RELATIONS BETWEEN MORAL REASONING,
PERSONALITY TRAITS, AND
JUSTICE-DECISIONS ON HYPOTHETICAL
AND REAL-LIFE MORAL DILEMMAS

by

Russell W. C. Day
B.A. (HON), University of Victoria, 1987
M.A., University of Guelph, 1989

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the Department
of
PSYCHOLOGY

©Russell W. C. Day 1997
SIMON FRASER UNIVERSITY
July 1997

All rights reserved. This work may not be reproduced
in whole or in part, by photocopy or other means,
without the express permission of the author.
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
APPROVAL

Name: Russell W. C. Day
Degree: Ph.D.
Title: Relations Between Moral Reasoning, Personality Traits, and Justice-Decisions on Hypothetical and Real-Life Moral Dilemmas

Examining Committee:

Chair: Dr. William Turnbull

Dennis Krebs, Ph.D.
Professor
Senior Supervisor

Steve Hart, Ph.D.
Associate Professor

Gary Pool, Ph.D.
Senior Lecturer

Charles Crawford, Ph.D.
Professor
Internal/External Examiner

Kimberly Schonert-Reichl, Ph.D.
Associate Professor
Department of Educational Psychology and Special Education
Faculty of Education, University of British Columbia
External Examiner

Date Approved ____________
ABSTRACT

Participants (106 women and 95 men) responded to a distributive justice dilemma that involved allocating money to self and three others and making judgments about allocation behaviours. Participants were randomly assigned to groups that responded to the dilemma in a purely hypothetical manner, in a real situation with play money, and in a real situation with real money. Participants also completed Kohlberg's Moral Judgment Interview (MJI; Colby & Kohlberg, 1987), the Revised Interpersonal Adjectives Scale--Big 5 (IASR-B5; Trapnell & Wiggins, 1990), and a study-specific measure of the moral reasoning used to justify allocation decisions. Participants made more self-benefitting decisions when the consequences were real than when the consequences were hypothetical or involved play money and they justified their decisions at relatively low levels of moral reasoning. The data were consistent with previous findings that: a) moral reasoning scores on Kohlberg's test were higher than scores on a non-Kohlbergian dilemma; b) there were no sex differences in reasoning on Kohlberg's dilemmas; and, c) moral reasoning scores were moderately predictive of behaviour. New findings were that: a) specific instructions to engage in moral reasoning either before or after making a moral decision had little effect on moral reasoning or allocation decisions; b) men exhibited more selfish behaviour than women did, yet showed similar levels of moral reasoning; and, c) Openness scores on the IASR-B5 were as predictive of allocation behaviour as MJI scores were. The findings are discussed in terms of Kohlberg's model and the additive/inclusive model (Levine, 1979) of moral development.
Acknowledgements

The existence of this work must be attributed to the support I received from family, friends, and colleagues. Sue Smyth, my partner and mate, gave me love and support over the long months when I was often home in body but not in mind. My parents, Ken and Kay Day, my sister Keli, her husband David Swinnard, and their children, Lisa and Alex, believed from the very beginning that I could do it.

The work might never have been completed without the guidance, support, and friendship of my supervisor, Dr. Dennis Krebs. His efforts on my behalf have been exceptional. Thank you.

The other members of my committee, Drs. Gary Poole, Bill Krane, and Steve Hart challenged me to do the best work possible. Grant Burt, in addition to conducting half of the research in a very profession manner, helped me to keep the project in perspective--thanks for your good humour. Dr. Gillian Wark and Danielle Krebs spent hundreds of hours scoring all the protocols; a job well done.

Finally, my sanity, or lack thereof, through the whole process was due to the support of friends and colleagues--in no particular order: Bev, Gill, Lorie, Cristine, Joan W., Elizabeth, Phil, Dave, Wendy, Anita, Joanne, Gloria, Richard W., Kathy D., Gary, Devorah, Wayne, Dianne, and all the people who, in one way or another, made the process more human. Thanks to all.
Table of Contents

APPROVAL ....................................................................................................................... ii
ABSTRACT ...................................................................................................................... iii
ACKNOWLEDGMENTS .................................................................................................... iv
TABLE OF CONTENTS ...................................................................................................... v
LIST OF TABLES ............................................................................................................... viii
LIST OF FIGURES ............................................................................................................ ix
INTRODUCTION ................................................................................................................ 1
Problems with Kohlberg and Candee’s (1984) Model ..................................................... 6
The Current Study ........................................................................................................... 8
Issues Related to Kohlberg and Candee’s (1984) Model ............................................... 10
Exploring the Relation between Personality, Moral Reasoning, and Behaviour ......... 16
METHOD ............................................................................................................................ 18
Participants ..................................................................................................................... 18
Instruments .................................................................................................................... 19
Procedure ....................................................................................................................... 22
Scoring and Analyses.................................................................................................... 26
RESULTS .......................................................................................................................... 28
Scores on Standard and Study-specific Measures ....................................................... 28
Examination of Issues .................................................................................................. 29
APPENDICES

A  Personality Measures: Conceptual Content and Related Research ...................... 54

B  Section A: Design of the Decision Reasoning Questionnaires .......................... 56
    Section B: Decision Reasoning Questionnaire-A ........................................... 58
    Section C: Performance Scales ........................................................................ 59
    Section D: DRQ-B Questions only ................................................................. 61
    Section E: DRQ-P Questions only ................................................................. 62

C  Third-person Hypothetical “Fred” Dilemma Instructions and Questions .......... 63

D  Section A: Complete Participant Feedback--Sample: Fred Condition ............ 66
    Section B: Participant Feedback--Questionnaire Study .................................. 67
    Section C: Participant Feedback--Allocation Study ....................................... 68

E  Procedural Details ............................................................................................. 70

F  Section A: Allocation Study Instruction Summary .......................................... 72
    Section B: Ranking Task Instructions ......................................................... 73

G  Section A: Allocation Task Instructions--Reasoning After/Real Money ............ 74
    Section B: Allocation Task Instructions--Reasoning Before/Play Money .......... 75

H  Hypothetical “Imagine Self” Instructions ........................................................ 76

FOOTNOTES .......................................................................................................... 78
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Description of Groups, DRQ Format, Context, and Tasks</td>
</tr>
<tr>
<td>2. Mean Moral Maturity across Sex and Kohlberg Dilemmas</td>
</tr>
<tr>
<td>3. Mean IASR-B5 Domain Scores by Sex compared to Trapnell and Wiggins (1992)</td>
</tr>
<tr>
<td>4. Mean NEO-PI-R Domain Scores compared to Costa and McCrae (1992) Table B-3 for College-Age Individuals</td>
</tr>
<tr>
<td>5. Pearson’s Correlations between NEO-PI-R Domain scores and IASR-B5 Domain Scores (N = 166)</td>
</tr>
<tr>
<td>6. Mean KMM and DJMM Scores across Conditions</td>
</tr>
<tr>
<td>7. Mean Money Kept Scores and DJMM Scores by Consequence</td>
</tr>
<tr>
<td>8. Mean KMM and DJMM Scores by Sex and Condition</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
</table>
For decades, psychologists have attempted, without much success, to explain the relation between moral judgment and moral behaviour. Many researchers, from Hartshorne and May (1929) to the recent attempts reviewed by Kohlberg and Candee (1984), have concluded that, in general, there is only a weak, if any, relation between moral judgment and moral behaviour. This has led some researchers to suggest moral reasoning and moral behaviour are on separate tracks; that "what people say and what people do" are basically unrelated (see Krebs & Denton, in press, for a review).

Kohlberg and Candee (1984) have argued that the failure to find a strong relation between moral judgment and moral behaviour results from using invalid measures of both moral judgment and moral behaviour. Over a period of three decades, Kohlberg (Colby & Kohlberg, 1984, 1987) developed a measure, that he argued was a valid and reliable measure of moral judgment. In addition, Kohlberg proposed a model of the relation between moral judgment and moral behaviour.

In Kohlberg’s model (Colby & Kohlberg, 1987) moral judgment is assumed to develop progressively in fixed stages, following an invariant sequence, with each more advanced stage structure completely replacing the previous stage structure; that is, each stage is a "structure of the whole," and the more advanced stage transforms and displaces the earlier developmental stage (Kohlberg, 1984). Initially, Kohlberg outlined six stages
necessary to define all possible levels of moral judgment, but after much criticism, Kohlberg modified the six stage approach, suggesting only five stages were necessary to measure moral maturity.

Kohlberg's measure of moral judgment. The most recent version of Kohlberg's Moral Judgment Interview (MJI; Colby & Kohlberg, 1987) poses hypothetical moral dilemmas based on competing moral norms (e.g., life vs. law). An individual responds to the dilemmas, either orally or in writing, by making a decision about what the protagonist should do, then justifying that decision. A familiar example is the dilemma faced by Heinz, who must decide whether or not to steal a drug to save his dying wife. People responding to the MJI dilemma must decide what Heinz should do, then justify the decision. The reasons used to justify the decision, not the decision itself, are scored for stage of moral maturity by matching the reasons or "judgments" with "criterion judgments" in Colby and Kohlberg's (1987) 1500 page scoring manual. Scores from all matched judgments are weighted and averaged to produce a global stage score from 1 to 5 and/or a weighted average score (WAS) from 100 to 500 (corresponding to stage 1 and stage 5). WAS are also called moral maturity scores. The WAS corresponds to the highest stage achieved, or moral competence.


Kohlberg and Candee (1984) outlined a four-step model of the relation between moral judgment and moral behaviour (see Figure 1). Kohlberg and Candee made several assumptions, namely: 1) that the MJI is a good measure of moral judgment, 2) that people
facing moral situations will desire to know the morally correct behaviour in that situation, and, 3) that people will use rational, deductive problem-solving processes to figure out the most moral course of action.

Step 1 in the process involves interpreting the information in the situation according to the stage of moral development achieved. “Stage of moral reasoning is a filter through which...situational forces are perceived, interpreted, and acted upon” (Kohlberg & Candee, 1984, p. 564). People at different stages perceive the situation differently, focussing on the aspects relevant to their stage. For example, a person at Stage 2 may focus on individual, instrumental purpose and exchange, potentially leading to behavioural preferences that maximize self-interest in terms of concrete fair exchange. A person at Stage 3 may focus on mutual, interpersonal expectations, relationships, and conformity or the Golden Rule, potentially leading to behavioural preferences that are attentive to the expectations of others.

Stage-based moral reasoning processes lead to a moral decision, which Kohlberg and Candee (1984) called a deontic choice (Step 2). Once the deontic, or “should” or “right” choice is made, the appropriate behaviour, being thus “prescribed,” should follow. Theoretically, depending on the dilemma, the specific deontic choice a person makes may or may not relate systematically to his or her stage of moral development. For example, people at all stages might be willing to tell a small lie to save a life; whereas, on more difficult dilemmas, only people who have reached high stages may make a “more moral” decision. Although deontic choice is necessary for moral behaviour, it does not lead directly to moral behaviour, making moral judgment even less related to moral behaviour.
At Step 3 of Kohlberg and Candee's (1984) model, people make judgments of responsibility or obligation; that is, whether or not, regardless of their stage of moral judgment or deontic choice, they feel compelled to "follow-through" and behave according to their decision about the morally "right" action.

At Step 4 of Kohlberg and Candee's (1984) model non-moral "follow through" processes labelled "ego control" (intelligence, attention, and delay of gratification) link judgments about the obligation to act on judgments of responsibility, and moral behaviour. Most researchers examining the relation between moral judgment and behaviour have ignored steps 3 and 4.

What is moral behaviour? The outcome variable in Kohlberg and Candee's (1984) model is moral action, which is defined very specifically. Kohlberg and Candee claim that "moral judgment is a necessary component of an action judged moral, but it need not be sufficient for evaluating the morality of an action or actor" (p. 512). Locke (1983) criticized Kohlberg's definition of moral action, arguing that Kohlberg's definition is similar to Kant's idea that "moral action consists in acting for the sake of morality itself, or as he would put it, acting out of respect for the moral law" (Locke, 1983, p. 114).

This narrow definition, that all moral action stems from moral reasoning, excludes actions that are moral, but are not derived from moral reasoning; which means that few, if any, everyday behaviours would be considered moral. An alternative model is based on the idea that people have any number of habits and stable behavioural patterns (perhaps traits) that lead to specific behaviours--many of which could be considered moral. The specific
behaviour elicited in a situation depends on the type and strength of the evoking stimuli, as well as on a complex interplay of internal trait strengths, inhibiting factors, and situational understanding. A more interactional model of moral judgment/action, which might include cognitive processes other than the rational processes of moral reasoning, is needed.

A resolution of the restricted-range of moral behaviour problem of Kohlberg and Candee’s (1984) model can be achieved by using a broader definition of what constitutes moral action. This broader definition recognizes that some behaviours are motivated by moral reasoning, and other behaviour may be classified as moral in an a posteriori fashion. The resulting definition, “moral action as action which is explicable, at least in part, by moral reasons” (Locke, 1983, p. 117) includes actions that might not be considered moral under the strict Kantian/Kohlbergian definition. However, this broader definition of morality has greater utility, especially if we are to accept the notion that people sometimes act without thinking. In summary, an action is not considered moral until we pass judgment on it before or after the action takes place.

Research on Kohlberg and Candee’s Model. As mentioned earlier, and as documented by Blasi (1980) and Kohlberg and Candee (1984), most past research on the relation between moral judgment and behaviour deemed more or less moral has examined the relation between scores on Kohlberg’s test (Step 1 in Kohlberg and Candee’s model), at one point in time, and some behaviour considered to be moral or immoral at another point in time. Blasi’s (1980) review of research found that most studies reported a weak positive relation between scores on Kohlberg’s test and measures of moral behaviour. Even Kohlberg
and Candee noted that attempts to predict moral behaviour from moral judgment scores have had only moderate success.

**Problems with Kohlberg and Candee’s Model**

The support reported by Kohlberg and Candee (1984) notwithstanding, more and more evidence has pointed to problems with their model. The first problem is that Kohlberg and colleagues assume people use the same level of moral maturity (stage structure) they display on Kohlberg’s test as they do on other dilemmas—Kohlberg’s structure of the whole assumption. Kohlberg supports this assumption by showing that moral reasoning is very consistent across the dilemmas on his test. However, the nine dilemmas on the MJI were developed specifically to produce consistency by assessing the highest level of competence available to an individual.

Krebs and his colleagues (Krebs et al., 1989; Krebs, Denton, Vermeulen, Carpendale, & Bush, 1991; Krebs, Vermeulen, & Denton, 1991; Krebs & Denton, in press) have found that people generally score lower on real-life dilemmas than they do on Kohlberg's test. Krebs et al. have concluded that people do not always use the same cognitive structures or perform at their level of moral competence when making real-life moral decisions (Krebs, Denton et al., 1991; Krebs, Vermeulen et al., 1991; Carpendale & Krebs, 1992, 1995). Indeed, Krebs, Denton et al. (1991) have argued that all the stages people have acquired are available to them and may be invoked when so demanded by the situation. While these findings do not challenge Kohlberg’s claim that his measure assesses the highest level of competence achieved, they do challenge his structure of the whole assumption. Krebs,
Individual Differences and Justice-Decisions

Denton et al. (1991) conclude that their findings are most consistent with an interactive, additive/inclusive model such as the one proposed by Levine (1979) in which moral development "can be understood as describing a process of stage acquisition in which higher stages include components of earlier stages but do not replace these stages" (p. 155).

The second problem with Kohlberg and Candee’s (1984) model is that they assume that moral reasoning produces deontic choices, that is, moral reasoning precedes deontic choice. However, on Kohlberg’s test (MJI), people make the deontic choice first (Should Heinz steal the drug?), then offer arguments in support of the choice—moral reasoning follows moral choice. Evidence from social psychology suggests that the content of self-justification may be affected by the preceding decision (e.g., cognitive dissonance; hindsight effects; Fischhoff, 1975). In addition, although Kohlberg assumes little or no systematic relation between stage and deontic choice at lower stages, evidence shows that some deontic choices (e.g., steal the drug) led to higher moral maturity scores than other choices (e.g., do not steal the drug) (Nisan & Koriat, 1989; Carpendale & Krebs, 1992).

A third problem is that Kohlberg and Candee’s (1984) third and fourth functions—judgments of responsibility or obligation and ego controls—are often neglected when examining the relation between moral judgment and behaviour. Although Kohlberg and Candee have argued that ego controls and judgments of responsibility mediate between moral judgment and moral action, most other researchers have not assessed either judgment of responsibility or ego controls in relation to moral judgment or in relation to behaviour. The failure to include some measures of the constructs might account for the weak-to-moderate
correlations between judgment and behaviour (Blasi, 1980; Kohlberg & Candee, 1984). The first step toward including judgments of responsibility and ego controls in Kohlberg's model might be to include standard, global measures of personality to confirm the relation between personality traits, moral reasoning, and behaviour.

The Current Study

The overriding goals of this study were to evaluate central aspects of Kohlberg's model, to explore possible revisions, and to examine one alternative--the interaction additive/inclusive model. The major focus was on the moral judgments made about a non-Kohlbergian dilemma and the relation between judgment and behaviour. A minor goal was to examine the possibility that personality measures could be used as a way of assessing judgments of responsibility and non-moral ego controls postulated by Kohlberg and Candee (1984), and consequently, to compare the ability of moral judgment alone to predict behaviour with the ability of moral judgment plus judgments of responsibility and ego control to predict behaviour.

The first step was to design a context in which participants who differed in personality could exhibit some level of moral behaviour that could be measured in some discrete way, and about which participants could generate moral arguments, which, in turn, could be scored for stage of moral reasoning. The goals of this research required the design to include measures of: 1) individual differences in personality, 2) competence on moral reasoning, 3) moral behaviour, and 4) reasoning about the moral behaviour. Standardized measures of personality and moral reasoning were available and a procedure for assessing
moral reasoning about real-life dilemmas offered some guidance for the development of a study-specific measure of moral reasoning. The second step was to create an appropriate dilemma.

**Distributive justice dilemma.** Colby and Kohlberg (1987) outlined three independent types of justice, distributive justice, commutative justice, and corrective justice, and a fourth dependent type, procedural justice. Distributive justice, "the way in which society or a third party distributes 'honor, wealth, and other desirable assets of the community'" based on "equality, desert, merit...and, finally, equity" is central to all stages of Kohlberg's model (Colby & Kohlberg, 1987, p. 24). Distributive justice dilemmas range in complexity from the fair division of candy between children to concerns about human rights and values. Previous research has used distributive justice dilemmas to study the relation between moral judgment and behaviour (Carpendale & Krebs, 1992, 1995; Damon, 1977).

I devised a distributive justice task that required participants to distribute a finite resource (money) to other members of their "working group." The money could not be distributed equally (10 X $2.00 bills/4 participants). The amount of money kept for self could be measured accurately and deemed as more or less self-benefitting. Participants were asked to generate arguments about the decisions they made--arguments that could be scored for stage of moral reasoning. The dependent variables were: 1) the amount of money kept for self--"money kept" (resource allocation behaviour), and 2) the level of moral reasoning used to justify the amount of money kept (distributive justice moral maturity--DJMM).
**Issues Related to Kohlberg's (1984) Model**

**Issue 1.** The first issue concerned whether people invoke the same level of moral reasoning on a hypothetical, resource allocation dilemma when responding to it from a first-person versus third-person perspective. Kohlberg's test (MJI) asks people to respond in the third-person, but real-life dilemmas generally invoke first-person reasoning. With content held constant, would the degree to which a person self-projects into the dilemma affect his or her moral judgment?

There has been only one study on the first-person/third-person perspective-taken issue using Kohlberg's test. Krebs, Vermeulen, Denton, and Carpendale (1994) failed to find any differences between people's level of moral judgment when they responded in the third-person versus in the first-person to four of Kohlberg's hypothetical dilemmas. In their review of the literature, Krebs et al. (1994) had found three studies of the effect of third-person versus first-person perspectives on Rest's Defining Issues Test (DIT); however, they concluded that the combined evidence from the studies was inconclusive; for example, perspective differences were evident for high school and college age participants, but were not evident for older participants, and, in one study, there were complex sex differences. I wondered whether or not the Krebs et al. findings would generalize to the resource allocation dilemma. Related to this issue was whether people would allocate more money to themselves in the hypothetical first-person dilemma than they would in the hypothetical third-person dilemma.

To test the comparison of interest, two groups completed the tasks under different
instructions. One group responded to a hypothetical, third-person “Fred” dilemma similar to the dilemmas on Kohlberg’s test, but based on the distributive justice task used in this study. A second group, operating under “imagine self” instructions, responded to the same hypothetical distributive justice dilemma as the “Fred” group, except they were asked to imagine that they were faced with the dilemma.

**Issue 2.** The second issue examined whether people invoke different structures of moral judgment before they make a moral decision than they do after making a decision, and whether this relation affects their decisions and their behaviour. As outlined earlier, Kohlberg and Candee (1984) assume that people engage in a rapid, unconscious, stage-consistent form of moral reasoning that gives rise to a deontic choice, which, in turn, leads to behaviour. However, some social psychological evidence suggests that people make (moral) decisions and act on those decisions without engaging in moral reasoning (e.g., emergency helping behaviour; Piliavin, Dovidio, Gaertner, & Clark, 1981).

In this study, I tested whether there were differences between level of moral reasoning indicated before versus after resource allocation decisions. It seemed plausible that evoking moral reasoning before the allocation decision would result in a higher level of moral reasoning and a less self-benefitting decision than evoking reasoning after the decision. Although the before/after question has been raised as being theoretically relevant (Saltzstein, 1994), to my knowledge, there is no past research on this issue.

To test the comparison of interest, two groups completed the tasks under different instructions. One group was asked to engage in and record their moral reasoning before
making the resource allocation decision, while a second group was asked to make their resource allocation decision first, then record their moral reasoning. From Saltzstein’s (1994) perspective, both the amount of money kept and DJMM could be influenced by time of reasoning, with participants engaged in moral reasoning before making the allocation decision keeping less money and invoking higher levels of moral reasoning to justify their decisions than participants who reasoned after the decision. From Kohlberg’s perspective, there should not be any time of reasoning differences.

**Issue 3.** The third issue was whether Kohlberg’s “structure of the whole” assumption, that is, that all moral judgments will be at the same or adjacent stages across dilemmas, would extend to the distributive justice dilemma used in this study. Consistent with the additive/inclusive model (Levine, 1979; Krebs, Denton et al., 1991) and the Krebs, Vermeulen et al. (1991) finding that moral reasoning scores on Kohlberg’s dilemmas are higher than scores on other moral dilemmas, the prediction was that Kohlberg’s Moral Maturity (KMM) scores would be higher than the DJMM scores across all groups.

**Issue 4.** The fourth issue, related to the first and third issues, was whether the degree of similarity between a non-Kohlbergian dilemma and Kohlberg’s dilemmas would affect the similarity between the respective moral reasoning scores. Logically, the more similar a dilemma is to Kohlberg’s dilemmas in terms of administration format, content, and structure (hypothetical to real-life), the more similar should be the moral reasoning scores. For example, Bush, Krebs, and Carpendale (1993) found that changing the content of one of Kohlberg’s dilemmas (i.e., replacing the victim—Heinz’s wife who is dying of cancer—with a
homosexual dying of AIDS) did not result in lower stage reasoning; participants invoked the same moral reasoning for both the AIDS dilemma and Kohlberg’s dilemma.

In this study, participants in the “Fred” and “imagine self” groups responded to a purely hypothetical resource allocation dilemma that was most similar in administration format to Kohlberg’s dilemmas. Another group responded to the resource allocation dilemma in a real group setting, but the consequences were limited to the distribution of play money. Participants in the remaining group were asked to distribute $20.00 real money, any or all of which they could keep. Because the consequences were real, the dilemma faced by this group was the least similar to those on Kohlberg’s test. The prediction was there would be a positive relation between KMM and DJMM scores, with the relation being the strongest in the hypothetical, then the play, then the real conditions.

**Issue 5.** The fifth issue concerned whether or not the consequences of a decision about how to allocate resources would have any impact on the decision and the moral reasoning used to justify it. A purely hypothetical or play-money allocation decision has no consequences compared to an allocation decision involving real money, which may have sufficient potential self-benefit to affect behaviour and reasoning about that behaviour.

Little research has examined the relation between level of consequence and the level of moral reasoning used to justify the moral decisions. Carpendale and Krebs (1995) created a real-life distributive justice dilemma, which asked how many of a product’s defects a seller should reveal to a buyer, that asked participants to allocate actual money to self, as the “seller” of defective goods, and to another individual, the “buyer.” Half the participants had
real money to allocate; the other half of the participants responded to the same dilemma in
the traditional hypothetical format. Contrary to their prediction, Carpendale and Krebs found
that people distributing real money actually kept less money and reasoned at a higher moral
level than people who responded to the hypothetical dilemma. There are several aspects of
the design of the Carpendale and Krebs study that may have produced the counter-intuitive
findings. First, the consequences for the “buyer” were salient in the study. Indeed
Carpendale and Krebs suggested that participants, who were “sellers,” identified with the
“buyers.” Second, the comparison between the hypothetical and consequential conditions
confounded administration format with degree of consequence. To test the effect of level of
consequence requires three levels of consequence: hypothetical--knowing there are no other
people involved; non-consequential--knowing there are other people involved but without
consequences; and, consequential--knowing there are other people involved, with
consequences.

If identifying with the recipient accounted for the Carpendale and Krebs (1995)
findings, then participants in this study responding to a real dilemma would be expected to
behave most generously because the (“real”) recipient would be more salient than the
hypothetical recipients. Alternatively, if people are motivated to advance their own interests
(economic model), then participants making hypothetical decisions should keep the least
money and people making real consequence decisions should keep the most money.

In summary, as originally predicted by Carpendale and Krebs (1995), I expected that
money kept and DJMM would be influenced by the reality of the consequences (hypothetical,
play, or real money). I expected participants who made the decision about how to distribute real money to keep more money and invoke lower levels of moral reasoning to justify the decision than participants making the decision about distributing play money. I expected participants who made purely hypothetical decisions (Fred and Imagine Self) to prescribe keeping the least money and invoke the highest levels of moral reasoning.

**Issue 6.** The sixth issue related to the link between behaviour and judgment; is self-benefitting behaviour related to low level moral reasoning? As outlined above, self-benefitting behaviour is most easily justified at Stage 2 in Kohlberg’s system. Carpendale and Krebs (1995) reported that people who kept more money for themselves (selfish behaviour) invoked lower levels of moral reasoning to justify their behaviour than people who kept less money (generous behaviour) and visa versa. I expected to find a similar relation even though it would be inconsistent with Colby and Kohlberg’s (1987) model (the structure of the whole assumption). I expected that the money kept scores would be inversely related to DJMM and KMM scores and further, there would be a stronger relation between the DJMM, which involved reasoning about the amount of money kept, and the amount of money kept than between KMM, which did not pertain to the behavioural decision, and the amount of money kept.

**Issue 7.** The seventh issue, although not a primary focus of this study, pertained to whether or not there would be sex differences across any of the variables of interest in the study. Kohlberg’s model and measure have been criticized for penalizing women (Gilligan, 1982), yet a number of researchers have suggested that no bias has been conclusively
demonstrated (Wark & Krebs, 1996; Walker, 1988). Carpendale and Krebs (1995) used an exclusively male sample to avoid sex difference complications. Major and Deaux (1982) reported complex sex interaction effects on allocation tasks, with women generally keeping less of a reward than men did. Therefore, participants completed this study in same-sex groups, and I predicted that women would keep less money than men kept and use higher levels of moral reasoning to justify their decisions.

Exploring the Relation between Personality, Moral Reasoning, and Behaviour

Issues 8 and 9. The eighth and ninth issues were whether personality traits would be related to moral maturity and moral behaviour. As mentioned earlier, Kohlberg and Candee (1984) have assessed the non-moral “follow through” aspects of their model; however, most other researchers have not examined them in any detail. It might be argued that these individual differences could be assessed using a global measure of personality. In this study, I used the most general, broadly based, easily administered measures of personality available to explore the relation between the Big 5 personality traits, moral judgment, and behaviour (the content of the measures used and related research are outlined in Appendix A).

To review, at step 3 of Kohlberg and Candee’s (1984) model people make judgments of responsibility or judgments of commitment to follow through on deontic choice. One trait domain, “conscientiousness,” seemed conceptually similar to Kohlberg and Candee’s (1984) step 3 functions. McCrae and Costa (1990) describe conscientious people as “adhering scrupulously to their moral precepts and rigorously fulfilling their social and civic duties” (p. 45), and, noted that “(L)ow scorers are not necessarily lacking in moral principles, but they
are less exacting in applying them... (Costa & McCrae, 1992, p. 16). I expected there would be a relation between conscientiousness and, a) moral maturity, and b) the amount of money allocated to others.

Step 4 in Kohlberg and Candee’s (1984) model are the non-moral “ego controls.”

One trait domain, labelled “openness” by Trapnell and Wiggins (1990) and “openness to experience” by Costa and McCrae (1992) seemed conceptually similar to ego controls. Openness to experience, as an individual difference domain, has been labelled a number of different ways including intellect, culture, and intelligence (Trapnell & Wiggins, 1990) and is positively correlated with measures of divergent thinking ability and sensation seeking (McCrae, 1987), and Personality Research Form needs for change, sentience, and understanding scales (McCrae, 1990). Johnson and Ostendorf (1993) have argued that the openness to experience domain might be better labelled as intellect, while McCrae (1993-94) reported only weak to moderate correlations with WAIS-R intelligence scores.

The Openness items of Trapnell and Wiggins (1990) measure seem to reflect the broad nature of the construct; for example, the positively loaded terms are: philosophical, abstract-thinking, imaginative, inquisitive, reflective, literary, questioning, individualistic, unconventional, broad-minded, and the negatively loaded terms are: conventional, unartistic, unliterary, unreflective, uncomplex, unimaginative, unabstract, unsearching, uninquisitive, and unphilosophical. Higher levels of moral reasoning require cognitive skills that might be captured by terms like philosophical, abstract-thinking, inquisitive, and reflective. I expected there would be a positive relation between openness and, a) moral maturity, and b) the
amount of money allocated to others.

**Summary of Study Elements**

All participants--university undergraduate students--completed one or two standardized measures of personality and Kohlberg's measure of moral maturity. Each participant was asked to make a decision about how $20.00 "bonus money" should be divided among the four members of a group, with the self included as one of the four. The dilemma participants faced resulted from the fact that the "bonus money," in $2.00 bills, could not be evenly divided among four group members.

To test all the comparisons of interest, six groups completed the tasks under different instructions as outlined above (Issues 1 - 4). All participants completed: 1) Kohlberg's MJI, 2) the Interpersonal Adjective Scale Revised--Big 5 (IASR-B5; Trapnell & Wiggins, 1990), and, with the exception of one group of participants (third-person hypothetical), 3) the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992). Additionally, I designed an instrument, called the Decision Reasoning Questionnaire (DRQ), to assess moral reasoning about the distributive justice behaviour. The various tasks and measures completed by each of the six groups are outlined in Table 1.

**Method**

**Participants**

The sample was composed of 229 (116 women, 113 men) undergraduate student volunteers from Simon Fraser University. Twenty-six participants (10 women, 16 men) completed only one of two components of the study, two male participants failed to follow
instructions, leaving 201 participants (106 women, 95 men). An examination of the demographic variables for all groups showed no consistent differences between participants who completed the research and participants who did not, with the distribution of incomplete participation across the six conditions seeming random. The participants in this sample were typical of the undergraduate population, with a mean age of 20.5 years (N = 229, range 16 - 49; women M = 20.5, men M = 20.6), and, of those participants reporting a GPA, women reporting a significantly higher GPA than men did (women M = 3.17, SD = .47; men M = 2.89, SD = .49), t(168) = 3.82, p < .001. Most of the participants were in their first semester and many had not declared a major (28%), with the remainder distributed across Psychology (20%), Business and Economics (17%), Arts (13%), Applied Sciences (13%), Basic Sciences (12%), and Education (1%).

In all groups, except the “Fred” hypothetical group, 166 participants (86 women, 80 men) were asked to complete two studies that were actually two phases of a single study. In one phase, participants were given course credit in exchange for their participation. In the other phase, participants were paid $6.00 for their participation, a necessary part of the distributive justice manipulation. The order of phase completion was counterbalanced and no order differences were found. In the Fred-Hypo group, 35 participants (20 women, 15 men) participated in exchange for course credit in a single session.

**Instruments**

Kohlberg's Moral Judgment Interview (Colby & Kohlberg, 1987) was used to assess each participant’s stage of moral development. Both the Revised NEO Personality Inventory
Individual Differences and Justice-Decisions

(NEO-PI-R; Costa & McCrae, 1992) and the Interpersonal Adjective Scale Revised--Big 5 (IASR-B5; Trapnell & Wiggins, 1990) were used to assess the non-clinical individual differences on five dimensions of personality. The Decision Reasoning Questionnaire (DRQ) was designed specifically for this research and is outlined in more detail below.

**Moral Judgment Interview.** Form “A” of Kohlberg’s test was administered following procedures outlined by Colby and Kohlberg (1987). Form “A” includes three hypothetical dilemmas--the Heinz dilemma (III), the Officer Brown/Judge dilemma (III') and the Joe dilemma (I)--each of which is followed by a series of probe questions about the dilemma. Responses were scored for stage according to the instructions outlined in Colby and Kohlberg’s (1987) scoring manual. Using all three dilemmas of Form “A” ensured the stability and reliability of the overall Moral Maturity Score.

**Interpersonal Adjective Scale Revised--Big 5.** All participants completed the IASR-B5 (Trapnell & Wiggins, 1990)--a standard, non-clinical measure of five domains of personality. This 124 item measure is used to assess personality and has been validated across many samples.

**Revised NEO Personality Inventory.** In all except the third-person hypothetical group (Hypo-Fred, see Table 1), participants also completed the NEO-PI-R, a 240 item measure used to assess normal personality that has been thoroughly validated across many samples (Costa & McCrae, 1992).

The IASR-B5 and the NEO-PI-R assess similar aspects of personality, with three of the five domains having common summary labels (Neuroticism, Conscientiousness, and
Openness). Trapnell and Wiggins (1990) report similar factor structures for both instruments, strong correlations between their IASR-B5 16 item Dominance scale and the NEO-PI-R Extraversion scale, and between their 16 item Nurturance scale and the NEO-PI-R Agreeableness scale. The IASR-B5 has the advantage of taking only 15 minutes compared to 60 minutes to complete the NEO-PI-R.

Decision Reasoning Questionnaire(s). I designed and pilot-tested a study-specific set of instruments for the current research. The Decision Reasoning Questionnaires (DRQs) asked participants to give the reasons for the justice decisions made about the allocation task. Responses to specific questions were successfully and reliably scored for moral maturity. Moral Maturity was calculated following the same procedure reported by Krebs, Denton et al. (1991) for the scoring of real-life moral dilemmas.

The DRQ asked 13 questions designed to elicit reasons for decisions made, or about to be made, and about the influence of moral principles on the decision made. In the reasoning before (T-Bef) conditions, 10 of the items were part of the DRQ-B (B = before), which was used to evoke moral reasoning before the money allocation task was completed. After completing the distributive justice (money allocation) task, the final three questions, as part of the DRQ-P (P = post-decision), asked the participants to judge the fairness or unfairness of their actual allocation decisions. The DRQ-A (A = after), which included the 13 questions mentioned above in past-tense form, was completed after the distributive justice task decision was made. For the Fred-Hypo group, a third version of the instrument (DRQ-Fred) contained the same questions as those asked on the DRQ-A, but each question was
rephrased to reflect the third-person hypothetical nature of the decision reasoning being invoked. Details about the design, content, and the questions are included in Appendix B.

**Demographic questionnaire.** To maintain the fiction of two separate studies, all participants except those in the Fred condition completed two short demographic questionnaires, one at the beginning of each phase in which they participated.

**Procedure**

**Third-person hypothetical—Fred condition.** Participants in the hypothetical Fred condition were given a take-home questionnaire package to complete in exchange for psychology course credit; 35 of 36 were completed and returned. Participants were instructed to complete the demographic questionnaire, then respond to a purely hypothetical dilemma faced by a character, Fred, who participated in a study that included a distributive justice task. After reading the scenario in which: 1) Fred had been randomly assigned the role of “coordinator” for a nominal group of three “writers” whose task it was to write arguments against the legalization of marijuana; 2) Fred had completed a personality questionnaire; 3) Fred had read and ranked the arguments produced by the three writers; then participants were asked to suggest how Fred should distribute some bonus money. After recording how much money Fred should give to each of the three writers and how much he should keep for himself, participants completed the DRQ-Fred by responding to the questions about what Fred should have done and why (see Appendix C). Participants then completed the IASR-B5, followed by Kohlberg's test in a fixed order to control any priming effects the DRQ-Fred might have on the MJI. Participants in the Fred condition received an extensive
written debriefing form when they returned their completed packages (see Appendix D).

**Experimental and "Imagine Self" conditions.** Participants in the remaining five conditions participated in two seemingly separate studies, with the only connection between the two studies being the "long and expensive" personality measure (NEO-PI-R), scores that would be "used by two different researchers."

**Questionnaire/manipulation order.** After consenting to participate in the research and completing research participation credit documentation, participants completed two paper-and-pencil measures, in counterbalanced order: Form A of Kohlberg's Moral Judgment Interview (MJI) and the NEO-PI-R (Costa & McCrae, 1992). Next, participants were partially debriefed (see Appendix D, Sections B & C) and reminded of their commitment to participate in the second study the following week—same time, different room, and with a different researcher.

When participants (all in same sex groups) arrived to participate in the second study, they were informed they had been randomly assigned to the nominal, four person condition, one of several conditions in the study (details of the bogus conditions and the explanations used to describe nominal groups are included in Appendix E). The researcher told the group "your task as a group is to write arguments opposing the legalization of marijuana," and that one member of the group would be randomly assigned the role of "coordinator." The coordinator’s task was described as assuming responsibility for the final product, which meant "picking the best argument from each set of arguments written by the other group members" (in fact, all participants were assigned the coordinator role and received pre-
written arguments).

Participants were asked to give consent to a random drawing for task assignment and to the idea that they might be evaluated by—or required to evaluate—the other participants. Participants were paid $6.00 each, in $2.00 bills, then asked to provide their home addresses on stamped envelopes in which the researcher would mail, in due course, the full debriefing information (see Appendix D, Section A). Each participant then picked, at random, a research package and went to a separate, sound-dampened research cubicle before opening it. Participants who had picked the “imagine self” condition package were instructed to move to a separate research space to complete the questionnaires at their own pace.

When participants opened their package, they all found they had been assigned the coordinator role and were given additional information about the coordinator’s responsibilities (see Appendix F, Section A). Coordinators were instructed to complete the demographic questionnaire and the IASR-B5 personality questionnaire while the other members of their group supposedly wrote their arguments against the legalization of marijuana.

Each participant completed the remaining tasks as outlined below—with the minor alterations for the T-Bef conditions (time of reasoning—Before) being noted. After completing the personality questionnaire, participants were instructed to pick the best argument from each set produced by each of the other three members of their nominal group within a 15 minute time limit (see Appendix F, Section B). At this point, approximately 20 minutes after participants had moved into the research cubicles, the researcher delivered a set
of freshly hand-written arguments for them to rank order.

After completing the ranking task and a final question about whether or not marijuana should be legalized in Canada (should be = 1 to should NOT be = 7, \(N = 166, M = 5.00, SD = 1.90\), unrelated to any other variable), participants sealed the handwritten arguments and the ranking sheet in a large envelope, then read a new set of instructions that had been delivered with the written arguments. The instructions informed the participants that they, as coordinators, were obliged to determine how to distribute the $20.00 bonus money the researcher had given to the group, but about which the other group members knew nothing. The coordinators had to determine how much each other group member deserved (and received), and how much they would keep for themselves. Participants were assured that all decisions they made would be completely confidential. The instructions made it clear that the 10 X $2.00 bills could be distributed in any fashion and noted the impossibility of distributing the $2.00 bills equally (see Appendix G, Section A for two examples).

Participants in the time of reasoning—Before (T-Bef) conditions were instructed to complete the DRQ-B before distributing the money. In the play conditions, the instructions noted that participants would be allocating play money only. The instructions asked participants to distribute the money by placing it in labelled envelopes (and their own pockets in the real conditions), then to seal the envelopes and slide them through the communication tube in the side of each cubicle so the researcher could distribute them accordingly.

After completing the allocation task and depositing the envelopes, participants were asked to complete the last component. For participants in the T-Bef conditions, this entailed
Individual Differences and Justice-Decisions

completing the DRQ-P; for the participants in the T-Aft and the imagine self condition, it entailed completing the DRQ-A. After participants signalled they had completed the DRQ (A or P), the researcher collected their final questionnaires, provided verbal and written debriefing information, requested that they refrain from discussing the research with classmates, answered immediate questions, thanked them for their efforts and ushered each of them, one at a time, out of the research reception area in order to maintain anonymity for “all participants.” Complete debriefing information was mailed to all participants at the end of the study.

Imagine Self condition. Participants in the imagine self condition completed all components of the play money-reasoning after condition, with three modifications: 1) the beginning of each set of instructions was preceded by either the phrase or sentence: “Please imagine that....” or, “Please imagine yourself in the following situation,” 2) the “workers” sheets were photocopies, and 3) participants completed the task knowing that there were no other group members (see Appendix H).

Scoring and Analyses

Kohlberg’s Moral Maturity. The dilemmas on Kohlberg’s test were scored following the procedure outlined by Colby and Kohlberg (1987). The scoring procedure produces two scores: a) global stage scores from Stage 1 to Stage 5, and b) Moral Maturity Scores (MMS) also called “weighted average scores (WAS)” which range from 100 to 500 (Colby & Kohlberg, 1987, pp. 158-188). Although WAS were calculated for each of the three dilemmas, “overall” WAS were used for the analyses (KMM).
One very experienced scorer, blind to the purpose of the study, scored photocopies of all dilemmas. A second expert scorer, also blind to the purpose of the study and blind to the first set of scores, scored a random selection (25 male and 25 female) of the dilemmas. Interrater reliability was 98% agreement within 50 WAS points (one-half stage), 92% agreement within 33 WAS points (one-third stage), with an overall \( r(50) = .71, p < .001 \).

**Personality measures.** Both the IASR-B5 (Trapnell & Wiggins, 1990) and the NEO-PI-R (Professional Manual; Costa & McCrae, 1992) were scored in accordance with provided instructions.

**Decision Reasoning Questionnaire.** To confirm that there were enough moral issues recorded by participants for accurate and reliable scoring, I examined the DRQ protocols. Content analysis of the responses to the question "What are the main issues involved in this decision?" revealed that many issues which might be considered non-moral were invoked in the decision-making process. When I defined the moral/non-moral distinction by differentiating between clearly moral issues (e.g.: fairness) versus less clearly moral issues (e.g.: quality of work), only 25% of the issues were rated as clearly moral issues. Although 75% of the issues recorded by the participants to this question were obviously not moral in nature, overall, the 13 questions of the DRQ did provide sufficient material for scoring.

The DRQ was scored by the same two expert scorers who scored the MJI, in reverse roles, with both being blind to the specific hypotheses of the study. The scorers worked together with the DRQ protocols from two pilot studies--Pilot-study One \((N = 30)\) and Pilot-study Two \((N = 25)\)--to develop a reliable scoring system for the content-specific elements of
the instrument. Each DRQ was given both a global stage score (e.g., 1, 1/2, 2, 2/3, etc.) and a WAS which I called the Distributive Justice Moral Maturity score (DJMM). Inter-rater reliability was 94% agreement within 50 points (one-half stage), 76% agreement within 33 points (one-third stage) with an overall $r(50)=.68$, $p<.001$.

**Analyses.** To examine differences between the two hypothetical groups, I compared their mean scores on the two dependent variables. To examine the effect of time of reasoning (before and after) and consequence (real and play), I computed 2 X 2 ANOVAs on the two dependent variables. To examine differences between KMM and DJMM, I compared mean scores for each condition and overall. To examine the strength of the relation between KMM and DJMM, I computed Pearson correlations for each condition. To examine the effect of consequence (real, play, hypo), I computed one-way ANOVAs on the two dependent variables followed by planned comparisons.\(^4\) To examine the strength of the relation between KMM, DJMM, and the amount of money kept, I computed Pearson correlations. To examine the effect of sex and consequence, I computed 2 X 3 ANOVAs on the two dependent variables, followed by planned comparisons. To examine the strength of the relation between personality traits, moral reasoning, and behaviour, I computed Pearson correlations and multiple regressions.

**Results**

**Scores on Standard and Study Specific Measures**

Scores on the three standard measures were compared with published reports. In all cases, the scores reported for this sample were comparable to the published values.
Kohlberg's Moral Maturity. KMM scores are reported in Table 2. There were no sex differences within or across the three dilemmas on Kohlberg's test, nor were there any significant differences between any of the conditions, $F(5, 195) < 1$.

Interpersonal Adjectives Scale Revised - Big 5. The IASR-B5 scores are reported in Table 3. In their large study ($N = 941$), Trapnell and Wiggins (1990) reported small to moderate sex differences on four of the five domain scales. Only one small and significant sex difference was replicated in this sample (which had a much smaller $N$) with women having slightly higher Neuroticism scale scores than men, $t(199) = 2.72, p < .01$.

Revised NEO Personality Inventory. The mean domain scale scores, standard deviations, and the direction and size of sex differences were similar to those reported by Costa and McCrae (1992, Table B-3, p. 77). Three small and significant sex differences were found, with women scoring higher than men on the Agreeableness, Openness, and Neuroticism domains (see Table 4).

As expected, the correlations between the five related-content domains of the NEO-PI-R and the IASR-B5 were strongly positive and larger than for any of the remaining correlations (see Table 5). Although the IASR-B5 and the NEO-PI-R do not measure exactly the same constructs, analyses revealed striking similarity between their relations to KMM, DJMM, and the mean amount of money kept. Examination of the NEO-PI-R domain and separate facet scores in relation to the variables of interest did not reveal any findings that varied significantly from those related to the IASR-B5. For this reason, and because the IASR-B5 scores were available for the full sample, it seemed most parsimonious to report...
only the analyses with the IASR-B5 scores.

**Examination of Issues**

**Hypothetical first-person versus third-person.** Although the mean DJMM scores of the “Fred” group ($M = 290.4$, $SD = 57$) were higher than the mean DJMM scores of the “imagine self” group ($M = 273.1$, $SD = 25.8$), the difference was not statistically significant, $t(64) = 1.5$, **ns**. The mean amount of money kept for the “Fred” group ($M = 5.60$, $SD = 3.84$) also was not significantly different from the mean amount of money kept for the “imagine self” group ($M = 5.48$, $SD = 3.50$), $t(64) < 1$, **ns**. Because no significant differences were found between the Fred and imagine self groups, the two groups were collapsed into a single condition (hypo).

**Time of reasoning.** Time of reasoning had no effect on amount of money kept or on DJMM, both $F(1, 131) < 1$, **ns**, nor were there any significant interactions with consequence. Therefore, time of reasoning was collapsed across consequences (real and play) in all further analyses.

To summarize, preliminary comparisons showed that: 1) the “Fred” and “imagine self” group scores were not significantly different on either dependent variable; therefore, the two groups were collapsed into a single “hypo” condition; 2) time of reasoning also failed to produce any significant differences, so time of reasoning was collapsed across levels of consequence. All further analyses were conducted with three consequence groups: hypo ($N = 66$), play ($N = 67$), and real ($N = 68$).
Moral Reasoning

**Decision Reasoning Questionnaire scores.** As shown in Table 6, the overall mean DJMM score was positively correlated with but significantly lower than the overall mean KMM score.

A comparison among the Pearson correlations between KMM and DJMM scores in each group revealed interesting, but non-significant differences in the strength of the relation between the two measures. The predicted pattern, with the strongest relation between KMM and DJMM in the hypo groups, followed by the play groups, and the weakest relation in the real groups, was not found. The relation between KMM and DJMM tended to be stronger in the play groups, $r(67) = .40, p < .001$, than in either the real groups, $r(68) = .33, p < .01$, or the hypo groups, $r(66) = .27, p < .05$, but the differences between correlations were not statistically significant.

**Distributive Justice Behaviour and Reasoning by Group**

**Money kept by condition.** There was a significant difference between groups for the amount of money kept $F(2, 198) = 9.26, p < .001$; although, the pattern of scores was not exactly as predicted (see Table 7). In line with my prediction, participants in the real groups kept more money than participants in the play or hypo groups, but the difference between groups was significant only for the comparison between the real and play groups, $t(133) = 4.49, p < .001$, with the hypo group scores falling between the two other groups.\(^5\)

**DJMM scores by condition.** Contrary to prediction, participants in the hypo groups did not have the highest DJMM scores (see Table 7). Although there was a significant
difference between groups, \( F(2, 198) = 3.61, p < .05 \), the hypo group scores and the real
group scores were both significantly lower than the play group scores, \( t(131) = 2.39, p = .042 \), and \( t(133) = 2.05, p = .018 \), respectively.

**Relation between Moral Judgment and Behaviour**

**Money kept and DJMM.** The relation between DJMM and the amount of money kept across all conditions was in the predicted direction, but very weak, \( r(201) = -.15, p < .05 \), one-tailed. The correlation was strongest in the real consequence condition, \( r(68) = -.31, p < .01 \). The correlations in the other two conditions approached zero and were not significantly different from the real consequence condition correlation.

**Money kept and KMM.** The relation between KMM and the amount of money kept across all conditions was also in the predicted direction, but very small, \( r(201) = -.18, p < .01 \), one-tailed. Again the correlation was strongest in the real consequence condition, \( r(68) = -.31, p < .01 \), and the correlations in the other two conditions were non-significant; play group, \( r(67) = -.02 \); hypo group, \( r(66) = -.16 \). These correlations were not significantly different from the real consequence condition correlation. The global measure of moral maturity—KMM—was not more strongly related to the amount of money kept than was the more specific measure of distributive justice reasoning (DJMM).

**Sex Differences**

**Money kept.** Examining the effect of sex and consequence on the amount of money kept revealed a main effect for sex, \( F(1, 195) = 6.35, p < .05 \), with no significant interactions. Men kept significantly more money, \( N = 95, M = $6.10, SD = 3.83 \), than women did, \( N = \)
Individual Differences and Justice-Decisions

106, $M = 4.91$, $SD = 3.11$.

**DJMM scores.** Examining the effect of sex and consequence on DJMM revealed a main effect for sex, $F(1, 195) = 3.94$, $p < .05$, qualified by a sex by consequence interaction, $F(2, 195) = 5.45$, $p < .01$. However, the main effect for sex and the sex by consequence interaction were wholly due to the hypo condition sex difference, $t(64) = -4.21$, $p < .001$, with no DJMM differences between men and women in the play condition or in the real condition (see Table 8).

Comparing the differences between KMM and DJMM mean scores across consequence conditions highlights the extent of the sex difference in the hypo condition (see Table 8). Post hoc analyses revealed that men in the hypo conditions had much greater discrepancy scores between their KMM and DJMM scores, $M = 54.7$, $SD = 39.9$, pair-wise $t(29) = 7.50$, $p < .001$, than men in the other two consequence conditions, $M = 24.7$, $SD = 51.3$, pair-wise $t(64) = 3.89$, $p < .001$, or than women in the hypo condition, $M = 16.5$, $SD = 43.1$, pair-wise $t(35) = 2.29$, $p < .05$, and the other two consequence conditions, $M = 22.3$, $SD = 42.4$, pair-wise $t(69) = 4.40$, $p < .001$.

**Relation between Personality, Moral Reasoning, and Behaviour**

**IASR-B5 trait domains and KMM.** The KMM scores were not significantly correlated with IASR-B5 Conscientiousness, $r(201) = .06$, *ns*. The KMM scores were positively correlated with IASR-B5 Openness across all conditions, $r(201) = .33$, $p < .01$. The correlations for each consequence condition were: real, $r(68) = .31$, $p < .05$; play, $r(67) = .32$, $p < .01$; hypo, $r(66) = .40$, $p < .01$, with no significant differences between any pair of
correlations. No other correlations between KMM scores and IASR-B5 domains were significant.

**IASR-B5 trait domains and DJMM.** The DJMM scores were not significantly correlated with IASR-B5 Conscientiousness, $r(201) = -0.04$, *ns*. The DJMM scores, compared across consequence conditions, were weakly correlated with IASR-B5 Openness, $r(201) = 0.187$, *p* < .01. In the hypo condition, the DJMM scores were positively correlated with IASR-B5 Neuroticism, $r(66) = 0.27$, *p* < .05, and in the real condition, negatively related to IASR-B5 Conscientiousness, $r(68) = -0.30$, *p* < .05, directly opposite to my expectation.

**IASR-B5 trait domains and money kept.** Across all conditions (*N = 201*), only IASR-B5 Openness was significantly, although weakly, correlated with the amount of money kept for self, $r(201) = -0.182$, *p* < .01. Only one condition-specific correlation achieved significance; in the hypo condition, IASR-B5 Neuroticism was related to the amount of money kept, $r(66) = 0.235$, *p* < .05.

**Predicting Behaviour**

KMM were predictive of the amount of money kept, but very weakly, $R = 0.18$, $F(1, 199) = 6.77$, *p* < .05, accounting for only 3% of the variance. Including IASR-B5 Openness with KMM increased the predictive ability, still very weak, $R = 0.22$, $F(2, 198) = 5.16$, *p* < .01, but they accounted for 5% of the variance. Examination of the correlations between the amount of money kept and KMM, $r(201) = -0.181$, *p* < .01, and the amount of money kept and Openness, $r(201) = -0.182$, *p* < .01, revealed that KMM and Openness were nearly equivalent predictors of the amount of money kept. Because the correlation between KMM and the
amount of money kept was strongest in the real condition, \( r(68) = -.31, p < .01 \), I examined KMM and Openness with the amount of money kept for each group. For the real condition only, KMM were predictive of the amount of money kept, \( R = .31, F(1, 66) = 7.13, p < .01 \), and including Openness accounted for a small increase, \( R = .33, F(2, 65) = 3.86, p < .05 \).

**Discussion**

The data confirmed several previous findings and offered new insights into some aspects of the relation between moral reasoning, personality, and decision-making behaviour. As found in previous research: a) moral reasoning scores on Kohlberg’s test were higher than scores on the non-Kohlbergian dilemma; b) there were no sex differences in moral reasoning on any of the Kohlbergian dilemmas; c) the moral reasoning scores were moderately predictive of moral behaviour; and, d) the degree of relation between Kohlberg’s measure of moral reasoning and moral behaviour was mediated by the consequences of the behaviour.

New findings suggested that: a) the relation between moral reasoning and moral behaviour was mediated by the **personal** consequences of the moral decision required by the situation; b) giving participants time and explicit instruction to engage in moral reasoning before they made a moral decision had little effect on reasoning or behavioural decisions; c) context and consequence differences influenced moral reasoning and decision-making behaviour; d) men exhibited more self-benefitting behaviour than women did across all circumstances, yet showed similar levels of moral reasoning; e) the IASR-B5 Openness score was as predictive of moral behaviour as Kohlberg’s test across all conditions. Each of the
original issues will be discussed in turn.

**Issue 1: First-person versus third-person.** Consistent with past research (Krebs et al., 1994), I found that the extent to which the self is projected into hypothetical distributive justice dilemmas--third-person versus first-person--did not affect moral reasoning and behaviour. One explanation for this similarity is that both groups responded to a hypothetical dilemma. Although the instructions asked participants in one group to make a decision about a hypothetical character, Fred, and in the other group to make a decision about self (imagine self), the participants most likely responded to both dilemmas as hypothetical, and of little personal consequence. When Krebs et al. found no difference between first versus third-person perspectives on four of Kohlberg’s dilemmas, they argued that people probably use the same reasoning about all hypothetical dilemmas, whether first-person or third-person. It seems that the perspective people adopt when considering hypothetical dilemmas, whether based on major issues like life and law in Kohlberg’s dilemmas, or on a less serious issue like the distribution of $20, has little impact on the level of moral reasoning.

The same argument--no personal consequences arising from hypothetical dilemmas--can be applied to the failure to find any differences between the two groups for the amount of money kept.

**Issue 2: Time of reasoning.** Time of reasoning about a moral dilemma did not affect level of moral reasoning and resource allocation behaviour. Kohlberg and Candee’s (1984) model assumes that moral reasoning precedes behaviour. Kohlberg argues that because information in moral dilemmas is processed through stage-structures, the moral maturity of
the arguments will be the same, regardless of when the decision or the reasoning is recorded.

The failure to find any reasoning-before versus reasoning-after differences is not surprising when examined from a more social psychological perspective (reviewed by Saltzstein, 1994). If participants in the reasoning before conditions were anticipating their decision before reasoning--maybe a tentative decision is a necessary starting point in the reasoning process--they may well have been engaging in the same processes as participants in the reasoning after condition (i.e., post hoc justification).

The nature of the cognitive processes mediating decisions is open to question. People might not be processing the information in the dilemma through moral stage structures, but rather basing their decisions on some other rapid and unconscious cognitive process. We do not have direct access to the processes involved, so it is difficult or impossible to determine whether people make the decision first, then justify it; or, alternatively, they may complete a rudimentary analysis before the decision, then revise the reasoning after the decision--which could account for the justification effects reported in the social psychological literature. This leaves open the question of whether or not we can ever successfully manipulate a variable like time of reasoning.

**Issue 3: KMM and DJMM scores.** I predicted and found that KMM scores were higher than DJMM scores. Past research has consistently found that people score higher on Kohlberg's test than on more real-life like dilemmas (Krebs, Vermeulen et al., 1991). This finding is consistent with Kohlberg's (Colby & Kohlberg, 1987) contention that the MJI is designed to measure an individual's highest moral competence. The moderate correlations
found between KMM and DJMM scores show that moral judgment on Kohlberg’s test was related to the allocation dilemma reasoning, suggesting that participants invoked moral reasoning on the allocation task. The magnitude of the relation between moral reasoning on the MJI and moral reasoning on the DRQ is similar to the relations between scores on the MJI and scores on other types of dilemmas reported and reviewed by Krebs and his colleagues (Krebs, Vermeulen et al., 1991, Bartek, Krebs, & Taylor, 1993, Carpendale & Krebs, 1992, 1995, Denton & Krebs, 1990).

While the finding that the KMM scores were higher than the DJMM scores is consistent with Colby and Kohlberg’s (1987) contention that the MJI assesses the highest level of moral capacity, it challenges their “structure of the whole” assumption. Krebs, Denton et al. (1991) have argued that the additive/inclusive model (Levine, 1979), which posits that all stages remain available, can better explain the discrepancy between moral-reasoning capacity on hypothetical dilemmas and moral reasoning about real-life dilemmas.

**Issue 4: KMM and DJMM scores by group.** I predicted, but did not find, a specific pattern of correlations between the KMM and DJMM scores across conditions—from strongest to weakest correlations—the hypothetical group, the play group, then the real group. One possible reason for the failure to find the predicted relation is that the distributive justice dilemma constrained the level of moral reasoning used to justify the behaviour. Carpendale and Krebs (1992) and Krebs et al. (1994) have argued that some dilemmas pull strongly for one or two stages of moral reasoning, leading to a restriction of the variance in the reasoning scores. The allocation task used here was a simple distributive justice dilemma, based on the
issue of equality (Colby & Kohlberg, 1987), which could be interpreted as "strict equality" (stage 2) or as "equality based on deservingness" (stage 3), thereby pulling for only two stages, which could have restricted the variance. However, although the allocation dilemma did pull mainly for stage 2 or stage 2/3 reasoning, the variance in DJMM scores was no smaller than the variance on Kohlberg's test.

**Issue 5: Effects of consequences.** Consistent with an economical model, real consequences--real money as opposed to play or hypothetical money, within the same research context--induced participants to keep more money and to lower the level of moral reasoning used to justify the decision. However, contrary to expectation, participants in the hypothetical groups did not have the lowest money kept scores and the highest DJMM scores, compared to the other groups. These two issues are examined below.

The finding that real consequences affected resource allocation behaviour and justice-decision reasoning seems, on the face of it, to be consistent with what Carpendale and Krebs (1995) predicted, but inconsistent with what they found. Carpendale and Krebs predicted participants would make a selfish decision, but these investigators found that participants who distributed real resources (money) actually kept less for themselves and gave more to another (i.e., behaved more generously) than participants who distributed hypothetical money. Consistent with the findings of this study, Carpendale and Krebs found an inverse relation between money kept and moral reasoning; that is, the participants in their hypothetical group used significantly lower levels of moral reasoning to justify their relatively selfish behaviour than participants in the consequential condition used to justify
their relatively generous behaviour.

Design differences between the two studies might explain the disparate findings. Carpendale and Krebs (1995) interpreted their unexpected findings by suggesting that the increase in the salience of the effect of the allocation decision on the recipient might have induced participants to take the recipients’ perspective, reducing the likelihood of selfish behaviour. In this study, however, there was no equivalent of a “particular other” who could be victimized; instead, there were three anonymous “others,” who would know nothing about the allocation decisions. Therefore, the design of this study may not have increased the salience of the social consequences of selfish behaviour to the victim(s), allowing participants to behave “as if” there were no real monetary consequences for a single individual (diffusion of cost, both social and monetary). A related possibility is that in the Carpendale and Krebs (1995) study the consequences to the recipient were more severe than the diffused consequences across three recipients in this study.

I also did not expect the play group to keep less money than the hypo groups. One explanation for this finding is that the recipients of the allocation money were more salient in the play group than in the hypothetical groups, even though recipients were not as salient as they were in the Carpendale and Krebs (1995) study. Participants in the hypothetical Fred group completed a take-home package and the members of the imagine self group were repeatedly informed that there were no other real group members. In contrast, all participants in the play groups actually saw the other three group members, and even saw them enter separate research cubicles, after having seen them receiving $6.00 cash for participation. The
increased salience of the other group members in the play condition may have induced participants to behave relatively generously, compared to the participants in the hypothetical condition. Although participants in the real condition also saw the other group members, the pull of the self-benefitting consequence may have been sufficient to entice the participants to keep more money. To summarize, the sight of potential victims of participants' selfishness may have induced them to behave more generously than they believed they would when considering the decision hypothetically, but only when there were no real costs to them.

A final possibility is that the consequences considered by the hypo groups were more real than the consequences considered by the play group because they were asked to imagine Fred or themselves allocating real money, whereas instructions for the play group participants referred to "play money," with no corresponding request that they pretend it was real money. It is possible that the participants in the play condition considered the decision they made to be less consequential psychologically than the decisions made by the members of the hypothetical groups.

These arguments also entail explanations for the group differences in DJMM scores—that is, that participants in the play group displayed a significantly higher level of moral reasoning about the distributive justice task than the hypo or real groups did. Although I did not expect the exact pattern of group differences, I expected the DJMM scores to covary inversely with the amount of money kept, and they did. Participants in the play group were the most generous, and they invoked the highest level of moral reasoning to justify their behaviour.
Issue 6: Relation between moral judgment and behaviour. Consistent with Carpendale and Krebs (1995), I predicted and found an inverse relation between the amount of money kept and DJMM and KMM scores, but the correlations were very low and roughly equal. These findings are not surprising given the long history of weak relations between hypothetical measures of a construct and the behaviour the construct is supposed to predict. Blasi’s (1980) review of the literature reported only moderate correlations between MJI scores and a variety of behaviours.

One interesting finding was that the strongest correlations, with both DJMM and KMM scores, were for the real consequence conditions. All other correlations between the amount of money kept and moral maturity approached zero. I expected the relation between DJMM and behaviour to be stronger in the real consequence condition, but I expected the relation between KMM and behavioural decisions to be weaker than the relation between DJMM and behavioural decisions in general, and weakest in the real group in particular. There are two types of explanations for the equally strong relation between KMM and behaviour and DJMM and behaviour: a) Kohlberg’s test assesses a more general moral quality, like moral maturity, than the parameters of his dilemmas suggest, and b) the behaviour assessed in this study was, at best, weakly determined by any kind of moral reasoning. The weakness of the correlations favours the latter explanation.

Issue 7: Sex differences. As predicted, men behaved in a more self-benefitting way than women did across all conditions. This finding is consistent with the conclusion of Major and Deaux’s (1982) review of sex differences in justice behaviour research. Messé
and Callahan-Levy (1979) also found that women were more generous than men.

Although I found an overall sex difference reported for DJMM, with men, as predicted, having lower scores than women, this difference was significant only for the hypo group. In the hypo group, men, on average, scored nearly one half stage lower (42 WAS) on the DJMM than women did. As shown in Table 8, men and women displayed predominantly Stage 3 moral judgment in the play and real conditions. Women also displayed Stage 3 judgment in the hypothetical condition, but men scores predominantly at Stage 2/3. The question, then, is why did men invoke more Stage 2-based judgments, or fewer Stage 3 judgments in the hypothetical group? One possibility is that Stage 2 forms of judgment are more acceptable to men than to women, perhaps more consistent with their self-concept, but invoked only when contextual cues are weak.

**Issues 8 and 9: Relation between personality, moral reasoning, and behaviour.**
IASR-B5 Conscientiousness was not significantly related to KMM, DJMM, or the amount of money kept. I also examined the NEO-PI-R (Costa & McCrae, 1992) Conscientiousness facet scores (e.g., dutifulness, orderliness) to see if there were any significant correlations between the domain facets and moral maturity or behaviour; there were none. One possible explanation for the lack of relatedness between Conscientiousness and moral maturity or behaviour is the general nature of the construct as assessed by global measures of personality. Schwartz and Howard (1981) have argued that only a specific measure of responsibility, assessing the personal norm evoked by the situation, will be related to behaviour in a specific situation.
I found a relation between Kohlberg's measure of moral maturity and the Openness domain score of the IASR-B5. The relation between Openness and DJMM was weaker, although not significantly, than between Openness and KMM. One explanation is that the DRQ is a single issue measure of moral maturity, therefore, less related to IASR-B5 Openness items like abstract-thinking than a more philosophical measure like the MJI.

I found a weak inverse correlation between Openness and the amount of money kept for self. This parallels the relation between Openness and moral reasoning (KMM and DJMM), suggesting that Openness is related to the allocation decision-making processes in ways similar to moral reasoning. Although the correlations between Openness and the KMM were moderate and the correlations between both individual difference measures (IASR-B5 and MJI) and resource allocation behaviour were weak, the overlap was sufficient to suggest that they were related in similar ways to the allocation decision.

**Summary.** Predicting behaviour from abstract measures of a related construct has always been difficult. Predicting behaviour that could be considered moral has been a particular challenge. This research has demonstrated that both Kohlberg's test and the IASR-B5 measure of personality are somewhat predictive of allocation decisions and reasoning about those decisions on a distributive justice dilemma. It also has shown that situational factors influence both behaviour and reasoning about the behaviour, especially real consequences. Finally, it has confirmed that the more selfishly participants behaved, the lower the level of moral reasoning they used to justify the behaviour.

Other findings are more difficult to interpret. Although the social psychological
Individual Differences and Justice-Decisions

literature suggests that people may engage in behaviour first, then justify their behaviour later, I found no differences between the justifications used before or after behaviour. The lack of a difference, as Kohlberg would predict, might be due to the way people process information about moral issues, or simply due to unknown cognitive processes.

**Serendipitous finding.** The design of this study was unusual in that it included a measure of abstract moral reasoning capacity (MJI), a measure of moral reasoning about a real-life-like dilemma at three levels: hypothetical, no-context; hypothetical, within-context; a measure of real-life, within-context reasoning (DRQ); and a behavioural measure. One unanticipated finding was that scores on Kohlberg's test were more strongly related to the real consequence behaviour than to the hypothetical or play consequence behaviour. Although Kohlberg's test only accounted for 3% of the variance overall, it accounted for 10% of the variance for the real consequence group, suggesting that future attempts to predict moral behaviour will benefit from including measures of moral maturity.

The finding that Openness contributed to the prediction raises the possibility that other individual difference measures will account for some of the unexplained variance, especially if the measures used are specific to elements within the dilemma as opposed to the omnibus personality measures used here. In this study at least, Openness predicted behaviour as well as moral maturity.

Finally, the finding that people invoked more non-moral issues than moral issues when faced with the allocation dilemma in this study suggests that people sometimes interpret seemingly moral dilemmas in non-moral ways. Although the positive correlation
between KMM and DJMM indicates that the allocation dilemma did invoke some moral reasoning for most participants, it is possible that many of the participants would not have spontaneously defined and interpreted the situational information as constituting a moral dilemma if not primed to do so by the questions of the DRQ. Future research could assess the content of participants' spontaneous interpretations of the situation information before priming the participants to define the situation in moral terms.

**Limitations and Future Directions.** There are a number of limitations in the design of this study. The limitations inherent in the attempt to assess the utility of Kohlberg and Candee's (1984) model are of two types, namely those that might be corrected by modifying the design and those that resulted from the lack of clearly specified constructs in Kohlberg and Candee's model. After outlining some of the modifications which might apply to future research, the limitations with Kohlberg and Candee's model will be discussed.

One modification is needed to address the first-person versus third-person issue, an important question that remains unanswered. The distinction used in this research between the Fred and imagine self conditions was based on the wording of the questions on the DRQ, with the wording on the DRQ being changed from “...all the things you considered....” in the imagine self condition to “...all the thing you considered when deciding how Fred should...;” the focus remained on what the participant thought about the dilemma. Future research should attempt to move the focus toward a third-person perspective. For example, the question could be rephrased to read, “...all the things Fred should have considered....” Such a modification might reveal the expected differences in perspective-taking.
Kohlberg and Candee (1984) have argued that the relation between moral stage and/or deontic choice becomes stronger the closer one approaches Stage 5 moral reasoning. The weak-to-moderate correlations between moral reasoning and moral action reported here might result from the difficulty of designing a “real-life-like” dilemma that would yield the data necessary to assess the whole model. The current study used a distributive justice dilemma which seemed to pull for Stages 2, 2/3, and 3 reasoning only. This restriction of range was compounded by the restriction of range in people’s MJI scores--there were no participants with Stage 1 or Stage 5 KMM scores. This restriction of range would reduce the correlations between moral reasoning and moral action, leading to a misinterpretation of the actual relation between the two variables. Designing a “real-life-like” dilemma which would be meaningful to people at all stages of moral development and about which their moral reasoning about their behavioural response to the dilemma could be measured for comparison to their MJI scores might be the only way to assess the model fully. While such a dilemma would be a challenge to create, it might be the only way to adequately assess the utility of the model.

The second limitation inherent in Kohlberg and Candee’s (1984) model is a lack of clear explication of steps 3 and 4 in their model--judgments of responsibility or obligation and ego controls--especially the latter. Although Kohlberg and Candee have discussed ego controls as including intelligence, attention, and delay of gratification, they have not clarified exactly how these constructs relate to moral reasoning or moral behaviour. In this study, Openness was related to both moral reasoning and moral behaviour; whether or not Openness
is related to ego controls or intelligence is unclear. Currently, there is no way to differentiate between the intelligence necessary for moral reasoning--Colby and Kohlberg (1987) report moral stage and measured intelligence correlations of .37 to .60 for young adults--and “intelligence” as Kohlberg and Candee use the construct in relation to ego controls.

More functional models of the relation between moral reasoning and behaviour are needed to adequately explain real-life dilemma reasoning and behaviour. Such models need to look at both moral and non-moral aspects of the types of decisions people make when confronted with dilemmas. For example, the finding that Openness and KMM were equally correlated with the amount of money kept might suggest that there were moral and non-moral issues involved in the distributive-justice decision-making process, or, alternatively, it might suggest that the common component of Openness and KMM, that is, intelligence, is the only factor related to the behaviour. This research, in concert with previous research (Carpendale & Krebs, 1992, 1995; Krebs, Denton et al., 1991; Krebs, Vermeulen et al., 1991; Krebs & Denton, in press; Krebs et al. 1997) suggests that more interactive, broadly-based models other than Kohlberg’s will be better able to account for everyday moral behaviour. Many real-life dilemmas are not easily defined in terms of distributive justice principles. Real people in real dilemmas are not limited by abstract philosophical or psychological models--thank goodness.
References


Appendix A

Personality Measures: Conceptual Content and Related Research

One of the most widely used instruments used to assess personality is the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992). The five domains of the NEO-PI-R are captured by the labels neuroticism (N), extraversion (E), openness to experience (O), agreeableness (A), and conscientiousness (C), leading to acronyms like OCEAN or CANOE.

Trapnell and Wiggins (1990) extended the Revised Interpersonal Adjectives Scales (IAS-R, Wiggins, Trapnell, & Phillips, 1988) to include neuroticism, conscientiousness, and openness domains, and used the existing terms from the IAS-R Assured-Dominant and Unassured-Submissive scales to create the DOM scale, equivalent to the NEO-PI-R E (extraversion) scale, and the existing terms from the IAS-R Warm-Agreeable and Cold-hearted scales to create LOV scale, equivalent to the NEO-PI-R A (agreeableness) scale. The five scales of the new measure, the IASR-Big 5 (IASR-B5) were strongly correlated to the five domains of the NEO-PI-R (McCrae & Costa, 1989; Trapnell & Wiggins, 1990; Costa & McCrae, 1992).

Johnson and Ostendorf (1993) examined several five factor models and measures, suggested how yet another measure might resolve some of the disputes about the content of each factor, and suggested different factor labels. Their factor labels (with NEO-PI-R labels following), social communication (E), softness (A), constraint (C), freedom from negative affect (N), and creativity (O) highlight the range of possible constructs and behaviours that
might be related to five factor model scores.

Global personality measures like the IASR-B5 and the NEO-PI-R have been related to a number of other constructs (Johnson & Ostendorf, 1993). Kohlberg and Candee’s (1984) judgment of responsibility might be similar to the trait construct of Conscientiousness. Kohlberg and Candee’s (1984) ego-controls included intelligence, attention, and delay of gratification, traits which are similar to the Openness to Experience domain of the NEO-PI-R and the Openness domain of the IASR-B5. McCrae (1987, 1990, 1993-94) and Johnson and Ostendorf (1993) have argued that openness to experience includes aspects of intelligence, creativity, and divergent thinking. There should be a strong relation between moral maturity and conscientiousness and openness.
Appendix B, Section A

Design of the Decision Reasoning Questionnaires

The questionnaires were designed to elicit sufficient reasoning about the resource allocation decision, that could be accurately and reliably scored using procedures developed for scoring real-life dilemmas, to result in a dilemma moral maturity score (WAS). The DRQ questions were progressively more focused on moral issues associated with the allocation decision. For example, question #1 (DRQ-B)/[DRQ-A], "Please explain, as fully as possible, all the things you (think you should consider)/[considered] when deciding how to allocate the bonus money" is very general, while question #5, "(Are)/[Were] there any moral issues involved in this decision? If so, please explain what they (are)/[were] and why they are moral issues" is focused specifically on the moral issues.

In all but the Fred-Hypo group, participants completed the DRQ, either as a single instrument (DRQ-A) worded in the past tense, or in two parts, with the DRQ-B worded in the future tense and the DRQ-P worded in the past tense. For example, question #1 was altered to read, "Please explain, as fully as possible, all the things you considered when deciding how Fred should allocate the bonus money."

All DRQ's included a set of group task performance rating scales. These performance rating scales asked the participant to rate self and the three argument writers, using a 1 to 8 Likert scale (scale anchors and original direction in brackets), on the amount of time on task (very long to very short), the energy expended on task (very little to very much), the quantity of contribution (very large to very small) and the quality of contribution (very
low to very high). All performance scale scores were re-coded with higher numbers indicating higher performance. Only one of the performance scales related in a significant way to any of the independent or dependent variables, and that relation was based on self-reported time spent completing a task, with no measured confirmation.

A final pair of questions assessed the effectiveness of the manipulation—to confirm that participants believed that there were real consequences (keeping the money). For the paid participation groups, real and play (N = 135), with the value of 1 assigned to "positive they will NOT receive the money" and 8 assigned to "positive they WOULD receive the money," participants indicated they believed the money they designated for the other participants would be given to the other participants (M = 5.55, SD = 2.26). In the real group, the mean was significantly higher (M = 5.93, SD = 2.43) than in the play groups (M = 5.16, SD = 2.0, t(135) = 2.0, p < .05), with a full 75% of responses being above the mid-point (5 to 8). The second question asked them to guess what the study was about. Analysis of their open-ended responses show that 39% of the participants believed the study had something to do with the connection between moral reasoning and moral behaviour or fairness, while the remaining 61% of participants believed the study had something to do with the manifest content (group task performance and personality), personality across situations, or personality and selfishness/authority. There were only small differences between the real and play groups, with 41% of participants in the real groups and 35% of participants in the play groups correctly identifying the latent content of the study.
Appendix B, Section B

Decision Reasoning Questionnaire-A

Instructions

Please read and answer the questions in the order given. We are particularly interested in the reasoning behind your answers, so please outline your reasoning and elaborate as fully as possible. If you need additional space to answer questions, write on the back of the page.

1. Please explain, as fully as possible, all the things you considered when deciding how to allocate the bonus money.

2. What were the main issues involved in this decision? Please outline the issues in detail.

   It was impossible in this situation to give everyone an equal share of the bonus money, a constraint that is typical of many real-life situations. However, you did allocate the money, and we would like to understand your reasoning.

3. Please give all the reasons you can think of, in order of importance, for giving yourself more money than you gave the workers.

4. Please give all the reasons you can think of, in order of importance, for giving the workers more money than you gave yourself.

5. Were there any moral issues involved in this decision? If so, please explain what they were and why they are moral issues.

6. Did you consider any moral issues when you made your decision? If so, how did they affect your decision? Please explain.

7. Regardless of what you did, what is the most moral way of allocating the money? Please explain why it is the most moral way by referring to specific principles.

   For questions 8 and 9, remember, it was not possible to divide the money equally.

8. Earlier you gave reasons for giving the workers more money than you gave yourself. Now please give the best reasons you can think of why it would be fairer to give the workers more money, on average, than you gave yourself.

9. Please give the best reasons you can think of why it would be fairer to give the workers less money, on average, than you gave yourself.
10. Thinking back over the decision you made about how to distribute the bonus pay, what is the most responsible thing to do? Please explain fully.

11. Given the constraints of the situation, did you allocate the bonus money fairly? Why or why not?

12. In what sense could the way you allocated the money be considered fair? Please explain fully.

13. In what sense could the way you allocated the money be considered unfair? Please explain fully.

Appendix B, Section C

Performance Scales

Instructions

Please read and answer the questions in the order given.

14. Please rate the performance of each worker.

14.1 Time spent on task.

<table>
<thead>
<tr>
<th>Worker A</th>
<th>Very long time</th>
<th>Very short time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Worker B</th>
<th>(same scale as above)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Worker C</th>
</tr>
</thead>
</table>

14.2 Energy expended on task.

<table>
<thead>
<tr>
<th>Worker A</th>
<th>Very little energy</th>
<th>Very much energy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Worker B</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Worker C</th>
</tr>
</thead>
</table>
14.3 Quantity of contribution.

Worker A  
Very large amount

Worker B  
Worker C

14.4 Quality of contribution.

Worker A  
Very low quality

Worker B  
Worker C

15. Please rate your contribution to the group project.

15.1 Time spent on task.

Self  
Very long time

15.2 Energy expended on task.

(scales and anchors as above)

15.3 Quantity of contribution.

15.4 Quality of contribution.

*Please turn over and answer the questions.*
Some Psychological research has a bad name for misleading participants about the purpose of the study. While taking part in this study, what did you think it was about?

How certain were you that the workers would, in fact, receive the money you gave them (in the brown envelopes)?

Positive they would not receive the money.

Positive they would receive the money.

1 2 3 4 5 6 7 8

Appendix B, Section D

DRQ-B Questions only.

1. Please explain, as fully as possible, all the things you considered when deciding how to allocate the bonus money.

2. What were the main issues involved in this decision? Please outline the issues in detail.

   It was impossible in this situation to give everyone an equal share of the bonus money, a constraint that is typical of many real-life situations. However, you did allocate the money, and we would like to understand your reasoning.

3. Please give all the reasons you can think of, in order of importance, for giving yourself more money than you gave the workers.

4. Please give all the reasons you can think of, in order of importance, for giving the workers more money than you gave yourself.

5. Were there any moral issues involved in this decision? If so, please explain what they were and why they are moral issues.

6. Did you consider any moral issues when you made your decision? If so, how did they affect your decision? Please explain.
7. Regardless of what you did, what is the most moral way of allocating the money? Please explain why it is the most moral way by referring to specific principles.

For questions 8 and 9, remember, it was not possible to divide the money equally.

8. Earlier you gave reasons for giving the workers more money than you gave yourself. Now please give the best reasons you can think of why it would be fairer to give the workers more money, on average, than you gave yourself.

9. Please give the best reasons you can think of why it would be fairer to give the workers less money, on average, than you gave yourself.

10. Thinking back over the decision you made about how to distribute the bonus pay, what is the most responsible thing to do? Please explain fully.

Appendix B, Section E

DRQ-P Questions only.

1. Given the constraints of the situation, did you allocate the bonus money fairly? Why or why not?

2. In what sense could the way you allocated the money be considered fair? Please explain fully.

3. In what sense could the way you allocated the money be considered unfair? Please explain fully.
Appendix C

Third-person Hypothetical “Fred” Dilemma Instructions and Questions

INSTRUCTIONS

Please read the following dilemma and complete the tasks in order. Answer all questions as fully as possible. Whenever possible, elaborate on your answers (use the back of the page, if necessary), but feel free to say "see above" if you have already answered a question.

At a medium sized university, an undergraduate student named Fred volunteered to be in a Psychology study about group structure and the effect of personality on group performance. He was told that he and three other volunteers had been randomly assigned to a four person nominal group. A nominal group was described as being a group in name only; people work on components of a group project, but work separately.

Fred's group's task was to write arguments opposing the legalization of marijuana, with one member of the group being randomly assigned to write economic, one to write medical arguments, and one to write social arguments. The fourth group member was assigned the role of "coordinator," whose task was to pick the three best arguments, one from each content area (economic, medical, social).

Fred was randomly assigned to the coordinator role. Fred completed a personality questionnaire while the other group members wrote their arguments. After 20 minutes, the researcher gave Fred the economic, medical, and social arguments. Each member of the group had generated five arguments, which Fred found approximately equal in quality. He picked the best argument from each set. Following this, Fred was given another task, which was described as follows:

"You have an additional responsibility as coordinator that the other group members know nothing about. Your group has been given $20 bonus money. You will have to determine how much money each member of the group, including yourself, deserves for his or her contribution to the final project. You will receive an envelope containing $20 dollars in ten $2.00 bills and three white envelopes with identifying codes on them. (All participants will receive $6.00 for participating in the study; the $20.00 is bonus money over and above the $6.00 participation money.) Your task will be to allocate the money to yourself and to the other members of the group as you see fit by putting whatever amount you choose to give each person in the labelled envelopes. The money CANNOT be divided equally—you will have to decide the best way to distribute the money. Your decisions will be completely private and completely confidential: NO ONE (not even the researcher, who will code all the data you supply by your subject number) will
ever know how much money you gave to each of the other three group members or how much you kept for yourself (which is why you were instructed NOT to identify yourself by name). In fact, the other members of the group will never be told that you allocated the bonus money. They will simply be told they have been awarded some extra money and given an envelope with their code on it."

1a) How much money should Fred keep for himself? (circle one)
$0  $2  $4  $6  $8  $10  $12  $14  $16  $18  $20

How much money should Fred give to each other group member?

Economic Argument Writer
1b) $0  $2  $4  $6  $8  $10  $12  $14  $16  $18  $20

Medical Argument Writer
1c) $0  $2  $4  $6  $8  $10  $12  $14  $16  $18  $20

Social Argument Writer
1d) $0  $2  $4  $6  $8  $10  $12  $14  $16  $18  $20

The total money distributed should add to $20.

2) Please explain, as fully as possible, all the things you considered when deciding how Fred should allocate the bonus money.

3. What were the main issues involved in this decision? Please outline the issues in detail.

It was impossible in this situation to give everyone an equal share of the bonus money, a constraint that is typical of many real-life situations. However, you did indicate how you believed Fred should allocate the money, and we would like to understand your reasoning.

4. Please give all the reasons you can think of, in order of importance, for having Fred give himself more money, on average, than he gave the writers.

5. Please give all the reasons you can think of, in order of importance, for having Fred give the writers more money, on average, than he gave himself.

6. Were there any moral issues involved in this decision? If so, please explain what they were and why they are moral issues.
7. Did you consider any moral issues when indicating what you thought Fred should do? If so, how did they affect the alternatives considered? Please explain.

8. Regardless of what you believe Fred should have done, what is the most moral way of allocating the money? Please explain why it is the most moral way by referring to specific principles.

For questions 9 and 10, remember, it was not possible to divide the money equally.

9. Earlier you gave reasons for why Fred should give the writers more money than he gave himself. Now please give the best reasons you can think of why it would have been fairer for Fred to give the writers more money, on average, than he gave himself.

10. Earlier you gave reasons for why Fred should give himself more money than he gave the writers. Please give the best reasons you can think of why it would be fairer for Fred to give himself more money, on average, than he gave the writers.

11. Thinking back over Fred's decision about how to distribute the bonus pay, what is the most responsible thing for Fred to do? Please explain fully.

12. Given the constraints of the situation, did Fred allocate the bonus money fairly? Why or why not?

13. In what sense could the way Fred allocated the money be considered fair? Please explain fully.

14. In what sense could the way Fred allocated the money be considered unfair? Please explain fully.

15. Is the fact that the other group members will never know who allocated the bonus money important? Please explain fully.
Appendix D, Section A

Complete Participant Feedback--Sample: Fred Condition

Measures of Moral Reasoning & Personality Study

Thank you for participating in my research. You have provided data that will contribute to the success of my dissertation research. The feedback information below is specific to the "condition" you were assigned--the hypothetical, content comparison condition.

In the past, both moral judgment research and personality trait research failed to live up to the promise of measuring people on meaningful dimensions. Part of the problem seems to have been in the way researchers measured personality and morality. Early attempts at measuring personality traits seldom defined the traits in ways that were acceptable or testable. The relation of measured traits to other equally questionable measures of intelligence, motivation, or morality were tenuous at best. In addition, Psychology researchers from the 1920's to the 1970's tended to avoid any area of research where the variable of interest could not be measured with a ruler (the behaviourist approach). The lack of common definitions, observable behaviours, and strong correlations between constructs meant that the relation between personality, morality, and behaviour has been ignored for several decades. The research you participated in attempts to overcome some of the problems of the earlier research and show that there are meaningful relations between personality and moral reasoning. You completed three measures, two of which are standardized measures, and one that was designed specifically for this research.

The Moral Judgment Interview (MJI, Colby & Kohlberg, 1987) included the Heinz dilemma. The MJI is a standardized and widely used measure that assesses the developmental aspects of moral reasoning. Researchers working with this instrument use a 1500 page scoring protocol to classify the moral reasoning used by a participant on hypothetical moral dilemmas into one of five moral stages. Although there have been some concerns about the applicability of this measure to various populations, it is the most reliable and empirically established instrument of its kind.

The personality questionnaire, the Interpersonal Adjectives Scale Revised--Big 5 (IASR-B5, Trapnell & Wiggins, 1990) measures normal personality and focuses on five domains of personality. In other conditions of the research, participants also completed the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992). This measure of personality focuses on the same five domains of personality as the IASR-B5. The NEO-PI-R describes the five domains of personality with the labels neuroticism (N), extraversion (E), openness (O), agreeableness (A), and conscientiousness (C). Each of the five domains has six facets that define the domain. One example from the Openness domain, is Values (O6), or openness to values, which measures the willingness to question social, political, and religious values. People with lower scores tend to accept authority and honour tradition. People with high scores seem willing to question both societal and personal values.

The hypothetical "Fred" dilemma is different in content to the MJI dilemmas and allows
me to assess moral reasoning about something that might happen in everyday circumstances. One concern that researchers have about previous research that looked at moral reasoning across situations is whether or not the level of reasoning that non-MJI dilemmas evoke is the same as would be evoked by MJI dilemmas. If the level of reasoning is different for different situations, then using the MJI scores to predict behaviour might not have great utility. One prediction I am making is that the level of moral reasoning will be different across the different dilemma types.

This overall study is a complex design, with many conditions. The amount of money you believed Fred should keep for himself is one DV, with the MJI, Fred dilemma scores, and the IASR-B5 being predictors. The study is also correlational, with several general predictions and many specific predictions. One general prediction is that scores Fred dilemma will be positively correlated with scores on the MJI. I am also interested in the relation between specific facets of the IASR-B5 and the scores of the MJI.

Thank you for participating in this research. You have participated in a complex study that asked you to perform certain tasks, some of which were very difficult. I recognize, and wish to remind you, that everyone who participates in this research contributes equally to the research, regardless of the task completed. As I wish to guarantee anonymity for all participants, I ask that you DO NOT discuss the role to which you were randomly assigned with anyone who participated in the present study, or who might participate in the future.

I hope this information has helped you understand the goals and methodology of this research. Again, I assure you that all of your responses will be kept anonymous and confidential. You are welcome to contact me if you are interested in the final results of the study (please wait several months); please contact Russell Day (291-3003).

Appendix D, Section B

Participant Feedback--Questionnaire Study

Personality & Measures of Moral Reasoning Study

In the past, both moral judgment research and personality trait research failed to live up to the promise of measuring people on meaningful dimensions. Part of the problem seems to have been in the way researchers measured personality and morality. Early attempts at measuring personality traits seldom defined the traits in ways that were acceptable or testable. The relation of measured traits to other equally questionable measures of intelligence, motivation, or morality were tenuous at best. In addition, Psychology researchers from the 1920's to the 1970's tended to avoid any area of research where the variable of interest could not be measured with a ruler (the behaviourist approach). The lack of common definitions, the lack of directly observable behaviours, and the poor correlations between constructs meant that the relation between personality and morality has been ignored for several decades.
The research you have just participated in attempts to overcome some of the problems of the earlier research and show that there are meaningful relations between personality traits and verbal expressions of moral capacity. The newer measures you have just completed are both valid and reliable. The measure of personality is the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992). This measure of normal personality focuses on five domains of personality. The five domains of the NEO-PI-R are captured by the labels neuroticism (N), extraversion (E), openness (O), agreeableness (A), and conscientiousness (C). Each of the five domains has six facets that define the domain. One example from the Openness domain, is Values (O6), or openness to values, which measures the willingness to question social, political, and religious values. People with lower scores tend to accept authority and honour tradition. People with high scores seem willing to question both societal and personal values.

The Moral Judgment Interview (MJI, Colby & Kohlberg, 1987) is a standardized and widely used measure that assesses the developmental aspects of moral reasoning. Researchers working with this instrument use a 1500 page scoring protocol to classify the moral reasoning used by a participant on hypothetical moral dilemmas into one of five moral stages. Although there have been some concerns about the applicability of this measure to various populations, it is the most reliable and empirically established instrument of its kind.

This study is correlational. There are several general predictions and many specific predictions. One prediction is that the Conscientiousness domain of the NEO-PI-R will be positively correlated with higher Stage Scores on the MJI. I am also interested in the relation between specific facets of the NEO-PI-R and the Stage Scores of the MJI. For example, the Openness to Values facet of the NEO-PI-R should be related to moral stages concerned with looking at moral dilemmas from more than one perspective. There should also be some negative correlations, with Neuroticism being negatively correlated with Moral Maturity.

Thank You for participating in this study. I hope this information has helped you understand the goals and methodology of this research. Again, I assure you that all of your responses are anonymous and confidential. You are welcome to contact me if you are interested in the final results of the study: please contact Russell Day (291-3003).

(Please do not discuss this research with your friends until they have participated.)

Appendix D, Section C

Participant Feedback--Allocation Study

Personality and Group Task Performance Study

In the past, both decision reasoning research and personality trait research failed to live up to the promise of measuring people on meaningful dimensions. Part of the problem seems to have been in the way researchers measured personality. Early attempts at measuring personality traits seldom defined the traits in ways that were acceptable or testable. The
Individual Differences and Justice-Decisions

The relation of measured traits to other equally questionable measures of intelligence or motivation were tenuous at best. In addition, Psychology researchers from the 1920's to the 1970's tended to avoid any area of research where the variable of interest could not be measured with a ruler (the behaviourist approach). The lack of common definitions, observable behaviours, and strong correlations between constructs meant that the relation between personality and behaviour has been ignored for several decades.

The research you have just participated in attempts to overcome some of the problems of the earlier research and show that there are meaningful relations between personality and decision behaviour. The newer measures you have completed (or will complete) are both valid and reliable. One measure of personality is the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992). This measure of normal personality focuses on five domains of personality. The five domains of the NEO-PI-R are captured by the labels neuroticism (N), extraversion (E), openness (O), agreeableness (A), and conscientiousness (C). Each of the five domains has six facets that define the domain. The other personality questionnaire, the Interpersonal Adjectives Scale Revised--Big 5 (IASR-B5, Trapnell & Wiggins, 1990) is also valid and reliable. Like the NEO-PI-R, the IASR-B5 measures normal personality and focuses on five domains of personality. The task and the questionnaire, the Decision Reasoning Questionnaire (DRQ) were constructed and carefully pilot-tested for this research.

This study is a mixed design. The NEO-PI-R and IASR-B5 should predict the types of decisions being made. The study is also correlational, with several general predictions and many specific predictions. One general prediction is that scores on the Conscientiousness scale of the NEO-PI-R will be positively correlated with scores on the DRQ.

Thank you for participating in this study. I hope this information has helped you understand the goals and methodology of this research. Again, I assure you that all of your responses will be kept anonymous and confidential. You are welcome to contact me if you are interested in the final results of the study (please wait several months!!); please contact Russell Day (291-3003).

(Please do not discuss this research with your friends until they have participated.)
Appendix E

Procedural Details

**Description of bogus groups.** The eight group conditions were described as varying in size (some larger, some smaller) and structure (some nominal, some intact). Groups A and B were described as nominal four person groups. Some participants were assigned to Group A and others to Group B, with all participants informed that some members of each group might be in different research rooms.

**Description of nominal groups.** In addition to providing written information, the researcher defined a nominal group as a group in name only, with the example of the "new electronic cottage industries" being used to illustrate the concept. A nominal group was described as being one where the members never met, but all members worked on components of the same project and communicated electronically. In contrast, the traditional intact group was described as one that met in one location, face-to-face, and worked on projects together. The intact group versus nominal group research paradigm was described and used as justification for the "test of the hypothesis that isolated contributions to a group effort can be as large and valuable as the contributions made by participants in the face-to-face group condition--and it might depend, to some degree, on individual differences in personality."

**Instructions regarding hand-written arguments.** Each set of freshly written arguments was allegedly just completed by the other group members and was delivered with the instructions, "Please do not write on these, we need to make photocopies for scoring"
purposes--Thanks." The arguments were generated by participants in pilot study one, the content rated as being typical of undergraduate students, and hand-copied by several research assistants. Each set of hand-written arguments was examined after being used, discarded if soiled or marked in any way, and replaced with new hand-written copies (see examples below).

**Bogus distribution of envelopes.** The researcher collected and hid the real allocation envelopes, with one or two *bogus* argument writer pay envelopes then being prominently displayed in Writer A, Writer B, or Writer C slots of racks labelled Group A, Group B, to Group H in the common reception area. During the initial briefing session, the labelled rack was clearly visible, but no explanation was offered as to its purpose.
Appendix F, Section A

Allocation Study Instruction Summary

Personality and Group Task Performance

**Instruction Summary—Please Read Carefully**

You have volunteered for a study of personality, group structure, and task performance for which you will be paid $6.00. Of the eight possible group structures, you have been randomly assigned, along with three other volunteers, to a "4 person nominal group." A **nominal** group is a group in name only as opposed to an "intact" group. In a nominal group situation, group members work on components of some larger project, but members of the group work separately (some past research has found that people work more effectively on separate parts of a group task when each member works in isolation).

Your group task is to draft a statement opposing the legalization of marijuana. Your role will be assigned randomly, with three members of your group being Writers and one of you being the Coordinator. Each of the three Writers will spend 15 minutes drafting his or her part of a statement opposing the legalization of marijuana—Writer A on the economic implications of legalizing the drug; Writer B on the medical implications of legalizing the drug; Writer C on the social implications of legalizing the drug. After the three Writers draft their statements, they will give them to the Researcher to give to the Coordinator, and the Coordinator will decide which three arguments are the best arguments to include in the final statement.

While the three Writers in the group are drafting their statements, the Coordinator will complete a personality test (the Writers will complete the personality test while the Coordinator picks the best arguments). Your random assignment to your role is below.

You have been randomly assigned the role of **Coordinator**.

Please complete the demographic and personality questionnaires now. DO NOT record any identifying information on the forms. Once you have completed the personality questionnaire, the researcher will supply you with the "Writer's" arguments.
Appendix F, Section B

Ranking Task Instructions

#:_____________ Coordinator

**Personality and Group Task Performance**

Please read the following task information carefully and answer all ensuing questions. You will have only 15 minutes to complete this task.

The researcher has given you the arguments produced by the other members of your group. Please read the arguments supplied by the other group members carefully. Each of their arguments is identified by a code number, and your task is to evaluate and **pick the best 3 arguments**, one from each set of arguments.

Please read each of the statements from a set, pick the argument you consider the best and write the code number for it on the line beside Economic # 1 below (for example, Best Economic Argument: Econ.-4A Econ. = Economic; 4 = the fourth argument; A = Writer A). Do the same for the Medical and Social arguments.

---

**Code Number**

Best Economic Argument: ________________

Best Medical Argument: ________________

Best Social Argument: ________________

---

**Do you believe that Marijuana should be legalized in Canada? (circle one)**

<table>
<thead>
<tr>
<th>Definitely should BE legalized.</th>
<th>Definitely should NOT BE legalized.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
</tbody>
</table>

*After* you have picked the 3 best arguments and answered the question, please put the arguments and this sheet into the brown envelope, **seal it**, and complete the next task.
Appendix G, Section A
Allocation Task Instructions--Reasoning After/Real Money

Read the following information very carefully.

You have an additional responsibility as coordinator. Your group has been given $20 bonus money and your job is to determine how much money to give each member of the group, including yourself, for his or her contribution to the final project. You will receive an envelope containing $20 dollars in ten $2.00 bills and three white envelopes with identifying codes on them. (All participants will receive $6.00 for participating in the study; the $20.00 is bonus money over and above the $6.00 participation money.) Your task will be to allocate the money to yourself and to the other members of the group as you see fit by putting whatever amount you choose to give each person in the labelled envelopes. The money CANNOT be divided equally—you will have to decide the best way to distribute the money. Your decisions will be completely private and completely confidential: NO ONE (not even the researcher, who will code all the data you supply by your subject number) will ever know how much money you gave to each of the other three group members or how much you kept for yourself (which is why you were instructed NOT to identify yourself by name). In fact, the other members of the group will never be told that you allocated the bonus money. They will simply be told they may have been awarded some extra money and given an envelope with their code on it.

Would you allocate the money now by putting it in the appropriate envelopes? Please seal the white envelopes and slide the envelopes through the tube in the wall so the researcher can distribute them. You may base your decision about how to allocate the money on whatever considerations you want.

Thank You

Now that you have allocated the money, we would like to understand as well as we can how you thought about this task and what sorts of things you think are important in making such decisions. Therefore, would you please complete the Decision Reasoning Questionnaire now?
Appendix G, Section B

Allocation Task Instructions--Reasoning Before/Play Money

Read the following information very carefully.

You have an additional responsibility as coordinator. Your group has been given $20 bonus money and your job is to determine how much money to give each member of the group, including yourself, for his or her contribution to the final project. You will receive an envelope containing $20 dollars in ten $2.00 bills (in play money) and four white envelopes with identifying codes on them. (All participants will receive $6.00 for participating in the study; the $20.00 is bonus play money over and above the $6.00 participation money.) Your task will be to allocate the play money to yourself and to the other members of the group as you see fit by putting whatever amount you choose to give each person in the labelled envelopes. The play money CANNOT be divided equally—you will have to decide the best way to distribute the play money. Your decisions will be completely private and completely confidential: NO ONE (not even the researcher, who will code all the data you supply by your subject number) will ever know how much play money you gave to each of the other three group members, or how much you kept for yourself (which is why you were instructed NOT to identify yourself by name). In fact, the other members of the group will never be told that you allocated the bonus money. They will simply be told they may have been awarded some extra money and given an envelope with their code on it.

Before you allocate the money, we would like to understand as well as we can how you are thinking about this task and what sorts of things you think are important in making such decisions. Therefore, would you please complete the Decision Reasoning Questionnaire now?

Thank You

Please complete the allocation task, following the instructions below.

Would you allocate the play money now by putting it in the appropriate envelopes? Please seal the white envelopes and slide the envelopes through the tube in the wall so the researcher can distribute them. You may base your decision about how to allocate the money on whatever considerations you want.

Once you have completed allocating the money, please complete the pink questionnaire (DRQ-P).

Thank You
Appendix H

Hypothetical "Imagine Self" Instructions

Personality and Group Task Performance

Instructions

Please imagine yourself in the following situation. You have volunteered for a study on personality and task performance for which you will be paid $6.00. You and three other volunteers will form a "group." The researcher explains that the four of you are part of a nominal group (a group in name only--past research has found that people work more effectively on separate parts of a group task when each member works in isolation). The four of you are members of a group whose task is to draft a statement opposing the legalization of marijuana. You will pick your assignment randomly, with three of you being workers and one of you being the Coordinator. The researcher explains that each of the three workers will spend 20 minutes drafting his or her part of a statement opposing the legalization of marijuana--worker A on the economic implications of legalizing the drug; worker B on the medical implications of legalizing the drug; worker C on the social implications of legalizing the drug. After the three workers draft their statements, they will give them to the Coordinator, and the Coordinator will decide which order the individual arguments should be in the final statement--with the best arguments first.

While the three workers in the group are drafting their statements, the Coordinator will complete a personality test (the workers will complete the personality test while the Coordinator ranks the arguments).

You pick the role of Coordinator. The three other volunteers pick the role of workers. Please imagine that the three other group members are drafting their statements now, while you are completing the personality questionnaire.

Please complete the personality questionnaire now. DO NOT record any identifying information on the questionnaire response form. Once you have completed the personality questionnaire, the researcher will supply you with photocopies of the "worker's" arguments.
Individual Differences and Justice-Decisions

Read the following information very carefully.

Please imagine that you have an additional responsibility as coordinator. Your group has been given $20 bonus money and your job is to determine how much money to give each member of the group, including yourself, for his or her contribution to the final project. You are told that you will receive an envelope containing $20 dollars in ten $2.00 bills and four small brown envelopes with identifying codes on them. (You are told that you will receive $6.00 for participating in the study; the $20.00 is bonus money over and above the $6.00 participation money.) Your task will be to allocate the money to yourself and to the other members of the group as you see fit by putting whatever amount you choose to give each person in the labelled envelopes. The play money CANNOT be divided equally—you will have to decide the best way to distribute the money. Your decisions will be completely private and completely confidential: no one will ever know how much play money you gave to each of the other three group members or how much you kept for yourself (which is why you were instructed NOT to identify yourself by name). In fact, the other members of the group will never be told that you allocated the bonus money. They will simply be told they may have been awarded some extra money and given an envelope with their code on it.

Please imagine that the attached envelope contains real money. Would you indicate how you think you would allocate the money now by putting the play money in the appropriate small envelopes? Please place the small envelopes into the white envelope and seal the white envelope. You may base your decision about how to allocate the money on whatever considerations you want.

Thank You

Now that you have allocated the money, we would like to understand as well as we can how you thought about this task and what sorts of things you think are important in making such decisions. Therefore, would you please complete the Decision Reasoning Questionnaire now?
Footnotes

1 The term sex is used throughout this research to reflect the question asked on the
demographic questionnaire; Circle One: Male or Female. No attempt was made to
assess aspects of gender like masculinity or femininity.

2 These identical twins continued to communicate with each other in spite of repeated
requests to complete the tasks alone. When separated, they responded to some
components in “mirror-image” fashion, e.g., one kept $20, the other kept $0.

3 All quotes in the procedure section are from the scripts used by the researchers.

4 The exact p value is reported for planned comparisons where appropriate.

5 To illustrate the group differences, participants’ money kept scores were dichotomized
as being more self-benefitting when values were greater than $5.00, the value marking a
conceptually equal allocation in the distributive justice task, and less self-benefitting
when less than $5.00. Across all conditions, 93 (46.3%) of participants had less self-
benefitting scores, while 108 (53.7%) of participants had more self-benefitting scores.
The distribution of less and more self-benefitting scores was different by consequence
condition with more real condition participants having more self-benefitting scores (51
of 68; 75%) than hypo condition participants (33 of 66; 50%) or play condition
participants (24 of 67; 35.8%).
Table 1

Description of Groups, DRQ Format, Context, and Tasks

<table>
<thead>
<tr>
<th>Group-Name</th>
<th>N (M/W)</th>
<th>DRQ-Format</th>
<th>Context and Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYPO-Fred</td>
<td>15/20</td>
<td>DRQ-Fred</td>
<td>Individual, take-home package, participants responded to a hypothetical dilemma similar to Kohlberg's dilemmas, but based on “Fred” distributing “$20.00 bonus money.”</td>
</tr>
<tr>
<td>HYPO-Imagine Self</td>
<td>15/16</td>
<td>DRQ-A (A=after)</td>
<td>Individual package, group context, participants responded to a hypothetical dilemma similar to the PLAY-After dilemma, differing in that the instructions asked them to “Imagine yourself in the situation.”</td>
</tr>
<tr>
<td>PLAY-After</td>
<td>16/17</td>
<td>DRQ-A</td>
<td>Group context, groups of 4 participants received common initial instructions before being isolated in individual research cubicles. Participants were asked to reason about allocating “$20.00 play bonus money” after dividing play money.</td>
</tr>
<tr>
<td>PLAY-Before</td>
<td>16/18</td>
<td>DRQ-B (B=before) DRQ-P (P=post)</td>
<td>Group context, like PLAY-After above, except participants were asked to reason before and after (post-decision) deciding how to divide the “$20.00 play bonus money.”</td>
</tr>
<tr>
<td>REAL-After</td>
<td>18/17</td>
<td>DRQ-A</td>
<td>Group context, like PLAY-After above, except the “$20.00 bonus money” was real. Participants could keep up to $20.00 real money.</td>
</tr>
<tr>
<td>REAL-Before</td>
<td>15/18</td>
<td>DRQ-B DRQ-P</td>
<td>Group context, like PLAY-Before above, except the “$20.00 bonus money” was real. Participants could keep up to $20.00 real money.</td>
</tr>
</tbody>
</table>
Table 2

**Mean Moral Maturity across Sex and Kohlberg Dilemmas**

<table>
<thead>
<tr>
<th>Kohlberg's Dilemmas</th>
<th>Women N = 106</th>
<th>Men N = 95</th>
<th>Both N = 201</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>III (Heinz)</td>
<td>309.27</td>
<td>48.90</td>
<td>312.35</td>
</tr>
<tr>
<td>III' (Officer Brown)</td>
<td>317.80</td>
<td>32.70</td>
<td>325.79</td>
</tr>
<tr>
<td>I (Joe)</td>
<td>320.30</td>
<td>28.80</td>
<td>318.28</td>
</tr>
<tr>
<td>Total MMS</td>
<td>316.38</td>
<td>27.88</td>
<td>317.95</td>
</tr>
</tbody>
</table>
Table 3

**Mean IASR-B5 Domain Scores by Sex compared to Trapnell and Wiggins (1992)**

<table>
<thead>
<tr>
<th>Domain Scale</th>
<th>Women (N = 106)</th>
<th>Men (N = 95)</th>
<th>Both (N = 201)</th>
<th>Sex Diff. t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>DOM</td>
<td>4.95</td>
<td>1.07</td>
<td>5.00</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>(5.03)</td>
<td>(0.97)</td>
<td>(5.15)</td>
<td>(0.88)</td>
</tr>
<tr>
<td>LOV</td>
<td>6.40</td>
<td>0.81</td>
<td>6.21</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>(6.45)</td>
<td>(0.68)</td>
<td>(5.89)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>CONSC</td>
<td>5.73</td>
<td>1.03</td>
<td>5.51</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(5.74)</td>
<td>(0.95)</td>
<td>(5.59)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>OPEN</td>
<td>5.61</td>
<td>0.98</td>
<td>5.48</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>(5.53)</td>
<td>(0.78)</td>
<td>(5.41)</td>
<td>(0.72)</td>
</tr>
<tr>
<td>NEUR</td>
<td>4.80</td>
<td>1.19</td>
<td>4.36</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>(4.75)</td>
<td>(0.90)</td>
<td>(4.50)</td>
<td>(0.86)</td>
</tr>
</tbody>
</table>

Note: *p < .01. Trapnell and Wiggins (1992) reported values; N = 941. *Comparisons for which Trapnell and Wiggins (1992) report a significant sex difference.
Table 4

**Mean NEO-PI-R Domain Scores compared to Costa and McCrae (1992) Table B-3 for College-Age Individuals**

<table>
<thead>
<tr>
<th>Domain Scale</th>
<th>Women N = 86 (Norm-Scores)</th>
<th>Men N = 80 (Norm-Scores)</th>
<th>Both N = 166 (Norm-Scores)</th>
<th>Sex Diff. t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Extraversion</td>
<td>117.5</td>
<td>22.2</td>
<td>113.6</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>(123.9)</td>
<td>(17.7)</td>
<td>(116.7)</td>
<td>(18.3)</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>117.3</td>
<td>19.1</td>
<td>111.1</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>(117.2)</td>
<td>(15.7)</td>
<td>(107.4)</td>
<td>(16.2)</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>114.8</td>
<td>20.2</td>
<td>114.7</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td>(115.1)</td>
<td>(20.6)</td>
<td>(113.5)</td>
<td>(22.0)</td>
</tr>
<tr>
<td>Openness</td>
<td>126.6</td>
<td>20.7</td>
<td>120.2</td>
<td>19.9</td>
</tr>
<tr>
<td></td>
<td>(118.6)</td>
<td>(17.1)</td>
<td>(113.9)</td>
<td>(18.5)</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>101.9</td>
<td>27.1</td>
<td>92.0</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>(99.8)</td>
<td>(20.9)</td>
<td>(90.5)</td>
<td>(22.1)</td>
</tr>
</tbody>
</table>

Table 5

Pearson's Correlations between NEO-PI-R Domain scores and IASR-B5 Domain Scores (N = 166)

<table>
<thead>
<tr>
<th>IASR-B5 Domains</th>
<th>Extraversion</th>
<th>Agreeableness</th>
<th>Openness</th>
<th>Conscientiousness</th>
<th>Neuroticism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dom. (Extraversion)</td>
<td>.60</td>
<td>-.21</td>
<td>.36</td>
<td>.24</td>
<td>-.33</td>
</tr>
<tr>
<td>Lov. (Agreeableness)</td>
<td>.25</td>
<td>.70</td>
<td>.16</td>
<td>.04</td>
<td>-.12</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>.24</td>
<td>.05</td>
<td>.75</td>
<td>.16</td>
<td>-.10</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-.00</td>
<td>.07</td>
<td>-.05</td>
<td>.79</td>
<td>-.09</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.24</td>
<td>-.14</td>
<td>-.11</td>
<td>-.10</td>
<td>.78</td>
</tr>
</tbody>
</table>

Note: Absolute coefficients greater than .15 are significant at \( p < .05 \). The IASR-B5 labels Dom. and Lov. refer to the 16 item scales assessing dominance and nurturance. The additional IASR-B5 labels (Extraversion/Agreeableness) are for convenience only.
<table>
<thead>
<tr>
<th>Condition</th>
<th>KMM Score</th>
<th>DJMM Score</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Hypo</td>
<td>316.11</td>
<td>24.79</td>
<td>282.27</td>
</tr>
<tr>
<td>Play</td>
<td>318.11</td>
<td>34.36</td>
<td>302.34</td>
</tr>
<tr>
<td>Real</td>
<td>317.13</td>
<td>34.21</td>
<td>286.06</td>
</tr>
<tr>
<td>Full Sample</td>
<td>317.12</td>
<td>31.34</td>
<td>290.24</td>
</tr>
</tbody>
</table>

Note: * p < .01, ** p < .001.
Table 7

**Mean Money Kept Scores and DJMM Scores by Consequence**

<table>
<thead>
<tr>
<th>Reality of Consequence</th>
<th>N</th>
<th>Money Kept</th>
<th>DJMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Hypo</td>
<td>66</td>
<td>$5.55</td>
<td>$3.66</td>
</tr>
<tr>
<td>Play</td>
<td>67</td>
<td>$4.18$_a$</td>
<td>$2.22$</td>
</tr>
<tr>
<td>Real</td>
<td>68</td>
<td>$6.68$_a$</td>
<td>$3.98$</td>
</tr>
</tbody>
</table>

Note: Means that share a subscript are significantly different at the $p < .05$ level.
Table 8

Mean KMM and DJMM Scores by Sex and Condition

<table>
<thead>
<tr>
<th>Condition by Sex</th>
<th>KMM Scores</th>
<th>DJMM Scores</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>HYPO Men</td>
<td>30</td>
<td>313.83</td>
<td>27.96</td>
</tr>
<tr>
<td>Women</td>
<td>36</td>
<td>318.00</td>
<td>22.05</td>
</tr>
<tr>
<td>Men</td>
<td>32</td>
<td>324.41</td>
<td>39.31</td>
</tr>
<tr>
<td>PLAY Women</td>
<td>35</td>
<td>312.34</td>
<td>28.50</td>
</tr>
<tr>
<td>Men</td>
<td>33</td>
<td>315.42</td>
<td>36.28</td>
</tr>
<tr>
<td>REAL Women</td>
<td>35</td>
<td>318.74</td>
<td>32.58</td>
</tr>
</tbody>
</table>
Figure 1. Kohlberg and Candee’s (1984) Model of the Relation between Moral Judgment and Moral Action