantly as a function of the gender of the protagonist of the dilemma, it is important to note that this variable accounted for 1% or less of the variance in the effect size estimates when controlling for other moderators. Although the perspective from which an individual responds to a dilemma may influence moral orientation, it does not appear to do so differently for males and females.

In summary, the results of the care and justice analyses suggest that researchers who study gender differences in moral orientation should be aware of the extent to which moral orientation is influenced by the content of specific moral dilemmas and by the context in which these dilemmas arise. Importantly, however, the moderator analyses demonstrate that even though the magnitude of the effect size does differ as a function of variables like age and dilemma type, among no group is the effect size for gender differences in care or justice reasoning ever more than moderate in size, and in most cases, it is considerably smaller. Thus, researchers may benefit from turning their attention away from the study of gender differences in moral orientation and toward a more sophisticated characterization of moral orientation or to questions of how moral orientations develop over time.

Implications

An important limitation of this meta-analysis is that we were able to determine only relative gender differences in the use of care and justice reasoning and not whether females predominantly use care reasoning and males predominantly use justice reasoning. In its strongest form, Gilligan's assertion is that females consistently use the care orientation and infrequently use the justice orientation and that males consistently use the justice orientation and infrequently use the care orientation. This meta-analysis was able to evaluate only whether females used the care orientation more than males and whether males used the justice orientation more than females. However, it is entirely possible that all participants used a mix of justice and care reasoning, with females using relatively more care reasoning than males and males using relatively more justice reasoning than females. Many of the findings regarding gender differences in moral orientation are consistent with this possibility (Walker, 1991; Wark & Krebs, 1996).

It is possible that we might have observed a gender-related pattern in justice and care reasoning if we had looked at differences in moral orientation as a function of gender-role identity (i.e., masculine, feminine, or androgynous). This, however, was not feasible because it would have required us to know the gender-role identity classification for each study member. It is worth noting, however, that the evidence for an association between gender-role identity and moral orientation is weak (Pratt & Royer, 1982), with most studies failing to find significant differences in moral reasoning as a function of gender-role identity (Abaris, 1990; Hagar, 1990; Stookey, 1994; Wark & Krebs, 1996). It is unlikely, then, that gender-role identity would have proved to be an important moderator of gender differences in moral orientation.

If, as the results of this meta-analysis imply, it is the case that most males and females use a mix of justice and care reasoning, the field should move beyond the debate about gender differences in moral reasoning and focus instead on how individuals integrate justice and care reasoning and under what conditions they determine which is a more adequate basis for moral action. For instance, Johnston (1988) found that boys and girls were equally capable of switching moral orientations when prompted to offer a different solution to a moral dilemma, and the solution that participants considered best was not always the one they offered originally. In follow-up interviews with her participants, Johnston found that boys, even if they thought the care solution was the best one, would not use it unless they believed the relationship between the characters in the dilemma could be salvaged. These findings suggest that the practical constraints of the situation may dictate one's choice of which moral orientation to implement.

What is most ironic is that the notion of moral voice that lies at the heart of the current debate over gender differences in moral reasoning has little to do with how Gilligan, Brown, and their colleagues on the Harvard Project on Women's Psychology and Girls' Development currently conceptualize moral voice (Brown, 1994; Brown & Gilligan, 1993; Gilligan, 1994). It is telling that not one of the authors represented in this meta-analysis is included in the body of work that, according to Gilligan (1994), best supports her theory of women's psychological development. Much of this discrepancy can be traced to the fact that Gilligan and her colleagues (Brown, 1994; Gilligan, 1994) argue from a different epistemological framework. They argue that quantitative tests of gender differences in moral reasoning miss the point because they construe justice and care as opposite ends of a single continuum instead of recognizing that these represent distinct moral perspectives. Quantitative researchers, on the other hand, argue that the conclusions reached by the Harvard Project are not falsifiable (Walker, 1995), that the care and justice constructs are insufficiently well defined and impossible to operationalize (Breakwell, 1994), and that measures such as the *Reading Guide* (Brown et al., 1988) are not well validated. This is not so much a debate about a specific theory or finding as it is a debate about epistemology. Unfortunately, the intersection between qualitative and quantitative methodologies is not well defined, and there are few common benchmarks for evaluating different epistemological approaches (Jaffee, Kling, Plant, Sloan, & Hyde, 1999). For instance, Brown, Tappan, Gilligan, Miller, and Argyris (1989) pointed out that traditional psychometric conceptions of reliability and validity are based on assumptions that "render them inappropriate for interpretive approaches" (p. 156). What this implies is that researchers who develop psychometrically valid and reliable measures of moral orientation will satisfy Gilligan's critics on the one hand but will fail to capture the construct of moral orientation as she and her colleagues have developed and refined it over the years.

Do We Need the Care Orientation?

Gilligan's critics suggest a number of reasons why the care orientation is unnecessary and/or insufficient as a mode of moral reasoning. For instance, Kohlberg and colleagues (Higgins, 1989; Kohlberg et al., 1983) have claimed that his cognitive-developmental framework encompasses both care and justice, suggesting that the ability to take another's perspective, or moral

role-taking, fosters a sense of responsibility to care. In addition, Colby et al. (1983) showed that considerations of care, relationship, and interpersonal trust are represented as norms and elements at each stage in the MJI scoring scheme, and various researchers have demonstrated that care reasoning can be elicited by the MJI dilemmas (Walker et al., 1987; Wark & Krebs, 1996, 1997). From this standpoint, the care orientation does not require a separate theoretical and psychometric structure.

On the other hand, these findings underscore that the care orientation does exist and can be represented in individuals' responses to moral dilemmas. Moreover, the suggestion that the care and justice orientations can be integrated at the highest levels of Kohlbergian reasoning does not adequately capture Gilligan's notion that these orientations represent distinct moral perspectives that exist in tension. However, even if one agrees that Kohlbergian moral reasoning does not encompass Gilligan's (1982) vision of moral maturity, it is not clear whether Gilligan's conceptualization forms an adequate basis for moral action in that it does not offer guidelines as to when one moral perspective should precede the other. Unfortunately, Gilligan and her colleagues (Brown & Gilligan, 1993) have moved away from work on moral development to a broader focus on girls' and women's psychological development, and it is unlikely that the necessary clarification will be offered.

Conclusion

The results of this meta-analysis do not indicate that the care and justice orientations are strongly gender differentiated. Moreover, the results of the moderator analyses support previous findings that the type of moral reasoning an individual uses is highly sensitive to the context and content of the dilemma. Although these results weaken Gilligan's strongest claims about the relation between gender and moral reasoning, they do not diminish her larger point. In a 1994 interview (Kitzinger & Gilligan, 1994), Gilligan highlighted the extent to which *In a Different Voice* was motivated by the observation that there was a "difference between the way psychologists talked about the self and about morality and the way people actually experienced it" (p. 409). This, perhaps, is the most important message to take from Gilligan's work. Debates over the association between gender and moral orientation and the moral adequacy of the care orientation, although important and necessary, miss the larger point that if psychological theories of human development intend to represent lived experience, then they must be constructed with the diversity of such experience in mind.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- *Abaris, N. L. (1990). Gender and sex-role differences in moral development among undergraduates: Ethics of care versus ethics of justice (Doctoral dissertation, Rutgers University, 1990). *Dissertation Abstracts International*, 51, 3588B.
- *Abide, M. M. (1994). Decisions about substance use: Moral reasoning and moral orientation (Doctoral dissertation, University of Virginia, 1994). *Dissertation Abstracts International*, 55, 905A.
- *Akman, D. (1991). Gender differences in moral orientation (Masters thesis, University of Calgary, 1991). *Masters Abstracts International*, 30, 965.

- *Albrecht, E. F. (1989). The relationship between protagonist gender in hypothetical moral dilemmas and the use of justice versus care moral orientation (Doctoral dissertation, Adelphi University, 1989). *Dissertation Abstracts International*, 49, 1734A.
- *Arvizu, D. R. (1995). The care voice and American Indian college students: An alternative perspective for student development professionals. *Journal of American Indian Education*, 34, 1-17.
- *Atunzu, K. O. (1986). The relationship among ethics of care, ethics of justice, psychosocial intimacy and gender (Doctoral dissertation, Claremont Graduate University, 1986). *Dissertation Abstracts International*, 47, 2505A.
- Auerbach, J., Blum, L., Smith, V., & Williams, C. (1985). On Gilligan's *In a Different Voice*. *Feminist Studies*, 11, 149-161.
- *Barnett, M. A., Quackenbush, S. W., & Sinisi, C. S. (1995). The role of critical experiences in moral development: Implications for justice and care orientations. *Basic and Applied Social Psychology*, 17, 137-152.
- Baumrind, D. (1986). Sex differences in moral reasoning: Response to Walker's (1984) conclusion that there are none. *Child Development*, 57, 511-521.
- *Beal, C. R., Garrod, A., Ruben, K., & Stewart, T. L. (1997). Children's moral orientation: Does the gender of the dilemma character make a difference? *Journal of Moral Education*, 26, 45-58.
- Benenson, J. F., Morash, D., & Petrakos, H. (1998). Gender differences in emotional closeness between preschool children and their mothers. *Sex Roles*, 38, 975-985.
- *Bollerud, K. H. (1987). Adolescent sexual decision-making: The influence of gender and pubertal development on moral reasoning (Doctoral dissertation, Harvard University, 1987). *Dissertation Abstracts International*, 48, 1704A.
- Breakwell, G. M. (1994). Review of *In a Different Voice* by Carol Gilligan. *Feminism and Psychology*, 4, 404-406.
- Brown, L. M. (1994). Standing in the crossfire: A response to Tavris, Gremmen, Lykes, Davis and Contratto. *Feminism and Psychology*, 4, 382-398.
- Brown, L. M., Argyris, D., Attanucci, J., Bardige, B., Gilligan, C., Johnston, D. K., Miller, B., Osborne, R., Tappan, M., Ward, J., Wiggins, G., & Wilcox, D. (1988). *A guide to reading narratives of conflict and choice for self and relational voice*. Cambridge, MA: Project on the Psychology of Women and the Development of Girls, Harvard Graduate School of Education.
- Brown, L. M., & Gilligan, C. (1993). Meeting at the crossroads: Women's psychology and girls' development. *Feminism and Psychology*, 3, 11-35.
- *Brown, L. M., Tappan, M. B., Gilligan, C., Miller, B. A., & Argyris, D. E. (1989). Reading for self and moral voice: A method for interpreting narratives of real-life moral conflict and choice. In M. J. Packer & R. B. Addison (Eds.), *Entering the circle: Hermeneutic investigation in psychology* (pp. 141-164). Albany: State University of New York Press.
- Bussey, K., & Maughan, B. (1982). Gender differences in moral reasoning. *Journal of Personality and Social Psychology*, 42, 701-706.
- *Carlo, G., Eisenberg, N., & Knight, G. P. (1992). An objective measure of adolescents' prosocial moral reasoning. *Journal of Research on Adolescence*, 2, 331-349.
- *Carlo, G., Koller, S. H., & Eisenberg, N. (1998). Prosocial moral reasoning in institutionalized delinquent, orphaned, and noninstitutionalized Brazilian adolescents. *Journal of Adolescent Research*, 13, 363-376.
- *Carlo, G., Koller, S. H., Eisenberg, N., DaSilva, M. S., & Frolich, C. B. (1996). A cross-national study on the relations among prosocial moral reasoning, gender role orientations, and prosocial behaviors. *Developmental Psychology*, 32, 231-240.
- *Carney, C. C. (1992). Children's concepts of good and bad: Age and sex differences in moral reasoning (Doctoral dissertation, University of California, Berkeley, 1992). *Dissertation Abstracts International*, 53, 5458B.
- *Cassidy, K. W., Chu, J. Y., & Dahlsgaard, K. K. (1997). Preschoolers' ability to adopt justice and care orientations to moral dilemmas. *Early Education and Development*, 8, 419-434.
- *Castor-Scheuffler, M. G. (1994). Looking for Gilligan's "other voice": The relation of gender identity, empathy, and moral development (Doctoral dissertation, University of Houston, 1994). *Dissertation Abstracts International*, 55, 4619B.
- Chodorow, N. (1978). *The reproduction of mothering*. Berkeley: University of California Press.
- *Clopton, N. A., & Sorell, G. T. (1993). Gender differences in moral reasoning: Stable or situational? *Psychology of Women Quarterly*, 17, 85-101.
- Colby, A., & Damon, W. (1983). Listening to a different voice: A review of Gilligan's *In a Different Voice*. *Merrill-Palmer Quarterly*, 29, 473-481.
- Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral development. *Monographs of the Society for Research in Child Development*, 48(1-2, Serial No. 200).
- Colby, A., Kohlberg, L., Speicher, B., Hewer, A., Candee, D., Gibbs, J., & Power, C. (1987). *The measurement of moral judgment: Standard issue scoring manual* (Vol. 2). Cambridge, England: Cambridge University Press.
- *Cole, E. H. S. (1987). In their own voices: Hawaiian-American adolescents speak of moral choice and personal changes (Doctoral dissertation, Harvard University, 1987). *Dissertation Abstracts International*, 48, 2576A.
- *Conley, T. D., Jadack, R. A., & Hyde, J. S. (1997). Moral dilemmas, moral reasoning, and genital herpes. *Journal of Sex Research*, 34, 256-266.
- *Craft, D. A. (1992). Children's moral orientation: An instrument validation study to assess school children's moral orientation (Doctoral dissertation, University of San Francisco, 1992). *Dissertation Abstracts International*, 54, 170A.
- *Crown, J., & Heatherington, L. (1989). The costs of winning? The role of gender in moral reasoning and judgments about competitive athletic encounters. *Journal of Sport and Exercise Psychology*, 11, 281-289.
- *Curror, S. (1994). The relationship between empathic listening styles and moral reasoning development in adolescents (Masters thesis, Wilfrid Laurier University, 1994). *Masters Abstracts International*, 33, 664.
- *Dekovic, M., & Gerris, J. R. M. (1994). Developmental analysis of social cognitive and behavioral differences between popular and rejected children. *Journal of Applied Developmental Psychology*, 15, 367-386.
- *Derry, R. (1989). An empirical study of moral reasoning among managers. *Journal of Business Ethics*, 8, 855-862.
- *de Vries, B., & Walker, L. J. (1986). Moral reasoning and attitudes toward capital punishment. *Developmental Psychology*, 22, 509-513.
- *Dezolt, D. M. (1992). Themes, age, and gender differences in children's stories about caring (Doctoral dissertation, Kent State University, 1992). *Dissertation Abstracts International*, 53, 2741A.
- *Diamonti, N. L. (1993). An investigation of gender differences in adolescent moral reasoning about rule breaking behavior (Doctoral dissertation, Boston College, 1993). *Dissertation Abstracts International*, 54, 2892A.
- *Dickey, B., Kroll, B. S., & Jenkins, L. (1987, November). *Gilligan revisited: Methodological issues in the study of gender and moral development*. Paper presented at the National Convention of the Speech Communication Association, Boston, MA.
- *Diederichs, M. (1993). Young children and moral understanding (Masters thesis, California State University, Fullerton, 1993). *Masters Abstracts International*, 32, 772.
- *Dohrenwend, A. M. (1995). Moral orientation in law students: A longitudinal study from a relational perspective (Doctoral dissertation, University of North Dakota—Grand Forks, 1995). *Dissertation Abstracts International*, 57, 1030A.
- *Donenberg, G. R., & Hoffman, L. W. (1988). Gender differences in moral development. *Sex Roles*, 18, 701-717.
- *Dossett, M. E. (1989). Moral orientation of males and females in parental

- caregiving, parental non-caregiving, and non-parental roles (Doctoral dissertation, University of Alabama, 1989). *Dissertation Abstracts International*, 51, 792A.
- *Eisenberg, N., Boehnke, K., Schuler, P., & Silbereisen, R. K. (1985). The development of prosocial behavior and cognitions in German children. *Journal of Cross-Cultural Psychology*, 16, 69–82.
- *Eisenberg, N., Carlo, G., Murphy, B., & Van Court, P. (1995). Prosocial development in late adolescence. *Child Development*, 66, 1179–1197.
- Eisenberg, N., Fabes, R., & Shea, C. (1989). Gender differences in empathy and prosocial moral reasoning: Empirical investigations. In M. Brabeck (Ed.) *Who cares? Theory, research, and implications of the ethic of care* (pp. 127–143). New York: Praeger.
- *Eisenberg, N., Hertz-Lazarowitz, R., & Fuchs, I. (1990). Prosocial moral judgment in Israeli kibbutz and city children: A longitudinal study. *Merrill-Palmer Quarterly*, 36, 273–285.
- *Eisenberg, N., Lennon, R., & Roth, K. (1983). Prosocial development: A longitudinal study. *Developmental Psychology*, 19, 846–855.
- *Eisenberg, N., Miller, P. A., Shell, R., McNalley, S., & Shea, C. (1991). Prosocial development in adolescence: A longitudinal study. *Developmental Psychology*, 27, 849–857.
- *Eisenberg, N., Pasternack, J. F., Cameron, E., & Tryon, K. (1984). The relation of quantity and mode of prosocial behavior to moral cognitions and social style. *Child Development*, 55, 1479–1485.
- *Eisenberg, N., & Shell, R. (1986). Prosocial moral judgment and behavior in children: The mediating role of cost. *Personality and Social Psychology Bulletin*, 12, 426–433.
- *Eisenberg, N., Shell, R., Pasternack, J., Lennon, R., Beller, R., & Mathy, R. M. (1987). Prosocial development in middle childhood: A longitudinal study. *Developmental Psychology*, 23, 712–718.
- Eisenberg-Berg, N. (1979). Development of children's prosocial moral judgment. *Developmental Psychology*, 15, 128–137.
- *Eisenberg-Berg, N., & Hand, M. (1979). The relationship of preschoolers' reasoning about prosocial moral conflicts to prosocial behavior. *Child Development*, 50, 356–363.
- *Eisenberg-Berg, N., & Mussen, P. (1978). Empathy and moral development in adolescence. *Developmental Psychology*, 14, 185–186.
- *Eisenberg-Berg, N., & Neal, C. (1981). Children's moral reasoning about self and others: Effects of identity of the story character and cost of helping. *Personality and Social Psychology Bulletin*, 7, 17–23.
- Eisenberg-Berg, N., & Roth, K. (1980). Development of young children's prosocial and moral judgment: A longitudinal follow-up. *Developmental Psychology*, 16, 375–376.
- *Estes, J. M. (1994). Gender differences in moral reasoning comparing pre-professional university students in nursing, education, and engineering: A content analysis (Doctoral dissertation, University of Alabama, Tuscaloosa, 1994). *Dissertation Abstracts International*, 55, 1866A.
- Fishkin, J., Keniston, K., & MacKinnon, C. (1973). Moral reasoning and political ideology. *Journal of Personality and Social Psychology*, 27, 109–119.
- Flanagan, O., & Jackson, K. (1987). Justice, care, and gender: The Kohlberg-Gilligan debate revisited. *Ethics*, 97, 622–637.
- *Ford, M. R., & Lowery, C. R. (1986). Gender differences in moral reasoning: A comparison of the use of justice and care orientations. *Journal of Personality and Social Psychology*, 50, 777–783.
- *Forsyth, D. R., Nye, J. L., & Kelley, K. (1988). Idealism, relativism, and the ethic of caring. *Journal of Psychology*, 122, 243–248.
- Freeman, S. J. M., & Giebing, J. W. (1979). Moral judgment as a function of age, sex, and stimulus. *Journal of Psychology*, 102, 43–47.
- *Friedman, W. J., Robinson, A. B., & Friedman, B. L. (1987). Sex differences in moral judgments? A test of Gilligan's theory. *Psychology of Women Quarterly*, 11, 37–46.
- *Fuchs, I., Eisenberg, N., Hertz-Lazarowitz, R., & Sharabany, R. (1986). Kibbutz, Israeli city, and American children's moral reasoning about prosocial moral conflicts. *Merrill-Palmer Quarterly*, 32, 37–50.
- *Galotti, K. M., Kozberg, S. F., & Farmer, M. C. (1991). Gender and developmental differences in adolescents' conceptions of moral reasoning. *Journal of Youth and Adolescence*, 20, 13–30.
- *Garmon, L. C., Basinger, K. S., Gregg, V. R., & Gibbs, J. C. (1996). Gender differences in stage and expression of moral judgment. *Merrill-Palmer Quarterly*, 42, 418–437.
- Garrod, A., & Beal, C. (1993). Voices of care and justice in children's responses to fable dilemmas. In A. Garrod (Ed.), *Approaches to moral development: New research and emerging themes* (pp. 59–71). New York: Teachers College Press.
- *Garrod, A., Beal, C., & Shin, P. (1990). The development of moral orientation in elementary school children. *Sex Roles*, 22, 13–27.
- Garwood, S. G., Levine, D. W., & Ewing, L. (1980). Effect of protagonist's sex on assessing gender differences in moral reasoning. *Developmental Psychology*, 16, 677–678.
- *Gibbs, J. C., Arnold, K. D., & Burkhart, J. E. (1984). Sex differences in the expression of moral judgment. *Child Development*, 55, 1040–1043.
- Gibbs, J. C., Basinger, K. S., & Fuller, D. (1992). *Moral maturity: Measuring the development of sociomoral reflection*. Hillsdale, NJ: Erlbaum.
- Gilligan, C. (1977). In a different voice: Women's conception of the self and morality. *Harvard Education Review*, 47, 481–517.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Gilligan, C. (1986a). On *In a different voice*: An interdisciplinary forum: Reply. *Signs*, 11, 324–333.
- Gilligan, C. (1986b). Remapping the moral domain: New images of the self in relationship. In T. C. Heller, M. Sosna, & D. E. Wellbery (Eds.), *Reconstructing individualism: Autonomy, individuality, and the self in western thought* (pp. 237–252). Stanford, CA: Stanford University Press.
- Gilligan, C. (1987). Moral orientation and moral development. In E. F. Kittay & D. T. Meyers (Eds.), *Women and moral theory* (pp. 19–33). Totowa, NJ: Rowman & Littlefield.
- Gilligan, C. (1994). Afterword: The power to name. *Feminism and Psychology*, 4, 420–424.
- *Gilligan, C., & Attanucci, J. (1988). Two moral orientations: Gender differences and similarities. *Merrill-Palmer Quarterly*, 34, 223–237.
- Gilligan, C., Brown, L. M., & Rogers, A. G. (1990). Psyche embedded: A place for body, relationships, and culture in personality theory. In A. I. Rabin, R. A. Zucker, R. A. Emmons, & S. Frank (Eds.), *Studying persons and lives* (pp. 86–147). New York: Springer.
- Gilligan, C., & Wiggins, G. (1987). The origins of morality in early childhood relationships. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 277–305). Chicago: University of Chicago Press.
- *Graham, S. (1995). The effectiveness and psychological impact of a high school community service course (Doctoral dissertation, University of California, Los Angeles, 1995). *Dissertation Abstracts International*, 56, 868A.
- Guttman, D. (1985). The parental imperative revisited. In J. Meacham (Ed.), *The family and individual development* (pp. 31–60). Basel, Switzerland: Karger.
- Haan, N. (1978). Two moralities in action contexts: Relationships to thought, ego regulation, and development. *Journal of Personality and Social Psychology*, 36, 286–305.
- Haan, N., Smith, B., & Block, J. (1968). Moral reasoning of young adults. *Journal of Personality and Social Psychology*, 10, 183–201.
- *Hagar, H. O. (1990). Exploring correlates of moral orientation: The roles of dilemma content, psychological gender, and attachment history (Doctoral dissertation, University of Wisconsin—Madison, 1990). *Dissertation Abstracts International*, 51, 6104B.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 7, 119–137.
- Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde & M. C. Linn

- (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14–50). Baltimore: Johns Hopkins University Press.
- Higgins, A. (1989). The Just Community Educational Program: The development of moral role-taking as the expression of justice and care. In M. Brabeck (Ed.), *Who cares? Theory, research, and educational implications of the ethic of care* (pp. 197–215). New York: Praeger.
- Holstein, C. B. (1969). *The relation of children's moral judgment to that of their parents and to communication patterns in the family*. Unpublished doctoral dissertation, University of California, Berkeley.
- Holstein, C. B. (1976). Irreversible, stepwise sequence in the development of moral judgment: A longitudinal study of males and females. *Child Development*, 47, 51–61.
- *Hopkins-Baker, C. (1995). An examination of coping strategies and other factors associated with Gilligan's moral orientations of care and justice (Doctoral dissertation, University of Mississippi, 1995). *Dissertation Abstracts International*, 56, 3476B.
- *Hosick, J. (1994). Moral development: The effects of gender and the transition to parenthood (Doctoral dissertation, Virginia Commonwealth University, 1994). *Dissertation Abstracts International*, 55, 3035B.
- *Jadack, R. A., Hyde, J. S., Moore, C. F., & Keller, M. L. (1995). Moral reasoning about sexually transmitted diseases. *Child Development*, 66, 167–177.
- Jaffee, S., Kling, K. C., Plant, E. A., Sloan, M., & Hyde, J. S. (1999). The view from down here: Feminist graduate students consider innovative methodologies. *Psychology of Women Quarterly*, 23, 425–432.
- *Janoff, S. (1991). The impact of a legal education on moral reasoning (Doctoral dissertation, Temple University, 1991). *Dissertation Abstracts International*, 52, 3891B.
- *Janssens, J. M., & Dekovic, M. (1997). Child rearing, prosocial moral reasoning, and prosocial behaviour. *International Journal of Behavioral Development*, 20, 509–527.
- *Johnston, K. (1988). Adolescents' solutions to dilemmas in fables: Two moral orientations. In C. Gilligan, J. V. Ward, J. M. Taylor, & B. Bardige (Eds.), *Mapping the moral domain: A contribution to psychological theory and education* (pp. 49–69). Cambridge, MA: Harvard University Press.
- *Kalkoske, M. S. (1992). The internal representation of attachment and its relationship to moral orientation in middle childhood (Doctoral dissertation, University of Minnesota, 1992). *Dissertation Abstracts International*, 53, 4393B.
- Kitzinger, C., & Gilligan, C. (1994). Listening to a different voice. *Feminism and Psychology*, 4, 408–419.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 348–480). Chicago: Rand McNally.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). New York: Holt, Rinehart & Winston.
- Kohlberg, L. (1984). Moral stages and moralization: The cognitive-developmental approach. In *Essays on moral development: Vol. 2. The psychology of moral development: The nature and validity of moral stages* (pp. 170–205). San Francisco: Harper & Row.
- Kohlberg, L., & Kramer, R. (1969). Continuities and discontinuities in childhood and adult moral development. *Human Development*, 12, 93–120.
- Kohlberg, L., Levine, C., & Hower, A. (1983). *Moral stages: A current formulation and a response to critics*. New York: Karger.
- *Korschgen, A. J. (1988). Concepts of morality in administration: The role of gender, rank, and personality type (Doctoral dissertation, University of Wisconsin—Madison, 1988). *Dissertation Abstracts International*, 49, 2474A.
- *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.
- *Lamm, C. L. (1992). Children as problem solvers: Voices of care and justice in young children (Doctoral dissertation, Claremont Graduate University, 1992). *Dissertation Abstracts International*, 53, 2675A.
- *Langdale, C. J. (1986). A re-vision of structural-developmental theory. In G. L. Sapp (Ed.), *Handbook of moral development: Models, processes, techniques, and research* (pp. 15–54). Birmingham, AL: Religious Education Press.
- *Lantsberger, T. J. (1993). Factors in the decision to use a justice or a care orientation in moral reasoning (Doctoral dissertation, University of Missouri, 1993). *Dissertation Abstracts International*, 52, 3076A.
- Liddell, D. L. (1990). Measure of Moral Orientation: Construction of an objective instrument measuring care and justice with an investigation of gender differences (Doctoral dissertation, Auburn University, 1990). *Dissertation Abstracts International*, 51, 5626B.
- *Liddell, D. L. (1998). Comparison of semistructured interviews with a quantitative measure of moral orientation. *Journal of College Student Development*, 39, 169–178.
- *Liddell, D. L., & Davis, T. L. (1996). The Measure of Moral Orientation: Reliability and validity evidence. *Journal of College Student Development*, 37, 485–493.
- Liddell, D. L., Halpin, G., & Halpin, W. G. (1992). The Measure of Moral Orientation: Measuring the ethics of care and justice. *Journal of College Student Development*, 33, 325–330.
- *Liddell, D. L., Halpin, G., & Halpin, W. G. (1993). Men, women, and moral orientation: Accounting for our differences. *NASPA Journal*, 30, 138–144.
- *Logan, C. E. (1990). Message design logic and moral reasoning: A social cognitive approach to understanding messages (Doctoral dissertation, University of Washington, 1990). *Dissertation Abstracts International*, 51, 3559A.
- Lollis, S., Ross, H., & Leroux, L. (1996). An observational study of parents' socialization of moral orientation during sibling conflicts. *Merrill-Palmer Quarterly*, 42, 475–494.
- Lonky, E., Roodin, P. A., & Rybash, J. M. (1988). Moral judgment and sex role orientation as a function of self and other presentation mode. *Journal of Youth and Adolescence*, 17, 189–195.
- *Lourenco, O. M. (1991). Is the care orientation distinct from the justice orientation? Some empirical data in ten- and eleven-year-old children. *Archives de Psychologie*, 59, 17–30.
- Lyons, N. P. (1982). *Conceptions of self and morality and modes of moral choice: Identifying justice and care judgments of actual moral dilemmas*. Unpublished doctoral dissertation, Harvard University.
- *Lyons, N. P. (1983). Two perspectives: On self, relationships, and morality. *Harvard Education Review*, 53, 125–145.
- *Maqsd, M. (1998). Moral orientation of Batswana high school pupils in South Africa. *Journal of Social Psychology*, 138, 255–257.
- Mason, A. (1990). Gilligan's conception of moral maturity. *Journal for the Theory of Social Behavior*, 20, 167–179.
- Mednick, M. T. (1989). On the politics of psychological constructs: Stop the bandwagon, I want to get off. *American Psychologist*, 44, 1118–1123.
- *Miller, P. A., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology*, 32, 210–219.
- *Moore, M. L. (1988). Sex roles and moral reasoning: The relationship between moral orientation and the social domain (Doctoral dissertation, University of North Carolina at Greensboro, 1988). *Dissertation Abstracts International*, 50, 551A.
- Morris, S. B., & DeShon, R. P. (1997). Correcting effect sizes computed from factorial analysis of variance for use in meta-analysis. *Psychological Methods*, 2, 192–199.
- *Norris, D. A. (1996). Factors related to the resolution of ethical dilemmas by school psychologists (Doctoral dissertation, Arizona State University, 1996). *Dissertation Abstracts International*, 57, 1019A.
- Orchowsky, S. J., & Jenkins, L. R. (1979). Sex biases in the measurement of moral judgment. *Psychological Reports*, 44, 1040.

- *Orenstein, S. H. (1991). Parental preference in moral dilemma resolution with girls and boys (Doctoral dissertation, New York University, 1991). *Dissertation Abstracts International*, 52, 2079A.
- *Perez, M. (1994). The impact of moral development, moral orientation, gender, and religiosity on reasoning about abortion (Doctoral dissertation, Loyola College, 1994). *Dissertation Abstracts International*, 56, 1721B.
- *Peter, E., & Gallop, R. (1994). The ethic of care: A comparison of nursing and medical students. *IMAGE: Journal of Nursing Scholarship*, 26, 47-51.
- Poppen, P. (1974). Sex differences in moral judgment. *Personality and Social Psychology Bulletin*, 1, 313-315.
- *Pratt, M. W., Diessner, R., Hunsberger, B., Pancer, S. M., & Savoy, K. (1991). Four pathways in the analysis of adult development and aging: Comparing analyses of reasoning about personal-life dilemmas. *Psychology and Aging*, 6, 666-675.
- *Pratt, M. W., Golding, G., & Hunter, W. (1983). Aging as ripening: Character and consistency of moral judgment in young, mature, and older adults. *Human Development*, 26, 277-288.
- *Pratt, M. W., Golding, G., Hunter, W., & Sampson, R. (1988). Sex differences in adult moral orientations. *Journal of Personality*, 56, 373-391.
- *Pratt, M. W., & Royer, J. M. (1982). When rights and responsibilities don't mix: Sex and sex-role patterns in moral judgment orientation. *Canadian Journal of Behavioral Science*, 14, 190-204.
- Puka, B. (1989). The liberation of caring: A different voice for Gilligan's "different voice." In M. Brabeck (Ed.), *Who cares? Theory, research, and educational implications of the ethic of care* (pp. 19-44). New York: Praeger.
- Puka, B. (1991). Interpretive experiments: Probing the care-justice debate in moral development. *Human Development*, 34, 61-80.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota Press.
- *Roberson, M. K. (1988). The adolescent's decision to engage in sexual behavior (Doctoral dissertation, Michigan State University, 1988). *Dissertation Abstracts International*, 49, 4577B.
- *Rogers, A. G. (1987). Gender differences in moral reasoning: A validity study of two moral orientations (Doctoral dissertation, Washington University, 1987). *Dissertation Abstracts International*, 49, 1410B.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- *Rothbart, M. K., Hanley, D., & Albert, M. (1986). Gender differences in moral reasoning. *Sex Roles*, 15, 645-653.
- *Saloner, M. E. (1988). The contribution of a prosocial component in a social skills training program for preschoolers (Doctoral dissertation, University of Wisconsin-Madison, 1988). *Dissertation Abstracts International*, 49, 2532A.
- *Schiller, D. C. (1991). The effect of additional context on the usage of moral orientation in the measurement of moral reasoning: An examination of sex differences (Doctoral dissertation, California School of Professional Psychology, Los Angeles, 1991). *Dissertation Abstracts International*, 52, 2784B.
- *Schlesinger, M. (1987). Gender differences in moral stage, moral orientation, and sex role identity of academic administrators: A comparison of Kohlberg's and Gilligan's theories (Doctoral dissertation, Temple University, 1987). *Dissertation Abstracts International*, 50, 3192A.
- *Self, D. J., Skeel, J. D., & Jecker, N. S. (1993). A comparison of the moral reasoning of physicians and clinical medical ethicists. *Academic Medicine*, 68, 852-855.
- Sichel, B. A. (1985). Women's moral development in search of philosophical assumptions. *Journal of Moral Education*, 14, 149-161.
- *Sitzer, D. L. (1990). A theoretical and empirical linkage between moral and personality development: A study correlating narcissism, affective empathy, and the justice and care moral orientations (Doctoral dissertation, California School of Professional Psychology, Los Angeles, 1990). *Dissertation Abstracts International*, 51, 5042B.
- *Skoe, E. E., & Diessner, R. (1994). Ethic of care, justice, identity, and gender: An extension and replication. *Merrill-Palmer Quarterly*, 40, 272-289.
- *Skoe, E. E., & Gooden, A. (1993). Ethic of care and real-life moral dilemma content in male and female early adolescents. *Journal of Early Adolescence*, 13, 154-167.
- Skoe, E. E., & Marcia, J. E. (1991). A measure of care-based morality and its relation to ego identity. *Merrill-Palmer Quarterly*, 37, 289-304.
- *Skoe, E. E., Pratt, M. W., Matthew, M., & Curror, S. E. (1996). The ethic of care: Stability over time, gender differences, and correlates in mid- to late adulthood. *Psychology and Aging*, 11, 280-292.
- Smetana, J. G. (1984). Morality and gender: A commentary on Pratt, Golding, and Hunter. *Merrill-Palmer Quarterly*, 30, 341-348.
- *Söchting, I., Skoe, E. E., & Marcia, J. E. (1994). Care-oriented moral reasoning and prosocial behavior: A question of gender or sex role orientation. *Sex Roles*, 31, 131-147.
- *Stander, V., & Jensen, L. (1993). The relationship of value orientation to moral cognition: Gender and cultural differences in the United States and China explored. *Journal of Cross-Cultural Psychology*, 24, 42-52.
- *Stiller, N. J., & Forrest, L. (1990). An extension of Gilligan's and Lyon's investigation of morality: Gender differences in college students. *Journal of College Student Development*, 31, 54-63.
- *Stookey, K. (1994). Gender, gender-roles, and moral judgments: An investigation of the theories of Kohlberg and Gilligan (Doctoral dissertation, University of Alabama, 1994). *Dissertation Abstracts International*, 56, 2178A.
- Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review*, 6, 165-180.
- *Tietjen, A. M. (1986). Prosocial reasoning among children and adults in a Papua New Guinea society. *Developmental Psychology*, 22, 861-868.
- Time 25. (1996, June 17). *Time*, 147, 52-79.
- Tronto, J. C. (1987). Beyond gender difference to a theory of care. *Signs: Journal of Women in Culture and Society*, 12, 644-663.
- Turiel, E. (1976). A comparative analysis of moral knowledge and moral judgment in males and females. *Journal of Personality*, 44, 195-208.
- Van Gelder, L. (1984, January). Carol Gilligan—Leader for a different kind of future. *Ms.*, 12, 37-40.
- *Vera, E. M., & Levin, I. P. (1989, May). *Sex differences in moral decision making*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.
- Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. *Child Development*, 55, 677-691.
- *Walker, L. J. (1986a). Experimental and cognitive sources of moral development in adulthood. *Human Development*, 29, 113-124.
- Walker, L. J. (1986b). Sex differences in the development of moral reasoning: A rejoinder to Baumrind. *Child Development*, 57, 522-526.
- Walker, L. J. (1989). A longitudinal study of moral reasoning. *Child Development*, 60, 157-166.
- Walker, L. J. (1991). Sex differences in moral reasoning. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Vol. 2. Research* (pp. 333-364). Hillsdale, NJ: Erlbaum.
- Walker, L. J. (1995). Sexism in Kohlberg's moral psychology? In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral development: An introduction* (pp. 83-107). Boston: Allyn & Bacon.
- Walker, L. J. (1997). Is morality gendered in early parent-child relationships? A commentary on the Lollis, Ross, and Leroux study. *Merrill-Palmer Quarterly*, 43, 148-159.
- *Walker, L. J., de Vries, B., & Trevethan, S. D. (1987). Moral stages and moral orientations in real-life and hypothetical dilemmas. *Child Development*, 58, 842-858.
- *Walker, L. J., & Moran, T. J. (1991). Moral reasoning in Communist Chinese society. *Journal of Moral Education*, 20, 139-155.
- *Wark, G. R. (1993). The effects of sex, sex-role, and type of moral dilemma on moral maturity and moral orientation (Masters thesis, Simon Fraser University, 1993). *Masters Abstracts International*, 33, 652.

- *Wark, G. R., & Krebs, D. L. (1996). Gender and dilemma differences in real-life moral judgment. *Developmental Psychology*, 32, 220-230.
- *Wark, G. R., & Krebs, D. L. (1997). *Sources of variation in real-life moral judgment: Toward a new model of morality*. Manuscript submitted for publication, Simon Fraser University, Burnaby, British Columbia, Canada.
- *Webster, P. A. (1996). Moral orientation, ego maturity, and religious commitment: An exploratory study from a developmental perspective (Doctoral dissertation, Andrews University, 1996). *Dissertation Abstracts International*, 57, 1982A.
- *White, J. (1994). Individual characteristics and social knowledge in ethical reasoning. *Psychological Reports*, 75, 627-649.
- *White, J., & Manolis, C. (1997). Individual differences in ethical reasoning among law students. *Social Behavior and Personality*, 25, 19-48.
- *Wilson, F. L. (1995). The effects of age, gender, and ethnic/cultural background on moral reasoning. *Journal of Social Behavior and Personality*, 10, 67-78.
- *Yacker, N., & Weinberg, S. L. (1990). Care and justice moral orientation: A scale for its assessment. *Journal of Personality Assessment*, 55, 18-27.

Received August 25, 1997

Revision received March 5, 2000

Accepted March 5, 2000 ■

Call for Nominations

The Publications and Communications Board has opened nominations for the editorships of *Journal of Applied Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Educational Psychology*, *Psychological Bulletin*, and *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes* for the years 2003-2008. Kevin R. Murphy, PhD, Philip C. Kendall, PhD, Michael Pressley, PhD, Nancy Eisenberg, PhD, and Chester A. Insko, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2002 to prepare for issues published in 2003. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

To nominate candidates, prepare a statement of one page or less in support of each candidate and send to

- Margaret B. Spencer, PhD, for the *Journal of Applied Psychology*
- Donna M. Gelfand, PhD, and Lucia Albino Gilbert, PhD, for the *Journal of Consulting and Clinical Psychology*
- Lauren B. Resnick, PhD, for the *Journal of Educational Psychology*
- Janet Shibley Hyde, PhD, and Randi C. Martin, PhD, for *Psychological Bulletin*
- Sara B. Kiesler, PhD, for *JPSP: Interpersonal Relations and Group Processes*

Address all nominations to the appropriate search committee at the following address:

[Name of journal] Search Committee
c/o Karen Sellman, P&C Board Search Liaison
Room 2004
American Psychological Association
750 First Street, NE
Washington, DC 20002-4242

The first review of nominations will begin December 11, 2000.