

Structural Flexibility of Moral Judgment

Dennis L. Krebs, Kathy L. Denton, Sandra C. Vermeulen, Jeremy I. Carpendale, and Alice Bush
Simon Fraser University, Burnaby, British Columbia, Canada

One of the central assumptions of Kohlberg's theory of moral development—that moral judgment is organized in structures of the whole—was examined. Thirty men and 30 women were given 2 dilemmas from Kohlberg's Moral Judgment Interview, a 3rd involving prosocial behavior, and a 4th involving impaired driving. Half the Ss responded to the prosocial and impaired-driving dilemmas from the perspective of a hypothetical character, and half responded from the perspective of the self. No sex or perspective differences in moral maturity were observed. Ss scored highest in moral maturity on Kohlberg's dilemmas, intermediate on the prosocial dilemma, and lowest on the impaired-driving dilemma. In partial support of Kohlberg's contention that his test assesses moral competence, there was a negative linear relationship between scores on his test and the proportion of Stage 2 judgments on the 2 other dilemmas. An interactional model of moral judgment is advanced.

Kohlberg's theory of moral development is based on two main assumptions. The first—that stages change in an invariant sequence—has received the most attention. The purpose of this study was to investigate the second assumption, that moral judgment is organized in "structures of the whole" such that "under conditions that support expression of the individual's most mature moral thinking, his or her reasoning will form a coherent system best described by one of Kohlberg's five stages or a mixture of at most two adjacent stages" (Kohlberg & Colby, 1983, p. 120). On the structure-of-the-whole assumption, people are expected to base most of their moral judgments to most moral issues on the stage structures they display on Kohlberg's test. This "horizontal" assumption is inextricably bound to the "vertical" assumption that as individuals develop, "higher stages displace (or, rather, integrate) the structures found at lower stages" (Colby & Kohlberg, 1987, p. 7).

Of course, no stage theory expects all individuals to base all their judgments on one stage structure (see Flavell, 1982). Because vertical change from old to new stages does not occur all at once, individuals are expected to base their judgments on the structures of both the stage they are in and the stage they are moving toward during periods of transition. Colby and Kohlberg (1987, p. 8) also acknowledge another source of heterogeneity—"performance factors." In "situations with a significant downward press" such as "the low-level 'moral atmosphere' of a traditional prison," individuals may base their moral judgments on stages lower than the ones they display on Kohlberg's test. Indeed, Colby and Kohlberg state that the "stage properties" of Kohlberg's theory, including the assumption of struc-

ture wholeness, "characterize competence though not necessarily performance in moral judgment" (p. 8). Although all measures of competence are based on tests of performance, Colby and Kohlberg argue that Kohlberg's test minimizes the gap between competence and performance "by using hypothetical dilemmas, by using probing questions that attempt to elicit the upper limits of the subject's thinking, and by . . . scoring rules according to which only the most mature expressed version of a particular moral idea is scored" (p. 7).

Other cognitive-developmental theorists have advanced models that allow for more contextual impact on social judgment than Kohlberg allows, and that, therefore, posit significantly less stage consistency. For example, Selman (1980) contends that the levels of perspective-taking individuals invoke are influenced by the demands of the tasks they face. Damon (1977) cautions, "We must not overlook the extent to which the child tailors his or her behavior appropriately to meet the specific demands of different social contexts" (p. 46). Rest (1983) offers a "layer-cake" model of moral development in which new stages are built on old stages, which are retained. Levine (1979) advances an "additive-inclusive" model of moral development that is based on the assumption that "higher stages include components of earlier stages but do not replace these stages" (p. 155) and "that the stability or variability of moral reasoning should be understood as a case of 'best fit' between one of several equilibrated moral structures, person characteristics, and recurring patterns of environmental stimuli" (p. 156).

The primary evidence adduced by Colby and Kohlberg (1987) in support of the structural consistency of moral judgment stems from the pattern of judgments made on their test by the subjects in their longitudinal study. As scored and counted by Colby, Kohlberg, and their colleagues (1987), 67% of the judgments were at the same stage and 99% were at adjacent stages. Only "9% [of the subjects] . . . show a third stage of reasoning greater than 10%" (p. 9). Establishing that moral judgment on Kohlberg's test is structurally homogeneous, however, is only a first step toward confirming the assumption that moral thought is organized in structures of the whole. Gilligan

This research was supported by Grants 410-87-1115 and 410-89-0649 from the Social Sciences and Humanities Research Council of Canada. The assistance of Marianne Schroeder, Michael Taylor, and Diana Doherty in early stages of the project and of Nancy Higgins and Robyn So in the later stages is gratefully acknowledged.

Correspondence concerning this article should be addressed to Dennis L. Krebs, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6.

(1982), Eisenberg (1982), and others have argued that Kohlberg's dilemmas neglect the prosocial and "care" side of morality and that prosocial and care-oriented moral judgments may be structured differently from prohibitive and justice-oriented moral judgments. Baumrind (1978) and Damon (1977) criticize Kohlberg's dilemmas for failing to reflect the types of moral conflicts people experience in their everyday lives. Baumrind (1978), Eisenberg (1986), Leming (1978), and others criticize Kohlberg's dilemmas for evoking third-person judgments about the moral obligations of hypothetical characters in extraordinary dilemmas, as opposed to first-person judgments about the self's moral obligation in the types of dilemma people usually encounter in their everyday lives.

More than a dozen studies have examined the consistency between stage use on Kohlberg's test and stage use on other types of moral dilemma. In eight studies (Gilligan, Kohlberg, Lerner, & Belenky, 1971; Gilligan & Belenky, 1980; Haan, 1975; Higgs, 1975; Kohlberg, Scharf, & Hickey, 1972; Leming, 1978; Lockwood, 1975; Smetana, 1982) researchers used outdated versions of Kohlberg's test and scoring system—versions that correlate only weakly with the present system (see Candee & Kohlberg, 1987, for a rescoring of the Haan, 1975, data). Among the studies in which Kohlberg's current scoring system were used, three—a study investigating care and responsibility in alternative high schools (Higgins, Power, & Kohlberg, 1984), a study investigating Israeli soldiers' justifications for refusing to fight in the Lebanon War (Linn, 1987a), and a study investigating actual moral conflicts (Walker, de Vries, & Trevethan, 1987)—supported Kohlberg's assumption that moral judgment is structurally homogeneous. The remaining two studies, involving decisions about day care (Linn, 1984) and an actual strike staged by 50 Israeli physicians (Linn, 1987b), found that subjects scored significantly lower on the non-Kohlberg dilemmas than on Kohlberg's dilemmas. Higgins, Power, & Kohlberg also found that students from traditional high schools scored lower on care and responsibility dilemmas. Reviewing the research on the consistency of moral judgment, Walker (1988) concludes, "Some research has indicated that occasionally some characteristics of dilemmas elicited varying amounts of reasoning at different stages . . . [but] the weight of the research evidence supports the viability of the structure criterion in that variability is minimal" (p. 42).

The central purpose of this study was to evaluate the structural homogeneity of moral judgment in adults by examining the consistency between stage scores on classic Kohlberg dilemmas and stage scores on two more familiar types of moral dilemma, one involving prosocial behavior (something people should do) and one involving prohibited behavior (something people should not do), using the new Colby and Kohlberg (1987) scoring system. The first question pertained to the capacity of Kohlberg's scoring system to classify moral judgments to dilemmas other than those on Kohlberg's test, especially those that are prosocial in orientation. Eisenberg (1982) and Gilligan (1982) argue that Kohlberg's system neglects and misclassifies prosocial morality. The second question concerned sex differences in moral judgment. Gilligan claims that women tend to score lower than men do on Kohlberg's test. A meta-analysis of relevant research (Walker, 1984) failed to reveal any consistent sex differences on Kohlberg's test. Walker's (1984) conclusions

were questioned by Baumrind (1986) and defended by Walker (1986). In this study, we examined sex differences on prohibitive as well as care-oriented moral dilemmas.

The third, and main, question concerned the structural consistency of moral judgment (a) between the dilemmas on Kohlberg's test, (b) between the Kohlberg dilemmas and the prosocial and prohibitive dilemmas, and (c) within each dilemma. The final question concerned the effect of a performance factor expected to affect the level of moral maturity on the prosocial and prohibitive moral dilemmas, that is, the social perspective from which they were viewed. Subjects were expected to invoke different stage structures when responding to dilemmas from a first-person perspective than when responding to dilemmas from a third-person perspective. In particular, we expected self-serving cognitive biases and prudential motives (see Krebs, Denton, & Higgins, 1988; Krebs, Vermeulen, Carpendale, & Denton, in press; C. R. Snyder & Higgins, 1988) to render first-person moral judgments less impartial and less objective than the third-person judgments on Kohlberg's test. The results of studies in which researchers used tests of moral judgment other than Kohlberg's test are consistent with these expectations (Eisenberg-Berg & Neal, 1981; Damon, 1977; Weiss, 1982).

Method

Subjects

The subjects were 30 men (age $M = 26.6$ years, $SD = 6.3$ years) and 30 women (age $M = 27.16$ years, $SD = 5.3$ years) between the ages of 17 and 40 who were recruited randomly from a large university and the surrounding community through the subject file of the psychology department and advertisements in the community. Subjects were paid \$5 for their participation.

Procedure

A trained interviewer read subjects either dilemmas III and III' from Form A of Kohlberg's Moral Judgment Interview (Should Heinz steal a drug to save his wife? Should Officer Brown report Heinz?) or dilemmas IV and IV' from Form B (Should Dr. Jefferson perform euthanasia? Should Dr. Rogers report him?) and then asked a series of probing questions in accordance with the procedure outlined by Colby and Kohlberg (1987, pp. 230–235). After the subjects responded to the dilemmas from Kohlberg's test, they were read a prosocial dilemma patterned after an incident investigated by McNamee (1978) and discussed at length by Kohlberg and Candee (1984, pp. 520–523) and a dilemma involving impaired driving that was patterned after incidents investigated by Denton and Krebs (1990). Half the subjects responded to the prosocial dilemma first; the other half responded to the impaired-driving dilemma first.¹

In the prosocial dilemma, a student who is participating in a psycho-

¹ Kohlberg's test was always given first to ensure a pure Kohlberg score, unconfounded by the perspective manipulation on the prosocial and impaired-driving dilemmas. Although it is possible that fatigue diminished moral maturity on the prosocial and impaired dilemmas, it seems improbable because (a) we have counterbalanced the order of different forms of Kohlberg's test in other studies and have found no hint of an effect (see Krebs, Vermeulen, Carpendale, & Denton, in press) and (b) there was no effect for order on the prosocial and impaired-driving dilemmas, which were given last.

logical experiment observes another student, who is sick from illicit drugs, approach the experimenter in search of help. The experimenter replies that she is not a clinical psychologist and cannot help. The question is, what should the first student do, continue to fulfill his or her contract with the experimenter or help the person in need? The dilemma was followed by a set of nine probe questions comparable to those on Kohlberg's test—questions such as, "What should he or she do, why?" "Were any moral issues raised in this situation, what makes them moral?" "What was each person's responsibility?" "In general should people help one another, why?" and "What if [the student sick from drugs] had taken prescription drugs, would that make a difference?"

The impaired-driving dilemma was selected as a prohibitive moral dilemma familiar to most adults and salient in the media (see Calvert-Boyanowsky & Boyanowsky, 1981). It read simply:

A person named Jack is out drinking with his friends. He doesn't keep track of exactly how much he drinks, but, when it comes time to go home, he senses that he has had more to drink than the legal limit. His car is outside.

This dilemma was followed by a set of 11 probe questions comparable to those on Kohlberg's test (e.g., "What should Jack do?" "Are there circumstances under which it is right to drive impaired. . . ?" "Should people be held responsible for their actions when they've been drinking?")

Half the subjects,² those in the *other* condition, were given the prosocial and impaired-driving dilemma in the same format as Kohlberg's test: They were asked to make judgments about a hypothetical protagonist, as indicated above. The other half of the subjects, those in the *self* condition, were asked to imagine themselves as the protagonist in the dilemma (e.g., the impaired-driving dilemma opened, "Imagine you are out drinking with your friends") and were asked questions about what they should do and why. The interviewers probed subjects' answers extensively, as outlined by Colby and Kohlberg (1987). The average time per interview was approximately 1½ hr. All interviews were tape-recorded and transcribed for scoring.

Scoring Kohlberg Dilemmas

The 17-step procedure for scoring Kohlberg's test is complex, exacting, and outlined in a two-volume, 1,200-page scoring manual (Colby & Kohlberg, 1987). In this study, moral judgments were scored over 4 months by a team of scorers trained by Krebs and another experienced scorer. In essence, scoring involves classifying prescriptive interview judgments by issue, norm, and element and finding matching criterion judgments in the Colby and Kohlberg scoring manual that specify their stage structure. For example, the chosen issue for the interview judgment, "Heinz should steal the drug because the guy was just trying to rip him off anyway" is life, because the subject advocates stealing a drug to save a life; the norm is property, because drugs and money are being valued; the element is retributing (getting even); the stage is 2, and the matching criterion judgment in the Colby and Kohlberg manual is "[Heinz should steal the drug] because the druggist was asking for it. . ." (p. 16).

After stage scores have been assigned to all scorable interview judgments, the scores are weighted and summed to produce either a global stage score or a weighted average score (also called a moral maturity score). The method of calculating global stage scores is specified by Colby and Kohlberg (1987, pp. 185–186). It involves deriving a representative stage score on each of the moral issues on Kohlberg's test, assigning weights (chosen issues \times 3, nonchosen issues \times 2, guess scores \times 1), summing the weighted scores, and transforming the product into a stage score on a 9-point scale (Stage 1 followed by Stage 1/2, Stage 2, etc.). We will call the stages on the 9-point scale *substages* to

distinguish them from global stages on a 5-point scale. Moral maturity scores are arrayed on a continuous scale ranging from 100 (corresponding to Stage 1) to 500 (corresponding to Stage 5). They are calculated by summing weighted issue stage scores, dividing by the sum of the weights, and multiplying by 100 (see Colby and Kohlberg, 1987, pp. 187–188).

Scoring the Prosocial and Impaired-Driving Dilemmas

In any study comparing stage scores on Kohlberg's test with stage scores on other dilemmas, it is imperative that the other dilemmas are scored validly. In this study, scorers identified judgments to the prosocial and impaired-driving dilemmas that were prescriptive in nature (*should* judgments), classified the prescriptive interview judgments by element, then perused the criterion judgments in the Colby and Kohlberg (1987) manual based on the element in search of a structural match. In cases in which a match could not be found, scorers searched elsewhere in the manual for criterion judgments that were similar in structure but had different elements. (All elements are not represented at all stages in the Colby & Kohlberg, 1987, manual.) To illustrate, consider the prosocial judgment, "you should help other people because it makes you feel good about yourself." Although this judgment relates to an issue (altruism) that is different from the criterion judgments in Colby and Kohlberg's scoring manual, it is similar in structure to Criterion Judgment 19, Form A, Contract: "[It is important to keep a promise] because it makes a person feel good inside." The two judgments are based on the same norm (conscience), element (upholding self-respect), and stage structure ("anticipation of approval from oneself if one lives up to conventional role expectations"; Colby & Kohlberg, 1987, pp. 210–211).

To obtain a measure of interrater reliability, we randomly selected 25% of the interviews (60 dilemmas) and rescored them independently, blind to other stage scores. Scorers agreed on 12 of 15 (80%) of the stage assignments on a 9-point scale for both the prosocial and impaired-driving dilemmas ($r_s = .90$ and $.80$), compared with 13 of 15 (87%) for the first dilemma on Kohlberg's test and 14 of 15 (93%) for the second dilemma on Kohlberg's test ($r_s = .87$ and $.91$). There were no cases in which a discrepancy in stage score was more than a minor stage. All differences were resolved by a third scorer. Scores obtained for matched judgments were converted into global stage scores and moral maturity scores in accordance with the Colby and Kohlberg (1987) scoring procedure.

Criteria of Structural Consistency and the Homogenizing Effects of Colby and Kohlberg's Scoring System

Because individuals are expected to base their moral judgments on the stage they are in as well as the stage they are moving toward during periods of transition, Colby and Kohlberg (1987, p. 90) set as their criterion for structural consistency adjacent stages on a 5-point scale. We consider this criterion too liberal. As pointed out by Cortese (1984), "With any system that effectively discriminates only a small number of categories, 'consistency' will be artificially elevated" (p. 241). Eisenberg (1986, pp. 180–183) makes a similar point. Assume that an individual takes two tests of moral judgment and is assigned one of Kohlberg's five stages on each. In all, 13 of the possible 25 combinations (1–1, 1–2, . . . 5–5) meet Colby and Kohlberg's "same or adjacent stage" criterion. Thus, the probability of scoring at the same or an adjacent stage by chance is more than .50. Considering only the stages at which most adults are classified—Stages 2, 3, and 4—the probability of meeting

² Because of a classification error, 32 subjects actually ended up in the self condition and 28 in the other condition.

the adjacent stage criterion is 9 of 15 or .60. Clearly, this criterion is too liberal, and it is inconsistent with the distinctions made by Colby and Kohlberg in their latest work: They classify the criterion judgments in their scoring manual on a 9-point scale and they use a 9-point scale when they report the results of their longitudinal study (see, e.g., Colby & Kohlberg, 1987, pp. 84–89). The 9-point scale differs from the 5-point scale because it includes transitional judgments (e.g., Stage 1/2 or 2/3). If moral judgment is organized in structures of the whole across varying content, individuals who are in transition on one dilemma should be in transition on another. Therefore, we set as our criterion for structural consistency the same or an adjacent stage on a 9-point scale.

It is important to note that even stage scores on a 9-point scale, as calculated in the Colby and Kohlberg (1987) scoring system, may mask underlying heterogeneity. In their scoring guidelines, Colby and Kohlberg instruct interviewers to elicit the highest stage available to interviewees: "We have attempted to minimize the gap between competence and performance by using . . . probing questions that attempt to elicit the upper limits of the subject's thinking" (p. 7). Low-stage judgments that are upgraded in response to probes are not assigned stage scores and are omitted in the calculation of global stages, according to the rule of inclusion (see Colby & Kohlberg, 1987, p. 177). In addition, low-stage interview judgments that have been assigned stage scores but that are similar in content to judgments scored at higher stages are eliminated: "Whenever interview material matches two CJs [criterion judgments] of the same norm and element but different stages . . . only the score of the highest match on the norm and element is included when assigning a score on the issue" (pp. 177–178). The upshot of using "scoring rules according to which only the most mature expressed version of a particular moral idea is scored" (p. 7) is to eliminate the less mature, heterogeneous, judgments.

Colby and Kohlberg's (1987) scoring system also eliminates heterogeneous stage responses at three different junctures in deriving global stage scores. First, when summing the stage scores of particular judgments to obtain mean stage scores for issues, scorers are instructed to drop anomalous scores unless they constitute 25% or more of the scored interview judgments. The stage represented by the highest percentage of judgments, whatever that percentage may be, is designated the major stage and is weighted accordingly. Second, when summing issue scores to obtain global stage scores, issue scores that fail to represent 25% of the total are discarded. Finally, scores on nonchosen issues are weighted only two thirds as much as scores on chosen issues, implicitly acknowledging that people may make higher level judgments when arguing in support of their position than when arguing against it (see de Vries & Walker, 1986; Nisan & Koriati, 1989).

Although issue scores that fall below 25% of the total are retained in the calculation of moral maturity scores, this measure may mask heterogeneity in another way. To illustrate, an individual making half of his or her judgments at Stage 2 and half at Stage 4 could receive a moral maturity score (300) identical to an individual making all his or her judgments at Stage 3.

Results and Discussion

The results are presented and discussed in four main sections, dealing with (a) the ability of Kohlberg's scoring system to classify judgments to the prosocial and impaired driving dilemmas, (b) the effect of sex, dilemma, and social perspective on moral maturity, (c) within-subject stage consistency across dilemmas, and (d) sources of inconsistency in stage use.

Structural Classification of Prescriptive Judgments from Prosocial and Impaired-Driving Dilemmas

The first finding of the present study was that subjects made prescriptive judgments to the prosocial and impaired-driving

dilemmas: Subjects considered the issues in the dilemmas moral in nature. The second finding was that virtually all prescriptive judgments to the prosocial and impaired-driving dilemmas could be matched reliably with one of the criterion judgments in Colby and Kohlberg's (1987) scoring manual. Although the content of moral judgments to the prosocial and impaired-driving dilemmas differed from the content of moral judgments to the dilemmas on Kohlberg's test (and therefore from the content of the criterion judgments in Colby and Kohlberg's 1987 scoring manual), parallels in the underlying structure of the two sets of judgments enabled scorers to classify them in accordance with Kohlberg's scoring system. A sample of moral judgments to the prosocial and impaired driving dilemmas and the criterion judgments from the Colby and Kohlberg manual with which they were matched is displayed in Table 1.

We and several other researchers (see Walker, 1988) have found that trained scorers can reliably match moral judgments from a variety of dilemmas for structure with the criterion judgments derived from moral judgments to Kohlberg's test: Individuals trained to identify Kohlberg's structures can find enough of them in judgments to other moral dilemmas to derive stage scores. It is, however, possible that scorers versed in Kohlberg's system overlook or misclassify moral judgments that, from another perspective, would be viewed in different terms. For example, scorers might neglect or misclassify prosocial and care-oriented moral judgments, as Eisenberg (1982) and Gilligan (1982) claim. We consider this possibility improbable for three reasons.

First, as pointed out by Nunner-Winkler (1984), the claim that Kohlberg's dilemmas deal exclusively with prohibitive and justice-oriented issues is incorrect. In fact, Kohlberg's classic dilemmas involve conflicts between prohibitive and prosocial choices: for example, should Heinz violate the prohibition against stealing to help his wife? Virtually all judgments in Colby and Kohlberg's scoring manual that uphold the value of life advocate helping someone in need. Thus, Colby and Kohlberg's (1987) scoring manual contains many prosocial and care-oriented criterion judgments—judgments such as "if he [Heinz] really loves her [his wife], it would be right [to help her] from his viewpoint" (p. 8). Indeed, Eisenberg (1986, p. 147) reports scoring judgments to Kohlberg's dilemma for her categories of prosocial moral reasoning, and Krebs and Denton (1990) have shown that all of the judgments in Eisenberg's classification of prosocial moral judgment can be classified reliably for stage in Kohlberg's system; that is, matches can be found for all of them among the criterion judgments in Colby and Kohlberg's scoring manual.

Finally, both Eisenberg (1986) and Gilligan (Gilligan & Attanucci, 1988) have moderated or relinquished whatever claims they may have advanced about discovering "hard" stages that differ from those of Kohlberg and his colleagues. The emphasis in Gilligan's recent work is on the difference between rights-based and care-based moral *orientations* (see Gilligan & Attanucci, 1988), not stages. And Eisenberg acknowledges (a) that "it is possible that individuals' responses to prosocial and Kohlberg dilemmas are more similar than the existing research would indicate" (p. 147), (b) that her categories of prosocial moral reasoning correspond more closely to Kohlberg's ele-

Table 1

Selected Responses to the Kohlberg, Prosocial, and Impaired-Driving Dilemmas

Subject	Kohlberg	Prosocial	Impaired driving
170	In terms of society, yes, because society has laid down a list of laws and regulations by which society is supposed to be run by, and if the laws are changed for certain people then that just breaks open the whole issue of if all the laws should be changed. You know, you have to draw the line. (Form A, Punishment, CJ 26, Stage 4, p. 170)	[In general should people help each other?] Yes, to some extent, I guess . . . I mean you may not think you do at the time [need help in the future] but I think there always is a time that arises where you are going to have to ask somebody for help and you would like them to give you the same consideration. (Form A, Contract, CJ 10, Stage 2/3, p. 202)	[Should Jack drive?] No, if by chance he gets into an accident or something, it'll be his fault and he could face a lot of repercussions. (Form B, Law, CJ 7, Stage 2, p. 334)
872	She's made all her decisions before and possibly this is the most important decision a person has to make. She has the right . . . her right to life or right not to life as she so chooses, in exercising her free will supercedes the law in that case. (Form B, Life, CJ 29, Stage 5, p. 322)	She has a responsibility, especially in her capacity as a professional . . . in that role she should take the guy and get some help for him. (Form B, Law, CJ 10, Stage 3, p. 300)	No, I should not [drive], because I may kill somebody or kill myself, or kill my passengers, or have an accident or do something. (Form B, Law, CJ 10, Stage 2/3, p. 347)
240	[Heinz should steal the drug] he's the one who's got everything at stake and perhaps his wife's going to die. He's going to be left all alone. (Form A, Life, CJ 7, Stage 2, p. 17)	You're not morally obligated to help him, but you should feel an obligation. I would feel an obligation to help a friend. (Form A, Life, CJ 13, Stage 3, p. 26)	Thank God the law is there trying to stop us from hurting other people. That's really what we're talking about here. I guess in a sense they are trying to protect us from ourselves in that circumstance anyway, but they're [the lawmakers] basically trying to protect everybody else from causing an accident or running over people or damaging property. (Form A, Punishment, CJ 11, Stage 2/3, p. 156)
112	It's against the law to steal drugs so that's the only reason he would do it is for the virtue of his love. (Form A, Life, CJ 12, Stage 3, p. 23)	[Jack should help] Just out of humanity, for the sake of responsibility to fellow man. (Form A, Life, CJ 26, Stage 4, p. 41)	[Should the probability that Jack will get stopped in a roadblock effect his decision to drive?] Yes, because he's going to get a big sentence and a fine and it's not worth it. (Form A, Law, CJ 8, Stage 2, p. 70)
516	[In general should people obey the law?] In general, yes, because the laws have been created by the will of society to help people within the society interact with the least possible interference and discomfort. (Form A, Law, CJ 26, Stage 4, p. 90)	If I help you, you are more likely to help me in some future date, should I need help. (Form A, Contract, CJ 6, Stage 2, p. 198)	I don't want to get hit with a drunk driving charge and the consequences that that entails. (Form B, Law, CJ 7, Stage 2, p. 344)
315	Considering she is married, her husband would most likely be the closest person to her and should be allowed a say in the decision. (Form B, Law, CJ 12, Stage 3, p. 350)	[Yes, people should help each other] because it makes for an even more orderly culture or society. (Form A, Life, CJ 24, Stage 4, p. 38)	I think that if someone is found to have a blood alcohol level over the limit there they should be charged criminally and held criminally liable . . . because there are so many cars on the road and so many human bodies involved, it would be the most sensible thing if the person did not drink and drive. (Form A, Punishment, CJ 11, Stage 2/3, p. 156)
414	[Should the doctor be given the death penalty?] I don't feel society would be served properly to remove a person that has a lot to give society with his skills where this issue isn't clear cut. (Form B, Morality, CJ 33, Stage 4, p. 430)	[In general should people help each other?] If you were asked to help someone, and they die, sure the dilemma is finished for the person that died, but now you must deal with the fact that you might have been able to prevent that death. (Form A, Life, CJ 17, Stage 3, p. 30)	[Are there any circumstances under which Jack should drive?] I say no, because there's a chance that you might kill yourself in the process of trying to help someone. (Form A, Law, CJ 8, Stage 2, p. 70)

Table 1 (continued)

Subject	Kohlberg	Prosocial	Impaired driving
764	[Should the doctor be given the death penalty?] What he was doing was at the insistence of the woman who wanted the injection. I don't feel it's murder in that sense because she wanted to die. (Form B, Morality, CJ 5, Stage 2, p. 401)	I think if somebody is in such desperate straits and if they feel totally out of control and they're asking for help, one should help them. (Form B, Life, CJ 11, Stage 3, p. 300)	I would be a hazard on the road, not only to myself but to other people and pedestrians and I would rather not risk that. (Form B, Law, CJ 10, Stage 2/3, p. 347)
888	[Thinking in terms of society, should people who break the law be punished?] [Yes] It seems to me what serves society better is an effort . . . to ensure that that individual won't be committing that act of crime, in other words, it's a matter of protecting society from acts of crime. (Form B, Punishment, CJ 40, Stage 4, p. 502)	Yes, because all the benefits that we collectively receive from our society depend on our willingness to live together and that includes the responsibility to help where we can. (Form B, Contract, CJ 48, Stage 5, p. 567)	The law has drawn the line and I think that's a good guideline, not simply because that's the law, but I presume that the law drew the guideline where it is after thinking carefully about the question of safety. (Form A, Law, CJ 26, Stage 4, p. 90)

Note. CJ = criterion judgment.

ments of morality than to Kohlberg's stages (p. 181), and (c) that both prosocial and prohibitive moral judgments are structured in terms of "a self-focused or egoistic orientation" at the lowest stages, "an orientation to the approval of and expectations of others and the adherence to global conceptions of societal norms and values" at intermediate stages, and "concern with abstract ethical principles, the social system, imperatives of conscience, and meeting one's obligations" at the highest stages (p. 145).

In conclusion, the results of this and other studies (see Krebs, Vermeulen, Carpendale, & Denton, in press) suggest that moral judgments to a wide range of moral issues are structured in terms of Kohlberg's stages. Although it is possible that someone will discover other stages, and although Kohlberg's stage structures may well be refined and elaborated, the bulk of evidence supports the conclusion that Colby and Kohlberg's (1987) scoring system is limited more in its ability to classify the content of moral judgments to dilemmas other than those on Kohlberg's test than in its ability to score such judgments for stage structure. The work of Eisenberg, Gilligan, and their colleagues seems valuable more in the refinements it suggests in the classification of prosocial and care-oriented content than in its ability to identify new moral structures.

Effect of Sex, Dilemma, and Social Perspective on Moral Maturity

Establishing that moral judgments about helping and impaired driving are structured in terms of Kohlberg's stages does not establish that subjects invoke the same stage in response to different dilemmas, as they should if moral judgment is organized in structures of the whole. As mentioned earlier, most of the evidence adduced by Colby and Kohlberg (1987) for the structural homogeneity of moral judgment stems from stage consistency on the dilemmas of Kohlberg's test. Consistent with this evidence, the mean moral maturity displayed by sub-

jects on the first dilemma of Kohlberg's test (324) did not differ significantly from their mean moral maturity on the second dilemma (327), $t(59) = .28$. To facilitate comparisons with moral maturity on the prosocial and impaired-driving dilemmas, we combined scores on the two Kohlberg dilemmas in accordance with the instructions in the Colby and Kohlberg scoring manual, and we used this score in subsequent analyses.

A $2 \times 2 \times 3$ analysis of variance (ANOVA; Form \times Sex \times Dilemma), with repeated measures on the last factor and moral maturity as the dependent variable, failed to reveal any significant main effects or interactions for form, $F(1, 54) = 2.09$, or for sex, $F(1, 54) = .42$. There was, however, a highly significant main effect for dilemma, $F(1, 108) = 48.58$, $p < .001$. Between-groups comparisons corrected for experimentwise error revealed that the impaired-driving dilemma evoked a significantly lower level of moral maturity (247) than the prosocial dilemma (284), $t(57) = 5.17$, $p < .01$, and the prosocial dilemma evoked a significantly lower level of moral maturity than the Kohlberg dilemmas, $t(59) = 5.55$, $p < .01$. The absence of any effects for sex is consistent with and extends the conclusions reached by Walker (1984) in his review and meta-analysis of relevant research: The level of moral maturity displayed by women on classic Kohlberg dilemmas, a prosocial dilemma, and a dilemma about impaired driving did not differ from the level of moral maturity displayed by men on the same dilemmas. Subjects did not, however, display the same level of moral maturity across dilemmas.

A $2 \times 2 \times 2$ ANOVA (Order \times Perspective \times Dilemma: Impaired and Prosocial) with repeated measures on the last factor failed to reveal any significant main effects or interactions for order or for perspective, $F_s(1, 54) = 1.79$ and 0.51 , respectively. Contrary to prediction, social perspective did not exert a significant effect on level of moral maturity: Subjects imagining themselves in the prosocial and impaired-driving dilemmas did not display lower levels of moral judgment than subjects responding to the dilemmas from the usual third-person perspective. Perhaps adults responding to hypothetical dilemmas

in the third person routinely imagine themselves as the protagonist, whereas children do not (see Damon, 1977; Eisenberg, 1986). Findings from other studies, however, reveal that social perspective exerts a significant effect on attributions of moral obligation and attributions of moral behavior: People tend to believe they exceed their moral obligation, but that others fall short, revealing a "self-righteous bias" in moral judgment (Denton & Krebs, 1990; Krebs et al., 1989).

Global Stage Consistency Within Subjects Across Dilemmas

As Walker, de Vries, and Trevethan (1987) have pointed out, mean differences in moral maturity across dilemmas do not necessarily disconfirm Kohlberg's claim that moral judgment is organized in structures of the whole. To evaluate Kohlberg's claim, the degree of consistency between the stage scores of individual subjects across dilemmas must be assessed.

Within-subject consistency across the two Kohlberg dilemmas. In all, 73% (44 of 60) of the subjects scored at the same major stage on the two Kohlberg dilemmas on a 5-point scale, and all but 2 (98%) scored at adjacent stages. These findings are very similar to those of Colby and Kohlberg (1987). Although there was, as would be expected, less consistency on a 9-point scale, with only 48% (29 of 60) of the subjects scoring at the same substage, 87% (52 of 60) nonetheless met our adjacent substage criterion.

Within-subject consistency across the Kohlberg and prosocial dilemmas. There was significantly less consistency between stage scores on the prosocial dilemma and stage scores on Kohlberg's dilemmas than there was between stage scores on the two dilemmas from Kohlberg's test. Only 28% (17 of 60) of the subjects scored at the same substage on the Kohlberg and prosocial dilemmas, with 63% (38 of 60) scoring within adjacent substages. All but 1 of the heterogeneous subjects obtained lower scores on the prosocial dilemma than on Kohlberg's test; 5 subjects scored a stage and a half (three substages) lower. The observed heterogeneity did not stem from stage transition: Among the 25 subjects solidly "in" a global stage on Kohlberg's test, only 7 scored at the same global stage on the prosocial dilemma.

Within-subject consistency across the Kohlberg and impaired-driving dilemmas. There was even less consistency between stage scores on Kohlberg's test and stage scores on the impaired-driving dilemma. Only 9% (5 of 58) of the subjects scored at the same substage, and only 40% (23 of 58) scored within adjacent substages. The vast majority of subjects obtained lower scores on the impaired-driving dilemma than on Kohlberg's test, with 8 subjects scoring a stage and a half or more lower. Only 1 of the 25 subjects solidly in a stage on Kohlberg's test scored at the same stage on the impaired-driving dilemma.

If valid, these findings reveal considerable stage inconsistency across dilemmas. It might, however, be argued that the inconsistency was caused by scoring error. To evaluate this possibility, a scorer trained by Kohlberg's Harvard group³ rescored the judgments of the 10 most heterogeneous subjects (i.e., those who made judgments that differed by four substages or more) blindly and independently from each other. Interrater reliabil-

ity was very high (90% exact stage matches on a 9-point scale), suggesting strongly that the observed heterogeneity was not the product of scoring error.

Stage Consistency Within Sets of Judgments

In addition to comparisons between global stage scores across dilemmas, Colby and Kohlberg (1987, p. 90) assess stage consistency in terms of the distribution of individuals' judgments within dilemmas. In these terms, their criterion for structural consistency is explicit: 90% or more of an individual's judgments must fall within adjacent stages. To obtain a measure of stage homogeneity relatively unaffected by the conventions of the Colby and Kohlberg scoring system, we compiled a stage distribution of all scored judgments made by each subject, first to the two dilemmas from Kohlberg's test, then to all four dilemmas (excluding, still, the judgments upgraded during the interview through probing). On Kohlberg's test, 88% of the subjects (53 of 60) based 90% or more of their judgments on adjacent major stages on a 5-point scale; however, only 25% (15 of 60) based 90% or more of their judgments on adjacent substages on a 9-point scale. On our interpretation, these results supply, at best, weak support for the "structure of the whole" assumption, even within Kohlberg's test. In a critique of Kohlberg's longitudinal study, Fischer (1983) reaches a stronger conclusion:

The moral judgment study, then, does *not* allow the conclusion that moral development demonstrates a strong form of the structured-whole hypothesis. With substantial variations in task and context, moral behavior seems instead to [show] . . . considerable variability in stage within an individual. (pp. 99-100)

As we would expect, inconsistency in the set of particular judgments made by each subject increased as the sample of judgments was expanded to include judgments from all four dilemmas. Only 68% (41 of 60) of the subjects met Colby and Kohlberg's (1987) 90% criterion on a 5-point scale, and only 2 subjects met this criterion on a 9-point scale. The mean percentage of judgments at each subject's modal substage was only 43%; the mean percentage at adjacent substages was 64%. As shown in Table 2, the range of stages evoked by most subjects increased across dilemmas, with 87% of the sample displaying a range of four or more substages across all four dilemmas.

Clearly, there is *some* consistency in moral judgment. On average, almost two thirds of each subject's judgments were at adjacent substages, and the correlations among moral maturity on Kohlberg's test, the prosocial dilemma, and the impaired-driving dilemma were moderately positive ($r_{s[60]} = .32, .51, \text{ and } .49$, respectively, $p_s < .01$). However, equally clearly, moral judgment is not as consistent as Colby and Kohlberg (1987) claim. The findings of this study are not consistent with the assumption that old stages are "transformed and displaced" when new stages are acquired. Rather, as argued by Rest (1983) and Levine (1979), individuals retain old stage structures and base a significant proportion of their moral judgments on them on some dilemmas.

The significance of within-subject inconsistency may get lost

³ Dennis L. Krebs.

Table 2
Percentage of Subjects Whose Judgments Fell Within Various Stage Ranges on Four Dilemmas

Dilemma	One substage		Two substages		Three substages		Four substages		Five substages		Six substages	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Kohlberg 1	23.4	14	28.3	17	33.3	20	10.7	6	5	3	0	0
Kohlberg 2	18.4	11	46.6	28	26.6	16	6.7	4	1.7	1	0	0
Kohlberg combined	3.3	2	20	12	41.7	25	25	15	10	6	0	0
Prosocial	13.3	8	25	15	46.7	28	6.7	4	6.7	4	1.7	1
Impaired driving	25.9	15	43.1	25	20.7	12	8.6	5	1.7	1	0	0
All four	0	0	1.7	1	11.7	7	36.6	22	40	24	10	6

in numbers. To appreciate the structural disparity between the judgments of some subjects, consider two examples. In response to the question on Kohlberg's test, "Should a law breaker be punished if he is acting out of conscience?" one subject gave the following Stage 5 judgment:

some of the greatest people have broken laws out of their conscience: Martin Luther [King] and Ghandi come to mind. So one has to say there are certainly times when [the] law is bad and the law has to be broken. I'm talking about not necessarily breaking the law, but changing the law.

(see CJ 34, Form A, Colby and Kohlberg, 1987, p. 138). Yet, in response to the prosocial question, "Why should you help a victim sick from drugs?" the same subject gave the following Stage 2 response: "Because he is freaked out enough to be lookin' . . . he's in real bad shape . . . [the experimenter should help] because she might find herself in this position" (see CJs 3 and 4 from Form B, *Life/Quality*, pp. 292–293). In the second case, a subject who responded to the question, "Should Officer Brown report Heinz for stealing?" with the Stage 4 response: "Yes, because that's his duty, what he's been hired for . . . The officer's duty, his obligation to his community, is to report the robbery" (CJ 29, Form A, *Punishment*, p. 174) responded to the question, "Why should people help each other?" with the classic Stage 2 response, "because it's sort of you scratch my back, I'll scratch yours" (CJ 6, Form A, *Contract*, p. 198).

Although Kohlberg's homogeneity claim pertains to general stage structures, not to specific judgments, it is particular judgments that make up global stage scores, and the "structure of the whole" claim is based on the assumption that general organizing principles structure specific judgments. For some purposes—for example, predicting the pattern of an individual's lifestyle approach to morality—it may be most appropriate to average across relatively large numbers of judgments and obtain global stage scores in much the same way it is appropriate to use general measures of attitudes when attempting to predict general patterns of behavior (see Ajzen & Fishbein, 1980). There also are purposes for which it may be most appropriate to consider only the highest level of moral judgment available to an individual—his or her moral competence. And for still other purposes, measures reflecting the proportion of preconventional or postconventional judgments, or the distribution of stage scores within individuals, may be most appropriate (cf. Eisenberg, 1986; Rest, 1983). However, for purposes such as predicting specific behaviors in specific situations, it

may well be necessary to attend to the specific judgments subjects make about the moral issues in question, in the same way it is necessary to attend to specific attitudes to predict specific behaviors (Ajzen & Fishbein, 1980).

Sources of Inconsistency in Stage Use

If most adults base their moral judgments on several stage structures, as we claim, it would seem appropriate to shift the focus in research on moral judgment from determining the stage individuals are in to determining the factors responsible for determining which of the several stage structures available to individuals they evoke when they make moral judgments about various issues and dilemmas. Our data are consistent with Colby and Kohlberg's (1987) contention that Kohlberg's test pulls for the highest stage available to individuals: Only 1 subject scored higher on the impaired-driving dilemma than on Kohlberg's dilemmas, and only 5 subjects scored higher on the prosocial dilemma (4 of whom scored only one substage higher).

If subjects performed at their level of competence on Kohlberg's test, what induced them to perform below their level of competence on the other dilemmas? Colby and Kohlberg (1987) acknowledge that performance factors such as the "low-level 'moral atmosphere' of a traditional prison" (p. 8) may constrain moral judgment, but there were no such constraints in this study: Testing conditions supported "expression of the individual's most mature thinking"; that is, the dilemmas were hypothetical, they were given in a supportive, academic atmosphere, and the moral judgments were well probed by trained interviewers. The factor that we expected would constrain moral judgment did not do so: Subjects did not make more lower level judgments from a first-person perspective than from a third-person perspective.

Colby and Kohlberg (1987, pp. 66–67) acknowledge that certain dilemmas "pull" for certain stages. Implicit in this acknowledgment is a softening of Kohlberg's constructivism. In this study, the prosocial and impaired-driving dilemmas pulled for Stage 2: 53% of the subjects scored at Stage 2 or Stage 2/3 on the prosocial dilemma, and 78% scored at these stages on the impaired-driving dilemma. The question, then, is what induced the lower stages in these dilemmas? A close examination of the types of moral judgments subjects made to these dilemmas suggests two answers: the openness of the prosocial dilemma to what Kohlberg calls "quasi-obligations"—"excuses,

generated by each stage, to justify failure to act in terms of the moral obligations generated by that stage" (Kohlberg and Candee, 1984, p. 522)—and the pull in the impaired-driving dilemma for normative defaults.

Quasi-obligations. In Kohlberg's dilemmas, responsibility is focused on the protagonists (Heinz and Dr. Jefferson), and performing a prosocial response—helping the victim in the dilemma—is generally consistent with the protagonist's role obligations (husband and doctor). However, in the prosocial dilemma, the formal role responsibilities of both the subject and the experimenter are inconsistent with helping the victim: The subject's role prescribes participation in an experiment, and the experimenter's role prescribes testing the subject. Subjects tended to use these role obligations as excuses for evading the responsibility to help, supporting individualistic Stage 2 judgments such as "he [the victim] got himself into it; he can get himself out" by invoking quasi-obligations such as "[the subject] had a responsibility to fill out the forms" and "if he [the subject] knows someone that can help, suggest it, but if not, just go about your business." Cognitive processes such as these may mediate the well-established reluctance of bystanders to intervene in emergencies when responsibility is diffused (Piliavin, Dovidio, Gaertner, & Clark, 1981). Individuals who are quite capable of engaging in relatively high-level moral judgment may "conveniently" invoke lower level judgments that supply excuses for evading responsibility in situations conducive to such excuses (i.e., in situations in which responsibility is not focused on the individual in question).

Normative defaults. Although the interviewers in the present study made every effort to probe subjects' responses to all dilemmas equally extensively, they reported meeting more resistance when probing the impaired-driving dilemma than when probing the other dilemmas: Subjects often became annoyed when probed, insisting that the judgments they made were perfectly adequate. This reaction may supply a clue to why the level of moral judgment on the impaired-driving dilemma was relatively low.

Whereas Kohlberg's dilemmas are "open" (see Blatt & Kohlberg, 1975, p. 134) in the sense of a classic moral dilemma (upholding one moral norm, such as life, entails violating another, such as law), the impaired-driving dilemma is more "closed": The values of upholding life, obeying the law, and avoiding punishment outweigh personal convenience and obligations to friends in most (but not all) subjects. Although subjects made judgments about the impaired-driving dilemma that matched criterion judgments in Colby and Kohlberg's (1987) scoring manual (thus qualifying as moral judgments on this criterion), the dilemma did not present a moral conflict in the classic sense. Thus, the judgments of most subjects to the impaired-driving dilemma seemed more mindless than their judgments to the other dilemmas, often parroting the arguments against drunk driving salient in the moral order of our culture (cf. Harré, 1983): "You could kill yourself or someone else"; "it's against the law"; "you could get a big sentence"; and "it's not worth the risk to wreck your life for one lousy night." Even though such judgments are based on low-level stage structures, they supply perfectly adequate, socially acceptable resolutions to the problem at hand. In the same way an expert mathematician will give essentially the same solution to a simple

problem in arithmetic as a 10-year-old child will, individuals may tend to evoke only the level of moral reasoning necessary to justify a particular moral choice.

One implication of this possibility is that extensive probing on Kohlberg's test may evoke upgraded moral judgments that are unrepresentative of the kinds of judgments individuals make in their everyday lives, which may be considerably more "mindless" (Langer, 1978). Although moral competence may be structurally homogeneous, moral performance may vary considerably across context and motivational state. Individuals need not be in extreme situations such as a traditional prison to make low-level moral judgments. The structure of the moral dilemmas they face and their moral choices (see deVries & Walker, 1986)—even hypothetically, in academic contexts—appear to exert considerable influence on the form of their moral judgments.

This does not, however, mean that all, or even most, of the variance in moral judgment is determined by the structure of dilemmas. Although a high proportion of the subjects appeared to succumb to the Stage 2 pull of the impaired and prosocial dilemmas, a significant number of subjects resisted it. The issue of the capacity to resist the pull of Stage 2 inducements is noteworthy. Many, if not most, of the moral dilemmas people face in their everyday lives involve a conflict between advancing their own interests (Stage 2 selfishness) and accommodating to the interests of others (Stage 3 altruism).

Resistance to Pull of Stage 2

What, then, increases the resistance of individuals to the pull or temptation of individualistic Stage 2 forms of justification? A number of personality traits come to mind, raising promising avenues for future research: Exchange versus commitment orientation (Murstein & MacDonald, 1983), self-monitoring (M. Snyder, 1987), self-awareness (Duval & Wicklund, 1972), need for consistency (Bem & Allen, 1974), coping and defensiveness (Haan, 1977), denial of responsibility (Schwartz, 1977), and cognitive rigidity (Goldstein & Blackman, 1978). In addition, as shown in Figure 1, our data suggest that level of moral competence may affect resistance to Stage 2: The higher each subject scored on Kohlberg's test (the higher his or her moral competence), the less inclined he or she was to invoke Stage 2 justifications on the prosocial and impaired-driving dilemmas, suggesting that people acquire a certain strength of conviction as they increase in moral development, as defined by Kohlberg's test.

Conclusions: An Interactional Model of Moral Judgment

We conclude that moral judgment is not organized in homogeneous structures of the whole. Because people tend to evoke different stages in response to different issues, they cannot be said to construct moral judgments in terms of their current stage. People are more morally flexible than Kohlberg's model of moral development implies; moral judgment is more plastic. Eisenberg and her colleagues reached a similar conclusion in the study of prosocial moral reasoning: "We do not find that stages are 'structured wholes' . . . a single individual often exhibits reasoning at a variety of levels" (Eisenberg, 1986, p. 182).

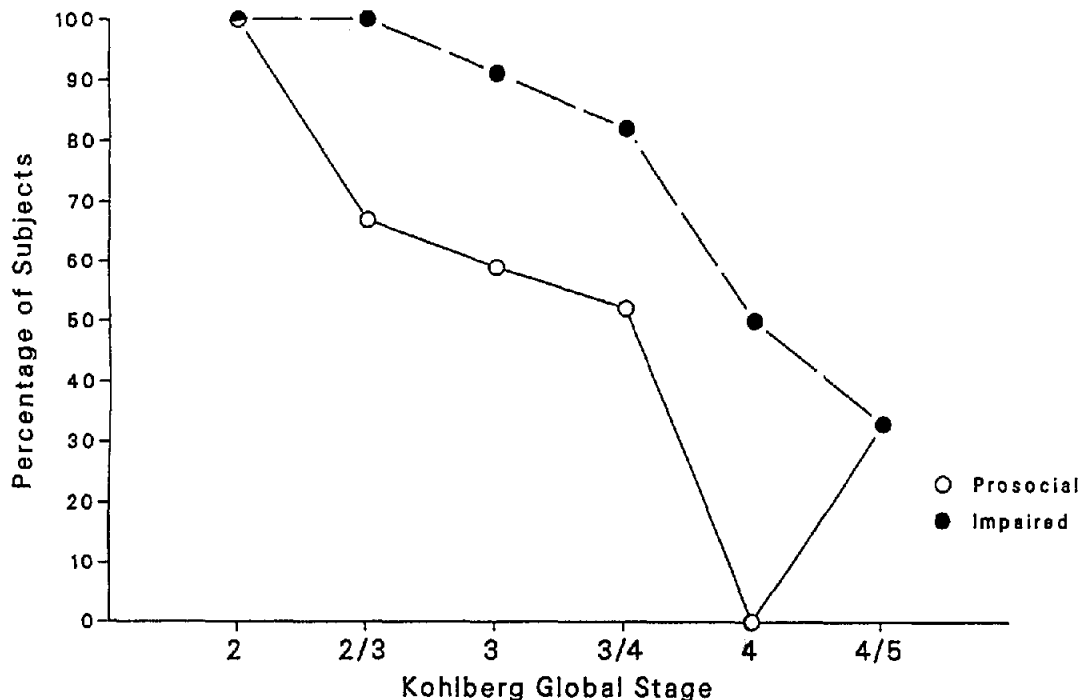


Figure 1. The percentage of subjects scoring at Stage 2 or 2/3 on the prosocial and impaired-driving dilemmas as a function of their scores on Kohlberg's test.

Our data also fail to support a strong contextual model of moral judgment: Different people invoked different stages in response to the same dilemma. The model that fits our data most comfortably is an interactional model that assumes (a) most adults base most of their moral judgments on a combination of three stage structures: Stages 2, 3, and 4 and (b) moral issues, dilemmas, and contexts differ in their pull or "resistance" (see Piaget, 1971) to these stage structures. Moral judgment results from an interaction among the interpretive structures available to people, the interpretability of the information individuals process in terms of these structures, and individuals' motivation to interpret information in particular ways.

The results of this study suggest that research on moral judgment would be more fruitfully directed toward the determination of the aspects of people and contexts that elevate and constrain the level of moral judgment individuals display in response to the types of dilemmas they face in their everyday lives than toward the determination of their level of competence on Kohlberg's test. In addition to level of moral competence, individual differences in values and personality traits relevant to various moral choices may affect moral maturity. For example, de Vries and Walker (1986) found that attitudes toward capital punishment affected the level of moral reasoning about capital punishment, and Eisenberg (1986) reports that materialistic values and need for approval affected the level of prosocial moral judgment. Contexts may vary in the strength of their pull for particular structures, with "strong situations" pulling uniformly for one interpretive structure and "weak situations" being more constructable in terms of different stages (cf. M. Snyder & Ickes, 1985). Factors that contribute

to the strength or pull of a context might include its moral order (Harré, 1983), normative structure (Backman, 1976), subjective norms and role-expectations (Ajzen & Fishbein, 1980; Schwartz, 1977), and the extent to which self-interest can be maximized through a particular course of action and type of justification (see Gerson & Damon, 1978). Self-interest may affect the attractiveness of various choices, and, as shown by de Vries and Walker (1986) and Nisan and Koriati (1989), the moral choices people make may affect the structures they invoke to justify them.

References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Backman, C. (1976). Explorations in psycho-ethics: The warranting of judgments. In R. Harré (Ed.), *Life sentences: Aspects of the social role of language* (pp. 98-108). New York: Wiley.
- Baumrind, D. (1978). A dialectical materialist's perspective on knowing social reality. In W. Damon (Ed.), *Moral development: New directions for child development* (No. 2) San Francisco: Jossey-Bass.
- Baumrind, D. (1986). Sex differences in moral reasoning: Response to Walker's (1984) conclusion that there are none. *Child Development*, 57, 511-521.
- Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistency in behavior. *Psychological Review*, 81, 506-520.
- Blatt, M., & Kohlberg, L. (1975). The effects of classroom moral discussion upon children's level of moral judgment. *Journal of Moral Education*, 4, 129-161.

- Calvert-Boyanowsky, J., & Boyanowsky, E. O. (1981). Tavern breath-testing as an alcohol countermeasure. *Transport Canada*, 1-100.
- Candee, D., & Kohlberg, L. (1987). Moral judgment and moral action: A reanalysis of Haan, Smith, and Block's (1968) Free Speech Movement data. *Journal of Personality and Social Psychology*, 52, 554-564.
- Colby, A., & Kohlberg, L. (Eds.). (1987). *The measurement of moral judgment*. (Vols. 1-2). New York: Cambridge University Press.
- Cortese, A. J. (1984). Standard issue scoring of moral reasoning: A critique. *Merrill-Palmer Quarterly*, 30, 227-247.
- Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.
- Denton, K., & Krebs, D. (1990). From the scene of the crime: The effect of alcohol and social context on moral judgment. *Journal of Personality and Social Psychology*, 59, 242-248.
- de Vries, B., & Walker, L. J. (1986). Moral reasoning and attitudes toward capital punishment. *Developmental Psychology*, 22, 509-513.
- Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. San Diego, CA: Academic Press.
- Eisenberg, N. (1982). The development of reasoning regarding prosocial behavior. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 219-249). San Diego, CA: Academic Press.
- Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Hillsdale, NJ: Erlbaum.
- Eisenberg-Berg, N., & Neal, C. (1981). The effects of person of the protagonist and costs of helping on children's moral development. *Personality and Social Psychology Bulletin*, 7, 17-23.
- Fischer, K. W. (1983). Illuminating the processes of moral development. In A. Colby, L. Kohlberg, J. Gibbs, & M. Lieberman (Eds.), *A longitudinal study of moral judgment. Monographs of the Society for Research in Child Development*, 48, (Serial No. 200).
- Flavell, J. H. (1982). On cognitive development. *Child Development*, 53, 1-10.
- Gerson, R., & Damon, W. (1978). Moral understanding and children's conduct. In W. Damon (Ed.), *Moral development: New directions for child development* (No. 2, pp. 41-59). San Francisco: Jossey-Bass.
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Gilligan, C., & Attanucci, J. (1988). Two moral orientations: Gender differences and similarities. *Merrill-Palmer Quarterly*, 34, 223-237.
- Gilligan, C., & Belenky, M. F. (1980). A naturalistic study of abortion decisions. In R. L. Selman & R. Yando (Eds.), *New Directions for Child Development: Clinical-developmental psychology* (No. 7, pp. 69-90). San Francisco: Jossey-Bass.
- Gilligan, C., Kohlberg, L., Lerner, J., & Belenky, M. (1971). *Moral reasoning about sexual dilemmas: The development of an interview and scoring system*. Unpublished technical report of the Commission on Obscenity and Pornography (Vol. 1).
- Goldstein, K. M., & Blackman, S. (1978). *Cognitive style*. New York: Wiley.
- Haan, N. (1975). Hypothetical and actual moral reasoning in a situation of civil disobedience. *Journal of Personality and Social Psychology*, 32, 255-270.
- Haan, N. (1977). *Coping and defending: Processes of self-environment organization*. San Diego, CA: Academic Press.
- Harré, R. (1983). *Personal being: A theory for individual psychology*. Cambridge, MA: Harvard University Press.
- Higgins, A., Power, C., & Kohlberg, L. (1984). The relationship of moral atmosphere to judgments of responsibility. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral development and moral behavior* (pp. 74-106). New York: Wiley.
- Higgs, A. C. (1975). An investigation of the similarities between altruistic and moral judgments. (Doctoral dissertation, University of Maryland at Baltimore, 1974). *Dissertation Abstracts International*, 35, 4269B (University Microfilms No. 75-7336).
- Kohlberg, L., & Candee, D. (1984). The relationship of moral judgment to moral action. In L. Kohlberg (Ed.), *Essays on moral development: Vol. 2. The psychology of moral development* (pp. 498-581). New York: Harper & Row.
- Kohlberg, L., & Colby, A. (1983). Reply to Fischer and Saltzstein. In A. Colby, L. Kohlberg, J. Gibbs, & M. Lieberman (Eds.), *A longitudinal study of moral judgment* (pp. 120-124). *Monographs of the Society for Research in Child Development*, 48 (Serial No. 200).
- Kohlberg, L., Scharf, P., & Hickey, J. (1972). The justice structure of the prison: A theory and intervention. *Prison Journal*, 51, 3-14.
- Krebs, D., & Denton, K. (1990, June). *Justice, care, and prosocial moral reasoning: Content vs. structure*. Paper presented at the Annual Convention of the Canadian Psychological Association.
- Krebs, D., Denton, K., Carpendale, J., Vermeulen, S., Bartek, S., & Bush, A. (1989). The many faces of moral judgment. In M. A. Luszcz & T. Nettelbeck (Eds.), *Psychological development: Perspectives across the life-span* (pp. 97-105). New York: Elsevier Science.
- Krebs, D., Denton, K., & Higgins, N. (1988). On the evolution of self-knowledge and self-deception. In K. MacDonald (Ed.), *Sociobiological perspectives on human development* (pp. 103-139). New York: Springer-Verlag.
- Krebs, D., Vermeulen, S., Carpendale, J., & Denton, K. (in press). Structural and situational influences on moral judgment: The interaction between stage and dilemma. In W. Kurtines & J. Gewirtz (Eds.), *The handbook of moral behavior and development: Theory, research, and application* (Vol. 2, 139-169). Hillsdale, NJ: Erlbaum.
- Langer, E. J. (1978). Rethinking the role of thought in social interactions. In J. H. Harvey, W. P. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 2). Hillsdale, NJ: Erlbaum.
- Leming, J. (1978). Intrapersonal variations in stage of moral reasoning among adolescents as a function of situational context. *Journal of Youth and Adolescence*, 7, 405-416.
- Levine, C. G. (1979). Stage acquisition and stage use: An appraisal of stage displacement explanations of variation in moral reasoning. *Human Development*, 22, 145-164.
- Linn, R. (1984). Practising moral judgment within the day care center: A look at the educator's moral decision under stress. *Early Child Development and Care*, 15, 117-132.
- Linn, R. (1987a). Moral disobedience during the Lebanon war: What can the cognitive-developmental approach learn from the experience of the Israeli soldiers? *Social Cognition*, 5, 383-402.
- Linn, R. (1987b). Moral reasoning and behavior of striking physicians in Israel. *Psychological Reports*, 60, 443-453.
- Lockwood, A. L. (1975). Stage of moral development and students' reasoning on public policy issues. *Journal of Moral Education*, 5, 51-61.
- McNamee, S. (1978). Moral behavior, moral development, and motivation. *Journal of Moral Education*, 7, 27-31.
- Murstein, B. I., & MacDonald, M. G. (1983). The relationship of 'exchange-orientation' and 'commitment' scales to marriage adjustment. *International Journal of Psychology*, 18, 297-311.
- Nisan, M., & Koriat, A. (1989). Moral justification of acts judged to be morally right and acts judged to be morally wrong. *British Journal of Social Psychology*, 28, 213-225.
- Nunner-Winkler, G. (1984). Two moralities: A critical discussion of an ethic of care and responsibility versus an ethic of rights and justice. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 348-361). New York: Wiley.
- Piaget, J. (1971). The theory of stages and cognitive development. In D. G. Green & M. P. Ford (Eds.), *Measurement and Piaget*. New York: McGraw-Hill.
- Piliavin, J. A., Dovidio, J. E., Gaertner, S. L., & Clark, R. D., III. (1981). *Emergency intervention*. San Diego, CA: Academic Press.

Rest, J. (1983). Morality. In J. H. Flavell, & E. Markman (Eds.), *Handbook of child psychology* (Vol. 3, 4th ed., pp. 556-629). New York: Wiley.

Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221-279). San Diego, CA: Academic Press.

Selman, R. (1980). *The growth of interpersonal understanding*. San Diego, CA: Academic Press.

Smetana, J. G. (1982). *Concepts of self and morality*. New York: Praeger.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin*, 104, 23-35.

Snyder, M. (1987). *Public appearances/private realities: The psychology of self-monitoring*. New York: Freeman.

Snyder, M., & Ickes, W. J. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed.). New York: Random House.

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

Walker, L. J. (1986). Sex differences in the development of moral reasoning: A rejoinder to Baumrind. *Child Development*, 57, 522-526.

Walker, L. J. (1988). The development of moral reasoning. *Annals of Child Development*, 5, 33-78.

Walker, L. J., de Vries, B., & Trevethan, S. D. (1987). Moral stage and moral orientations in real-life and hypothetical dilemmas. *Child Development*, 58, 842-858.

Weiss, R. J. (1982). Understanding moral thought: Effects on moral reasoning and decision making. *Developmental Psychology*, 18, 852-861.

Received May 3, 1990
 Revision received January 22, 1991
 Accepted February 14, 1991 ■



**AMERICAN PSYCHOLOGICAL ASSOCIATION
 SUBSCRIPTION CLAIMS INFORMATION**

Today's Date: _____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION _____			MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL) _____		
ADDRESS _____			DATE YOUR ORDER WAS MAILED (OR PHONED): _____		
CITY _____ STATE/COUNTRY _____ ZIP _____			P.O. NUMBER: _____		
YOUR NAME AND PHONE NUMBER _____			<input type="checkbox"/> PREPAID <input type="checkbox"/> CHECK <input type="checkbox"/> CHARGE CHECK/CARD CLEARED DATE: _____		
			(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.) ISSUES: <input type="checkbox"/> MISSING <input type="checkbox"/> DAMAGED		

TITLE _____	VOLUME OR YEAR _____	NUMBER OR MONTH _____
_____	_____	_____
_____	_____	_____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.

(TO BE FILLED OUT BY APA STAFF)			
DATE RECEIVED: _____	DATE OF ACTION: _____	_____	_____
ACTION TAKEN: _____	INV. NO. & DATE: _____	_____	_____
STAFF NAME: _____	LABEL NO. & DATE: _____	_____	_____

SEND THIS FORM TO: APA Subscription Claims, 1400 N. Uhle Street, Arlington, VA 22201-2969

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.