

JUSTICE, FAIRNESS, AND MORAL DEVELOPMENT: DIFFERENCES IN THE  
GENERATION OF EXEMPLARS

by

Beatriz Irene Alvarado

---

Copyright © Beatriz Irene Alvarado 2012

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2012

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Beatriz I. Alvarado

entitled Justice, Fairness, and Moral Development: Differences in the Generation of Exemplars

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

\_\_\_\_\_ Date: 7/24/12  
Aurelio Jose Figueredo

\_\_\_\_\_ Date: 7/24/12  
Connie Beck

\_\_\_\_\_ Date: 7/24/12  
Judith Becker

\_\_\_\_\_ Date: 7/24/12  
W. Jake Jacobs

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_ Date: 7/24/12  
Dissertation Director: Aurelio Jose Figueredo

### STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Beatriz Irene Alvarado

### Acknowledgements

The author would like to acknowledge the contributions of Aurelio J. Figueredo and Bruce D. Sales toward the authorship of this paper. The author would like to thank research assistants: Robert Kincannon, Kelly Obert-Klugheit, Paul Vega, Kevin Vidal, and Emmitt Williams, for their months of hard work on this project. This paper is made possible by the emotional and academic support of Raymundo Alvarado, Connie Beck, Debra Cox-Howard, Arlene Fresquez, Stephanie Schneider, and countless others too numerous to name.

## TABLE OF CONTENTS

LIST OF TABLES .....	7
LIST OF FIGURES .....	8
ABSTRACT .....	9
INTRODUCTION .....	10
“Justice” Research .....	12
Current study.....	15
Hypotheses and Dimensions Coded.....	17
Sex and Moral Development .....	20
METHOD .....	24
Participants.....	24
Materials .....	24
Procedure .....	26
Rater Training .....	27
Story Rating .....	30
RESULTS .....	32
Descriptive Statistics.....	32
Multiple Regression Models .....	34
DISCUSSION .....	42
Limitations of the current analysis.....	49
CONCLUSION.....	52
Future directions .....	52

## TABLE OF CONTENTS (CON'T)

APPENDIX A: TABLES AND FIGURES .....	61
APPENDIX B: ORIGINAL CODING BOOKLET .....	77
APPENDIX C: CODING BOOKLET – FINAL REVISION .....	81
REFERENCES .....	87

## LIST OF TABLES

Table 1 .....	61
Table 2 .....	62
Table 3 .....	63
Table 4 .....	64
Table 5 .....	65
Table 6 .....	66
Table 7 .....	67
Table 8 .....	68
Table 9 .....	69
Table 10 .....	70
Table 11 .....	71
Table 12 .....	72

## LIST OF FIGURES

Figure 1 .....	73
Figure 2 .....	74
Figure 3 .....	75
Figure 4 .....	76

## ABSTRACT

Fairness and justice are often used interchangeably in socio-legal research. The goal of this study was to use 157 student-produced examples of either “injustice” or “unfairness” to determine whether differences exist in the content of the stories, and by extension, the definitions of these terms, and on participants’ scores on modified versions of Kohlberg and Gilligan’s levels of moral development. As hypothesized, the two terms were related, yet significantly different, with “unfairness” stories highlighting violations of equality, and “injustice” stories highlighting legal interactions and violations of equity. Sex differences were also found such that females were more likely to write stories rated high on unfairness and therefore equality, but no sex differences were found in level of moral reasoning reached by this sample. Future research is aimed at developing theory to explain differences, including the possible innate nature of “fairness” and environmental requirements leading to a concept of “justice”.

*Keywords:* justice, fairness, equity, equality, moral development

## INTRODUCTION

Understanding the characteristics of the terms fairness and justice is more than just an academic exercise. These are fundamental concepts that are used daily to describe the goals of the US legal system. In our democratic society, court systems, judges, lawyers, as well as law enforcement, and even parents are consistently asked to ably provide both fairness and justice to society and its citizens. To know what these people are asking for, it is essential that psychologists and legal scholars, as well as legal actors commit to research investigating the layperson's understanding of these terms, as well as how they differ from that of the legal professionals. This first step is required to be able to even consider delivering on these requests, and whether it is even possible to deliver on both fairness and justice simultaneously if they might be discrepant.

Recent research in the broadly defined area of "justice" finds that increased levels of perceived procedural fairness (beliefs) and/or procedural justice can increase compliance (behaviors) with court orders (Makkai & Braithwaite, 1996; Kitzmann & Emery 1993). Emery (2012) specifically focuses on the higher rates of compliance by both parents when child support and custody agreements are reached in mediation. He calculates huge potential savings in terms of the manpower, financial output, and individual suffering that would result from not having to enforce such orders, and argues that it is in the society's best interests to increase the perception of fairness and justice in all areas of the law (Emery, 2012). Other studies have shown that increased procedural fairness and outcome favorability can increase the perceived legitimacy of authority figures in the judgments of individuals (Van der Toorn, Tyler, & Jost, 2011), which may

be of special interest to government agencies that deal with groups that feel disenfranchised with the criminal justice system.

Increased feelings of justice and fairness among individuals appear to lead to many consequences that are considered beneficial by the above-cited authors. However, it is important to question *what exactly* is being increased when these feelings of either justice or fairness or both are being increased. Are these authors advocating policies that increase only the *perception* of justice or fairness, or instead endorsing policies that change the *actuality* of the equity or equality achieved by court decisions?

The first, and most important concern of this study, is how the common person, meaning the non-legal scholar, in North American society understands and applies the terms fairness and justice. Given the alleged importance of the constructs of justice and fairness in our legal system, it is unfortunate that no *empirical* research to date addresses what these terms mean to laypersons (non-legal scholars or actors), to different cultural/ethnic groups, or to lawyers, judges and law enforcement. Any possible discrepancies between the definitions of justice, fairness, or both, among these various groups can clearly constitute a source of social conflict.

This study provides an important first step in identifying these possible discrepancies, should they exist. It seeks to answer several fundamental questions: How does a common person think of fairness as opposed to justice? Are the terms synonymous, completely different, or overlapping? Are the terms fairness and justice guided by conceptions of daily life or infused with ideas from the justice system? If so, where did the justice system itself come from, if not from these everyday intuitions?

## “Justice” Research

Although the two words are distinct, U.S. legal scholars and philosophers often use the terms interchangeably (Finkel, 2001). For example, in Lerner’s (1980) well-known book on conceptions of justice, *The Belief in a Just World: A Fundamental Delusion*, the author states that people in Western society share a fundamental, yet unrealistic belief that the world is just, with individuals deciding what is just based on the “appropriateness” (p. 11) of a situation. Because Lerner goes on to define appropriateness as judgments based on the learned societal concepts such as what is “fair” and “just”, his underlying definition of a “just world” is circular. John Rawls, an eminent philosopher, further complicates the issue by defining fairness as a distinct subset of the overall concept of justice, thereby making the terms related, but separate (Rawls, 2001; 1971).

Other socio-legal scholars eschew the issue entirely and spend their time and energy parsing out the elements that make up the larger concepts of “fairness” and “justice”, while not furthering any definitions of the terms. For instance, the allocation of resources in a society falls under the purview of “Distributive Justice” or “Distributive Fairness” (per author preference) (see Fleischacker, 2004; Hegtvedt & Cook, 2001; Roemer, 1996, etc). Scholars in this field debate various philosophies and practices used to find the “best” method for spreading out goods and services in a defined society. Concepts such as equity, equality, societal norms, and even morality have been drawn into the discussion of distributive justice/distributive fairness. While these scholars take care to critically analyze the differences between “equality” and “equity” or various

morality practices, no such attempt is made to distinguish between the over-arching terms fairness or justice.

Another example of this conflation of terms can be found within the legal system. Whether the conflict is between two citizens who cannot reach an amicable agreement on their own (as in civil proceedings), or when a member of society has been harmed by the actions of another (as in criminal proceedings), the goals of the justice system are to reach fair and just *outcomes* for the petitioners. However, these goals appear to mean different things to different people. For example, the model of retributive justice presupposes that individuals who have been wronged will determine the fairness of the punishment based upon the severity of the crime. In some societies, so-called "cultures of honor", the principle of "an eye for an eye" is widely endorsed (Figueredo, Tal, McNeill, & Guillén, 2004; Vidmar, 2001). In current North American society, the more likely solution is that the perpetrator is removed from society via incarceration for a time commensurate with the crime (Walker, 2001). Restorative justice is a different path with the same overall goal. According to Morrison and Ahmed (2006) the final outcome should be one where all parties are "restored" to their original states by the efforts of the parties involved.

In direct contrast to the discussions of outcome justice/fairness, the study of procedural justice (see Tyler & Lind, 2001; Lind & Tyler, 1988; Thibaut & Walker, 1975, etc), focuses on the methods used for resolving disputes among parties. According to procedural justice theory, as long as an individual in a dispute feels that he or she was heard and treated fairly in the process of presenting their case, the individual will be

satisfied with the outcome, regardless of whether that outcome benefits him or her (Tyler, 1990; Lind, Kurtz, Mustane, Walker, & Thibaut, 1980). Thibaut and Walker (1975) take care to differentiate procedural justice from outcome justice. They define a “just procedure” as one in which conflict is resolved in a process decided on and guided by the disputants themselves (p. 1-2), thus a definition dependent on legal or quasi-legal “process”. However, in their explanation of procedural justice, Tyler and Lind (2001), rely heavily on the term fairness, describing the fairness of the procedures, the fairness of the interactions between parties, and the fairness of the legal system. In fact, in *The Social Psychology of Procedural Justice*, Lind and Tyler (1988) state: “...we frequently use the term *fairness* to describe a social situation in which norms of entitlement or property are fulfilled, and we will use the terms procedural *justice* and procedural *fairness* interchangeably in this book to describe such situations,” (p. 3).

However, some scholars have adopted the language of this procedure versus outcome discussion to redefine what fairness and justice mean. Sales and Shuman (2005) argue that the Federal Rules of Evidence, which have been adapted for use in the majority of states, consider justice and fairness distinct constructs. Fairness is the construct used to evaluate the *procedures* followed in the law while justice addresses the *outcomes*. While this attempt to add clarity is commendable, the conclusion is premature. There is no empirical evidence to support whether the layperson, or even whether the legal scholar, envisions these terms the same way. It would be unfortunate if a good-intentioned policy were put in place to increase fairness or justice when that is not what the public wants, needs, or expects. This study was designed to begin to resolve such definitional

confusion by putting forth and testing my own hypotheses about how non-legal scholars use these terms below.

The above research, while interesting, has done little to further the goal of defining fairness or justice, and especially in distinguishing between them, if any distinction is in order. It has, however, set the groundwork for understanding the component pieces of these terms. For example, if a person argues that an outcome is “unfair” or “unjust”, they may be judging that outcome on any number of variables, such as inputs and outputs, discrimination, or the process by which the decision was made. The current study examined fairness and justice in terms of the literature’s well-known and well-researched sub-headings – while paying particular attention to both similarities *and* differences. With further investigation, a causal model of these terms can be developed that will answer the fundamental question of whether the terms fairness and justice are interchangeable, distinct, or related in some way.

#### Current study

This study utilized many of the investigative techniques in Norman Finkel’s (2001) *Not Fair! The Typology of Commonsense Justice*. One of the pioneering methodologies laid out in his book is the use of the open-ended question as a means of gathering data. In many psycho-legal studies, participants are asked to evaluate the “fairness” or “correctness” of a decision based on a short vignette that the researcher selects. In these studies, the researcher then uses Likert-type scales or multiple choice questions to evaluate the responses of the participants. While this method lends itself to easy quantification and analysis, it often depends on preconceived labels of fairness,

justice, proportional punishment, or revenge. The current study, like Finkel's, does not introduce any extraneous labels or storylines that could contaminate the responses of the participants to the prompts.

According to Finkel (2001), it is much easier for an individual to remember instances of unfairness rather than fairness. If something is considered fair, it has occurred in the way things *ought to be*, and is not exceptionally noteworthy. However, when individuals feel that things have not occurred as they should, it goes against the common belief in the "just world" and people are more likely to remember it (Finkel, 2001; Lerner, 1980). Similarly, Auerbach (1983) stated: "Conceptions of justice that rest almost entirely upon legal procedure...still trouble ordinary citizens who have difficulty defining justice but know injustice when they receive it," (p. 143).

Based upon the above findings, this study asked participants to respond to an open-ended prompt (thereby injecting no prior definitional information). Study participants were split into two groups who either produced instances of unfairness or instances of injustice. Content analysis was then used to evaluate how individuals conceptualized the terms unfairness and injustice. An additional benefit of using this method is that by asking participants to provide their own examples of unfairness and injustice, participants generated their own "moral dilemmas". The ultimate goal of this procedure was to examine the generated stories in an objectively observable and measureable way.

One possible explanation for this behavior can be found in the literature surrounding psychological modules or mechanisms. MacDonald (2008) summarizes

psychological modules as domain specific areas of the mind that have aided our ancestors to solve adaptive problems in survival and reproduction over evolutionary time. These areas of the mind become specialized due to novel stimuli, such as how to deal with a new type of conflict that arises. According to MacDonald, implicit psychological modules are automatic, fast, probably evolved early in human development and are largely effortless. Explicit psychological modules are far more conscious and controllable, relatively slow, require effortful control of the individual, and are acquired by culture and formal knowledge. Because explicit psychological modules require higher order executive function, it is believed that these modules may have been socially selected towards a higher level of complexity only recently in human history, under the evolutionary pressures of a more challenging social environment.

Dealing with conflict in an efficient manner would aid in the survival of both the individual and the society over time. Therefore, an underlying theory to be tested is whether may be the case that “fairness” as a psychological module is implicit, while “justice” -- as defined in this study -- is explicit.

#### Hypotheses and Dimensions Coded

Based upon the wealth of “justice” research, my determination was that there are eight main areas that should be used to create a definitional model for both fairness and justice. The main hypothesis of this study was that fairness and justice are related, but not equivalent terms. The hypothesis was that when the layperson speaks of fairness or unfairness, he or she uses a broad stroke concept (an implicit psychological module), which includes legal issues, distribution, equality, equity, procedures, and outcomes. A

further hypothesis was that justice or injustice is conceptualized by a legal component (an explicit psychological module). According to this hypothesis justice is a subset of fairness, and the two prompts should result in different types of responses. If one group should report more stories than another, this study may be able to provide empirical evidence that one term is more encompassing than the other. Further this study analyzed the content of the stories in order to determine what components make up each term in the minds of the participants. For example, this study investigated whether the injustice prompt created more stories with legal content than the unfairness prompt. Therefore:

1. *Injustice* was operationalized as: *any and all things that require formalized institutions, such as laws, legal professionals, and law enforcement, in order to create a sense of balance and reciprocity among the members of a given society.*

The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of justice; (for more information on coding please refer to the Appendix and Method section)

2. *Unfairness* was operationalized as: *inclusive of the concepts of justice, but also involves equity, equality, and the sense that a procedure or outcome is right, proper or correct.* The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of fairness;

Because the legal system inherently provides an outcome and the procedures for reaching that outcome, legal scholars agree that both procedures and outcomes are important components of justice/fairness. As stated previously, Sales & Shuman (2005) consider fairness and justice as distinct legal constructs that are uncorrelated. According

to their theory, fairness is a construct used to describe the procedures followed in the law, and justice is a legal construct used to address the outcomes. If this hypothesis is correct, injustice stories should receive higher ratings on outcome content while fairness stories should receive higher ratings on procedure content. A competing hypothesis was also tested by this study. It is my belief that procedure and outcome should be equally represented in the generated stories, regardless of prompt condition. Therefore:

3. *Procedure* was operationalized as: *any and all content that discusses the who, what, how and why of a decision or process that leads to unfairness or injustice.*

The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of procedure;

4. *Outcome* was operationalized as: *any and all content that discusses unfair or unjust gains, losses, punishments, rewards, or conclusions to a dispute.* The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of outcome;

Equity and equality are considered components of distributive justice. As such, I expected to see similar numbers of stories of equity and equality violations in the unfairness and injustice prompts. However, distributive justice is necessarily an outcome measure, where some third party has decided what outcome the target should receive in a certain situation. Therefore, I expected to see a high correlation between stories of equity and equality and outcome, and no correlation between equity and equality and procedure.

Research into distributive justice has shown a number of sex differences (Hegtvedt & Cook, 2001). For example, in cases where sex is not mentioned as a factor

in their decision-making, female decision makers choose to give themselves less reward than their male counterparts, regardless of the amount of effort put into the task.

Distributive justice research assumes that females will be more likely to endorse an equality-based value system than males, who will be more likely to value equity based value system. Because the current study makes no mention of sex in either prompt, I expected these differences to hold true, such that stories generated by females should score higher on “equity” than “equality” measures and stories generated by males should score higher on “equality” measures than on “equity” measures. Therefore:

5. *Equity* was operationalized as: *any and all content that discusses unfair or unjust proportion of work to reward or inputs to outputs.* The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of equity.
6. *Equality* was operationalized as: *any and all content that discusses unfair or unjust treatment of things that should be equal. For example, preferential treatment or discrimination stories would violate someone’s sense of equality.*

The coders evaluated the vignettes on a scale of 0-4 describing how well the story relates to this definition of equality.

### Sex and Moral Development

Another well-known area of research that is highly related to fairness and justice is in the realm of moral development. Kohlberg postulated that humans move through sequential levels of moral reasoning, paralleling the passage through universal stages of moral development that existed outside of the rules of any particular culture (Kohlberg,

1981). His six stages of morality are still taught in psychology classrooms and are regularly tested on various participant groups. However, Kohlberg's theories of moral development are most often challenged on the grounds that they may not be truly universal but are culturally-bound and gender-biased. Gilligan argued that women reason differently than men when presented with moral dilemmas (Gilligan & Attamucci, 1988; Gilligan, 1982). She presents her own stages of moral development specifically aimed at women. Gilligan's position was that men and women conceive of fairness and justice quite differently, especially in contexts or circumstances that are of particular interest to women. In her book, *In Another Voice* (1982), she concludes that the male approach to morality has a "justice orientation" and that female morality is bound in a "responsibility orientation".

According to Kohlberg, moral reasoning is a universal trend that develops with age regardless of the type of moral dilemma or sex differences. As the study participants are all roughly the same age, no predictions are made. However, with regard to sex, Gilligan has asserted that females will respond differently than males in different contexts and simply because they are female. In terms of this study, it was predicted that Gilligan would expect to see sex differences with regard to the prompt, such that males will be more responsive to the justice prompt and females more responsive to the fairness prompt. Also, Gilligan would expect to see females score lower on Kohlberg's moral reasoning scales, and higher on her own, while Kohlberg would predict no such differences.

However, I hypothesized that there should be fewer differences among males and females than Kohlberg or Gilligan would expect. First, both Kohlberg and Gilligan's scales require the individual to systematically move through a series of stages from very concrete, self-involved thinking, to more abstract, global thinking. It was also hypothesized that there would be no sex differences found between the two models. Therefore, Kohlberg and Gilligan's respective coding schemes could be used (with only slight modifications) to determine whether there is evidence in support of one or both theories to research on moral development.

7. *Kohlberg's moral development level (Kmdl)* – Kmdl is based on six stages of thinking about moral dilemmas. This scale is already developed and validated, however, for the purposes of this study, the raters were kept blind to Kohlberg's terminology for each of the six stages. This decision is based on the fact that Gilligan also uses a six stage model of moral development which she also named, pre-conventional, conventional and post-conventional. In order to avoid confusion among the coders, they were only presented with the distinction between the levels, and used a rating scale from 1-6 to describe the moral reasoning level exemplified in each vignette. (see Appendix for more detail)
8. *Gilligan's moral development Level (Gmdl)* – as stated above, this study employed Gilligan's descriptions of the stages, but did not use the definitional terminology, instead asking the coders to rate each vignette on a scale from 1-6.

This study had both quasi-experimental and observational elements. The quasi-experimental portion is: (1) the prompt randomly given to each participant, and (2) the

sex of each participant. These were correlated with each of the eight dimensions. The observational element of this study is an analysis of the correlations among the eight dimensions coded as well as an examination of interaction terms with sex to determine if males and females conceive of these concepts or their components differently.

For more detail, Table 1 shows the expected correlations among all the predictor (quasi-experimental) and criterion (observational) variables. In some cases, there are alternative hypotheses being tested that are derived from rival theoretical perspectives. Those are coded as follows: (A) Alvarado, (SS) Sales & Shuman, (K) Kohlberg, and (G) Gilligan.

## METHOD

### Participants

Participants were volunteers from The University of Arizona who were enrolled in an introductory psychology class. As part of the course, introductory psychology students were required to participate in several hours of research. Participants signed up for this study via a psychology department-run website which advertised a survey that would investigate the “basic socio-legal attitudes of undergraduate students”. In return for one-hour of time, participants would receive one hour’s worth of research credit. The participants were: 97 freshmen, 36 sophomores, 10 juniors, 11 seniors, two graduate student participants, and one student who did not identify his class. Participants ranged in age from 18 to 45. Of the N=157 participants, 91 self-identified as female and 65 as male. Four self-identified as African-American, 9 as Asian-American, 123 as White or Caucasian, 15 as Hispanic, 3 as Native-American, and 1 as Pacific-Islander.

### Materials

**Demographic Information.** Participants in this study were asked to fill out a packet designed by the research team which gathered demographic information such as age, sex, education level, participants’ first languages, socio-economic status during childhood and at present. As discussed previously, race, ethnicity and ethnic group affiliation were considered vital at the time of data collection, therefore, each participant was further asked to complete measures of race, ethnicity, and information about the number of generations that the participant’s family had spent in the United States.

Unfairness or Injustice Stories. This study employs a between groups design where participants are assigned to one of two main conditions for the predictor variable; they were either asked to contribute instances of unfairness or instances of injustice. The prompts for the two conditions were adapted from Finkel's (2001) original prompt on fairness to meet our time constraints and incorporating the concept of justice. The two revised prompts were presented as follows:

For the unfairness condition:

“Among members of our society, a commonly heard lament is, “BUT IT’S NOT FAIR!” In this research we are examining what people mean by “it’s not fair.” We are going to ask you to think about, and then write in some detail examples of situations, events or circumstances that would lead you to say that this is an instance of unfairness.

On the blank pages that follow, we want you to detail as many situations as you can where the phrase “but it’s not fair!” applies, in your opinion. The situations can be personal, where they happened to you or a family member, or a friend; they can also be something you heard about or read about. For each instance,

- 1) Briefly describe the event;
- 2) tell why, in your opinion it was unfair; and
- 3) explain what would have to happen to make it fair.”

For the injustice condition:

“Among members of our society, a commonly heard lament is, “JUSTICE WAS NOT SERVED!” In this research we are examining what people mean by “injustice”. We are going to ask you to think about, and then write in some detail examples of situations, events or circumstances that would lead you to say that this is an instance of injustice.

On the blank pages that follow, we want you to detail as many situations as you can where the phrase “justice was not served!” applies, in your opinion. The situations can be personal, where they happened to you or a family member, or a friend; they can also be something you heard about or read about. For each instance,

- 1) briefly describe the event;

- 2) tell why, in your opinion it was unjust; and
- 3) explain what would have to happen to make it just.”

### Procedure

Groups of between 1-10 participants began the study in a large conference room where they were instructed to select their own seating. A research assistant then explained that in the hour that the students were there, they would be filling out a lengthy questionnaire about demographics and would then provide essay-type responses. Participants were directed to stay for the entire hour of the study and continue generating as many responses as possible. Participants were asked to read an informed consent form which instructed them of their rights as study participants and if they agreed to participate, to sign and date the form.

In a prior study using similar techniques (Alvarado, Sales, & Figueredo, 2006), a number of participants would generate a few examples and then request to leave the study after only a few minutes had passed. In order to reduce the number of individuals employing this tactic, the author and research assistants agreed that if a participant approached the researcher asking to leave because they felt they were “out of ideas” or for a similar reason, they were encouraged to give it “one more try” since they were being given one hour’s worth of credit. If a participant approached the researcher for a second time with the same request, they were then debriefed and allowed to leave. Any student requesting to leave for any other reason was dismissed immediately, and their results were not used (one occasion). At the end of 50 minutes all participants were debriefed as to the purposes of the study, and course credit was assigned via the psychology department website.

The questionnaire packets which contained the prompts were stacked in alternating sequence containing the “unfairness prompt” or the “injustice prompt” and were distributed to the participants according to their seating. The predictor variable manipulation was whether the participant was asked to generate instances of unfairness or instances of injustice.

Once the packets were collected, each participant’s packet was assigned a participant number and entered into a database by a different research assistant. All fairness or justice stories were typed into word processing documents in order to remove any bias that the participant’s handwriting or vocabulary may have introduced to the coding process.

#### Rater Training

Three undergraduate and one graduate research assistants were trained by the author to rate the stories produced by the participants. All research assistants were kept blind to the hypothesis for the duration of the study. Raters were trained based on a set of instructions (found in the Appendix) which asked them to provide eight ratings ordinal scale of the above operationalized definitions. It is important to note that the raters were trained to use only the definitions provided, and not their commonsense understandings of the terminology to provide a single ordinal number on each dimension. At the first meeting, the raters met with the author to go over the coding booklet. After all questions were answered, the raters were provided with ten sample stories and the author explained how each story should be rated according to the instructions. The stories used during the training period were collected from a set of samples that had been collected for a prior

study (Alvarado et al., 2006). As a group, the raters then went over the ten stories, providing their individual ratings for comparison, and the coding manual was revised to clarify definitions, instructions, and increase inter-rater reliability with the author. Each rater was encouraged to reason as the author would in order to create a “gold-standard”, or starting point of agreement from which to proceed.

After the first meeting with the author, raters were provided stories from the prior study in sets of ten. Raters would then meet as a group with the author to record rating agreements and disagreements on the eight dimensions to the “gold-standard”. Discussion would follow, including clarification on any particular rating, and revision of the coding manual would take place as needed. Any rater who systematically deviated from the group was asked to meet with the author for retraining to increase agreement rates.

Two different reliability coefficients were computed during the training period. For the first training period, an extremely rigorous method for determining inter-rater reliability was chosen. Prior to any discussion, all rater’s scores on the eight dimensions for each story were collected and compared with the author’s in an “all-or-nothing” comparison. Each match was coded as a 1 and each mismatch a 0. The “all-or-nothing” average was then computed to see how well each coder’s training evolved over the training period. This “all-or-nothing” approach did not factor in the complexity of the scales used. For example, according to the rating instructions, dimensions 1-6, (injustice through equality) could have any of 5 possible ratings, while dimensions 7 & 8, (Kmdl or Gmdl), could have 6 possible ratings. Any near-matches, for example, a rating of 1 from

the author and 2 from rater X would produce a 0, or “no match” despite being extremely close on a scale from 0-4 or 0-6. This rigorous standard for training was chosen for ease of analysis, and showed that over time, average rater agreement with the author increased over the 3 month training period from about 34.4% to about 46.3% (see Fig. 1 for a graphical representation this process). At this point, the raters were required to re-rate each story from the training set in order to measure their intra-rater reliabilities as well as their inter-rater reliability – without any influence from the author.

The second type of reliability coefficient was calculated with the less rigorous, but far more accurate General Linear Model (GLM) procedures in the SAS statistical package for psychological research (SASInstitute Inc., 2011). The GLM models allowed for data analysis that incorporated “near matches”, or partial agreements into the data analysis, as well as taking into account the number of choices each rater had for each of the eight scales per story. These ratings were then designated as Time 1 (Training period) and Time 2 (Reliability Check). (The re-rating period lasted approximately one month.) The results of the reliability check can be found in Table 2. Of particular note, intra-rater reliabilities from Time 1 to Time 2 ranged from .954 to .818, which shows a high degree of internal consistency over the four month period. Of interest are the inter-rater reliabilities from Time 1, which range from .910 to .580, with an average of .691 over the four raters. This range was expected due to the length and nature of the training process as well as the number of possible ratings each dimension could have. It was also expected that dimensions Kmdl & Gmdl would have lower inter-rater reliabilities due to the additional rating option and their adaptations for this study.

Time 2 inter-rater reliabilities ranged from .815 to .540, with an average of .668 over the four raters. Although the slight drop in reliability was unexpected, comparisons on the individual dimensions show that training focused on raising reliabilities on dimensions Kmdl and Gmdl, perhaps to the detriment of the reliabilities on other dimensions. The overall inter-rater reliability average decreased by .023, showing remarkable consistency from Time 1 to Time 2, given that there were four raters. In addition, intra-rater reliability was extremely high and each rater had become more consistent with the author's ratings over time. It was therefore considered safe to proceed with the data collected for this study.

### Story Rating

In order to control for the variability in the number of stories generated by each participant (which ranged from 1-21), only the first story generated was rated by the trained research assistants. Each rater was responsible for roughly 70 stories. Each story was rated by a minimum of two raters at least once. Certain stories were given to all raters to check for inter-rater reliability, and some stories were repeated in order to check intra-rater reliability over the analysis period, which lasted approximately four months. This resulted in a minimum of 2 ratings per dimension per story and a maximum of 8 ratings per dimension per story.

The reliability data for the analysis period can be found in Table 3, which shows intra-rater reliability ranging from .992 to .953, with an average across the four raters as .970, showing extremely high consistency within each rater over the analysis period. The inter-rater reliability for the analysis period ranged between .896 and .715 with an overall

average of .785, which is considered high and acceptable for four raters. Among those stories that were rated for a second time, (primarily for the purpose of collecting data on intra-rater reliability) the inter-rater reliability ranged between .914 to .619, with an average of .701. All of these data show that over the analysis period, the raters were even more consistent than they were during the training period.

## RESULTS

### Descriptive Statistics

The sample of 157 participants that were included in the study created 863 stories (5.5 instances/per participant), with large variability among participants ranging from 1 to 21 stories.

The 81 participants who received the unfairness prompt produced 550 stories, or 6.79 stories per person. However, the 76 participants who received the injustice prompt created only 313 stories, or roughly, 4.12 stories per person. This difference seems to indicate that there is something about injustice that is less accessible to the participant, reducing the number of overall stories produced.

The sample of participants included 91 females that created 543 instances of unfairness (5.90 instances per participant), and 65 males that created 320 instances of unfairness (4.92 instances per participant).

As discussed above, each participant contributed one story that was rated between two to eight times. In order to begin the statistical analysis, the first step included aggregating the data from the four hypothesis-blind raters into a single, workable number. This was accomplished by taking the mean score of the ratings and standardizing the number according to how many ratings it had received, ranging from 2-8, (see Method section, Story Rating subheading). Once an aggregate score was achieved, the SAS statistical package (SASInstitute Inc., 2011) was used to begin the statistical analysis with a bivariate correlation matrix.

*Bivariate Correlations.* According to the correlation matrix (see Table 4), Injustice and Unfairness are related, although not in any extreme way. At a correlation of  $r = .25$ , it is fair to say that the two terms share some characteristics, however, they are NOT the same concept. This has large implications for this study, which will be discussed later.

Unfairness is also significantly correlated with Outcome ( $r = .55$ ). This correlation indicates that a person's views on Fairness are possibly colored by the *outcome* of the situation, and likely not as much by the process. The correlation between Injustice and Outcome ( $r = .28$ ) was not as high as that of Unfairness and Outcome. Injustice, as defined by this study (stories including interactions with the law and legal actors) shows low level correlations with Procedure, Outcome, and Equity, with  $r$ 's ranging between .28 and .36. These correlations appear to show that all of these components are likely important to the concept, but none is dominant in the way that Outcome was for Unfairness.

As one might predict from the literature, Procedure and Outcome are uncorrelated. This finding lends support to other research that shows that high levels of perceived procedural justice can create a satisfactory or unsatisfactory experience without regard to whether the individual wins or loses his or her case (Walker, 2001; Lind & Tyler, 1988; Lind, et al., 1980).

Unfairness is correlated to Equality ( $r = .29$ ) while Injustice is not. However, both are correlated to Equity (Equity to Unfairness [ $r = .36$ ]; Equity to Injustice [ $r = .36$ ]). This correlation shows that the story contents of Injustice appear to be driven by "proper"

inputs and outputs than by equal treatment for equal circumstances. This finding is somewhat unexpected, but will be discussed further later in the paper.

With regard to Kmdl and Gmdl, it is important to note that the only significant and large correlations of the predicted elements were found between Equality and with each other (Equality to Kmdl [ $r = .39$ ]; Equality to Gmdl [ $r = .46$ ]) In fact, the highest correlation in the entire matrix is between Kohlberg and Gilligan's scales ( $r = .75$ ). This finding shows that there is less of a distinction between the two scales than may have been previously reported.

#### Multiple Regression Models

In order to conduct a causal analysis of any differences found in the instances of unfairness, a series of hierarchical regressions was performed that is the conceptual, if not the technical, equivalent of a sequential canonical analysis (Figueredo & Gorsuch, 2007; Figueredo & Sage, 2007; Gorsuch, 1991). This analysis requires that criterion variables are analyzed according to a predetermined, hypothesized causal order. The choice to use this method was made because it was anticipated that the criterion variables (the ratings on each story) measure psychological processes that are hypothesized to causally impact each other when activated. For example, a story that rated extremely high on the unfairness scale would be likely to score very high on either procedure or outcome and high on equity or equality, or both. Therefore, the criterion variables were entered one at a time into a system of multiple hierarchical regression equations. This process allows each previous criterion variable to act as a predictor for the next. With this procedure, the estimated effect of each predictor is limited only to the direct effect on the successive

criterion variables. This statistically controls for any of the indirect effects of the predictor variables due to the criterion variables hypothesized to be causally prior.

It is important to note that there is no implication being made that one rating is causing another rating, or one word influencing another word, instead, the prediction is that the ratings on these key terms represent internal psychological mechanisms that are hypothesized to operate on each other, and the observations capture an indirect reflection of these mechanisms in the verbal behavior recorded by the raters. Nevertheless, ratings on these terms will be described in the presentation that follows as acting directly and indirectly on each other, because to phrase each sentence with greater precision (in terms of the underlying psychological processes) would produce such extremely complex sentences that the arguments would be too difficult to follow.

A generic form for the equations used in this type of hierarchical multiple regressions is:

$$Y1 = X1 + X2 + X3$$

$$Y2 = Y1 + X1 + X2 + X3$$

$$Y3 = Y2 + Y1 + X1 + X2 + X3$$

where the Y's are equal to the criterion variables, and the X's the predictor variables.

In this analysis, the multiple regression of all the criterion variables was created according to the previously discussed hypothesized relationships. In order to create the most stringent tests of the hypotheses, the causal order of the criterion variables were as follows: 1) injustice (INJUSTICE), 2) unfairness (UNFAIRNESS), 3) procedural fairness/justice (PROCEDURE), 4) outcome fairness/justice (OUTCOME), 5) equity

(EQUITY), 6) equality (EQUALITY), 7) Gilligan's moral development level (GMDL), and 8) Kohlberg's moral development level (KMDL). It should be noted that two hierarchical regressions were run, one with (KMDL) first, the second with (GMDL) first. The equation that placed (GMDL) first yielded the most explanatory and parsimonious model, and is therefore, reported here. (For a graphical representation of the cascade model, please refer to Fig. 2.) Predictor variables included type of story (unfairness/injustice) (TYPE), and sex (SEX). Inputting these variables into general linear models led to the following equations:

$$1) \text{ INJUSTICE} = \text{TYPE} + \text{SEX} + \text{TYPE} * \text{SEX};$$

$$2) \text{ UNFAIRNESS} = \text{INJUSTICE} + \text{TYPE} + \text{SEX} + \text{INJUSTICE} * \text{SEX} + \text{TYPE} * \text{SEX};$$

$$3) \text{ PROCEDURE} = \text{UNFAIRNESS} + \text{INJUSTICE} + \text{TYPE} + \text{SEX} + \text{UNFAIRNESS} * \text{SEX} + \text{INJUSTICE} * \text{SEX} + \text{TYPE} * \text{SEX};$$

$$4) \text{ OUTCOME} = \text{PROCEDURE} + \text{UNFAIRNESS} + \text{INJUSTICE} + \text{TYPE} + \text{SEX} + \text{PROCEDURE} * \text{SEX} + \text{UNFAIRNESS} * \text{SEX} + \text{INJUSTICE} * \text{SEX} + \text{TYPE} * \text{SEX};$$

$$5) \text{ EQUITY} = \text{OUTCOME} + \text{PROCEDURE} + \text{UNFAIRNESS} + \text{INJUSTICE} + \text{TYPE} + \text{SEX} + \text{OUTCOME} * \text{SEX} + \text{PROCEDURE} * \text{SEX} + \text{UNFAIRNESS} * \text{SEX} + \text{INJUSTICE} * \text{SEX} + \text{TYPE} * \text{SEX};$$

$$6) \text{ EQUALITY} = \text{EQUITY} + \text{OUTCOME} + \text{PROCEDURE} + \text{UNFAIRNESS} + \text{INJUSTICE} + \text{TYPE} + \text{SEX} + \text{EQUITY} * \text{SEX} + \text{OUTCOME} * \text{SEX} +$$

PROCEDURE\*SEX + UNFAIRNESS\*SEX + INJUSTICE\* SEX +  
TYPE\*SEX;

7) GMDL = EQUALITY + EQUITY + OUTCOME + PROCEDURE +  
UNFAIRNESS + INJUSTICE + TYPE + SEX + EQUALITY\*SEX +  
EQUITY\*SEX + OUTCOME\*SEX + PROCEDURE\*SEX +  
UNFAIRNESS\*SEX + INJUSTICE\* SEX + TYPE\*SEX;

8) KMDL = GMDL + EQUALITY + EQUITY + OUTCOME + PROCEDURE +  
UNFAIRNESS + INJUSTICE + TYPE + SEX + GMDL\*SEX +  
EQUALITY\*SEX + EQUITY\*SEX + OUTCOME\*SEX +  
PROCEDURE\*SEX + UNFAIRNESS\*SEX + INJUSTICE\* SEX +  
TYPE\*SEX;

Note that among the criterion variables, in addition to the previously mentioned ratings, interactions by sex were also investigated. For example, the term INJUSTICE\*SEX checked whether being male or female interacted with ratings of injustice. This type of interaction term was used throughout the model to determine whether males and females think about fairness or justice fundamentally differently. These interaction terms also helped investigate whether there were sex differences in reasoning about these terms according to both Kmdl and Gmdl.

The effects of receiving the unfairness prompt or the injustice prompt INJUSTICE was modeled first, given the predication that the prompt would cause the participants to vary in the content of stories they report. UNFAIRNESS was modeled next, as it was meant to be directly compared to INJUSTICE, first as a manipulation

check, to be sure that the study participants were writing about the assigned task, and second, to directly compare whether those who received the injustice prompt had more legal content in their stories than participants in the unfairness group, as predicted. In turn, it was hypothesized that stories high in equity content would be low in equality content and vice versa, and that stories high in procedural content would be low in outcome content and vice versa. Finally, both Kohlberg and Gilligan's models were investigated for their content and for comparisons to each other, predicting that males would perform higher on Kohlberg's scales and females would perform higher on Gilligan's scale.

Each equation was then entered into the SAS (SASInstitute Inc., 2011) statistical program. In this section, I will only report the significant parameters of each equation in this section, where  $p < .05$  (for specific p-values, see the corresponding table specified below). All other parameters are excluded from the Result section as their effect sizes are not significantly different than zero. For a full listing of all variables, see the corresponding table as listed below. Full output files are available from the author.

*Injustice.* See Table 5. ( $R^2 = .243$ ) This equation shows that there was a significant influence of the predictor variable, the type of story (TYPE) and INJUSTICE ( $\beta = .480$ ). Participants who were assigned to write stories about injustice, as opposed to unfairness produced a story that was fundamentally different on the INJUSTICE scale. Put another way, participants in the injustice condition wrote stories with more legalistic content than individuals in the unfairness category. For this model, no significant sex differences were found.

*Unfairness.* See Table 6. ( $R^2 = .101$ ) Given the prediction that unfairness and injustice are terms used in different ways, INJUSTICE was modeled first. In turn, the ratings for UNFAIRNESS were predicted to affect the types of stories that were written, by way of equity, equality, procedure, or outcome content, as well as their scores on Kmdl and Gmdl. Therefore, the criterion variable, UNFAIRNESS was examined next. Findings were that both INJUSTICE ( $\beta = .203$ ) and the SEX ( $\beta = .202$ ) of the writer influenced UNFAIRNESS ratings. According to the analysis, those stories that are rated high on injustice also tended to be rated high on unfairness, possibly indicating that the concepts are similar, if not the same. In addition, stories written by females tended to receive higher scores on unfairness than stories written by males. These sex differences will receive greater attention in the Discussion section.

*Procedure.* See Table 7. ( $R^2 = .189$ ) According to this analysis, the criterion variable PROCEDURE, was significantly influenced by stories that are rated high on UNFAIRNESS ( $\beta = .235$ ) and high on INJUSTICE ( $\beta = .183$ ). A close look at the beta-weights shows that the relationship is slightly higher for UNFAIRNESS than for INJUSTICE, however, violations of procedure are important to both terms.

*Outcome.* See Table 8. ( $R^2 = .386$ ) The regression results for OUTCOME were similar to those for PROCEDURE in that the only significant predictor variables were UNFAIRNESS ( $\beta = .579$ ) and INJUSTICE ( $\beta = .248$ ). Those stories that were rated high on UNFAIRNESS and high on INJUSTICE influenced a high rating on OUTCOME violations. However, the effect size for OUTCOME was much larger  $R^2 = .386$  than for PROCEDURE, where  $R^2 = .189$ . This relationship showed that while both procedure and

outcome significantly influenced UNFAIRNESS and INJUSTICE, OUTCOME accounted for more of the variance in the model.

*Equity.* See Table 9. ( $R^2 = .266$ ) The equation for the criterion variable, EQUITY (violations of the balance of inputs to outputs), had the highest number of significant predictor variables of all the equations. According to the regression model tested, both OUTCOME and PROCEDURE were significantly positive causal influences ( $\beta = .181$ , and  $\beta = .108$ , respectively). Stories with high INJUSTICE ratings (stories with legal content) also causally influenced high ratings on the EQUITY scale ( $\beta = .171$ ), which measures equity violations. It was also noteworthy that stories with EQUITY content tended to be reported more by males than by females ( $\beta = -.155$ ). This finding may not be surprising given Kohlberg's theories, which will be covered in greater detail later.

*Equality.* See Table 10. ( $R^2 = .152$ ) The criterion variable, EQUALITY produced some intriguing results. First, the relationship between UNFAIRNESS and EQUALITY ( $\beta = .261$ ) was consistent with the causal hypothesis that higher ratings on UNFAIRNESS would tend to lead to higher ratings on violations of EQUALITY. However, the relationship of INJUSTICE was negative ( $\beta = -.207$ ), such that higher ratings of INJUSTICE (legal content) led to *lower* ratings of equality violations; or, in the reverse, that stories written with less legal content influenced higher ratings of equality violations. This was the only instance of a negative causal influence of one of the "violation" predictor variables in any of the models. This finding suggests that violations of EQUALITY are the domain UNFAIRNESS, not INJUSTICE.

*Gmdl*. See Table 11. ( $R^2 = .297$ ) For *Gmdl*, the only significant predictor variables were EQUALITY ( $\beta = .465$ ) and the interaction term EQUITY\*SEX ( $\beta = .336$ ). Violations of EQUALITY exerted a strong causal influence on *Gmdl*, such that these stories create a higher rating on her scale. For the interaction term, the relationship indicated that the combination of female writers producing a story that was high on violations of EQUITY also led to high ratings on GMDL.

*Kmdl*. See Table 12. ( $R^2 = .633$ ) A significant predictor variable for *Kmdl* was *Gmdl* ( $\beta = .719$ ). In this case, the causal influence of *Gmdl* was extremely high, as expected due to the high bivariate correlation ( $r$ ) observed. Because of this fact, it should be noted that two hierarchical regressions were run for this final stage of the cascade model, one with *Kmdl* first, the second with *Gmdl* first. The equation that placed *Gmdl* first yielded the most explanatory and parsimonious model, and is therefore, reported here. Other significant predictor variables for high ratings on *Kmdl* were high INJUSTICE ratings ( $\beta = .075$ ), and the interaction term TYPE\*SEX ( $\beta = .318$ ), where females who received the “injustice” prompt also tended to rate highly on *Kmdl*. Graphical representations of all significant effects can be found in the appendix in Figures 3 (without sex differences) and 4 (sex differences only).

## DISCUSSION

The most important finding from this research is support for the primary hypothesis that the words injustice and unfairness are not understood to be synonymous by many persons other than legal scholars, as would be suggested by the research of Tyler, Lind, or Lerner. By using the multiple competing hypotheses of legal scholars, legal philosophers, and my own suppositions, there is now preliminary data to support the hypothesis that “fairness” terminology, while inclusive of “justice” terminology, is not the same concept as justice. Though the correlation matrix finds the terms highly related, it should be noted that when prompted, individuals assigned to the injustice category produced stories that had much higher levels of legal content than the stories produced by the individuals in the unfairness category. This pattern holds for both males and females.

These results are generally consistent with the hypothesis advanced that the term "justice" is subsumed in common language under the more encompassing term, "fairness" – which is the exact reverse of the definition posited by Rawls. This finding implies that the interchangeable use of the terms fairness and justice in the “justice” literature is inaccurate at best and misleading at worst. When a legal scholar speaks of the death penalty in terms of “increasing procedural fairness”, this might be better understood by an audience as “procedural justice” due to the legal context of the decision-making process. Clarity in this area is not only important for communication amongst the legal community, but in communication with policy makers and society at large. When socio-legal scholars and lawmakers talk about increasing “justice and fairness” are the same concepts being discussed? Are authors taking into account inputs and outputs, procedures

and equality, or are they only discussing the realm of the legal system? Are jury members thinking in terms of what would be most fair, or most just?

Socio-legal scholars find it important to identify any possible differences between a respondent's understandings of complex terms like fairness and justice because they assume a direct link between a person's understanding of these concepts and his or her actions. I hypothesize that the same mechanisms that underlie the observed verbal behavior of participants in legal research might also control certain non-verbal behaviors in the form of actions such as: 1) individuals making decisions on juries, or 2) the behavior of non-jurors to avoid coming into contact with the criminal justice system by virtue of a better understanding of the principles under which it operates. For example, English and Sales (2005) discuss at length how high levels of confusion over jury instructions (a matter of *justice* as defined by this study) may lead many jurors to abandon the legal technicalities of a case and render a verdict based on their commonsense understandings of the instructions whether or not they are accurate. In the legal realm, English and Sales argue that with a lack of comprehensible legal guidance, jurors' decisions may include too many factors that are legally irrelevant. The authors argue that this could lead to inconsistent and capricious behavior by juries.

Another example of the link between verbal and non-verbal behavior can be found in cases of jury nullification. In his book, *Commonsense Justice*, Finkel (1995) examines several cases where juries appear to ignore the evidence and instructions presented to them, returning a verdict that instead reflects the personal values of the jurors. In studies of jury nullification, the socio-legal scholars that Finkel reviews find

that jurors' deliberations may regularly take them outside of the limitations of their court instructions and into larger considerations of fairness (as defined by this study). In either case, miscomprehension of instruction, or the deliberate application of jury nullification, the authors claim that the behavior of jurors in alignment with their beliefs can create unanticipated consequences for the legal system and for the public at large when a certain result was expected and not received. For example, the acquittal of the police officers in the Rodney King case is considered by many to be the main factor in sparking the 1992 LA riots (Klotkin & Friedman, 1993). The 1992 LA riots themselves may be considered an example of non-verbal civil disobedience, where illegal acts were committed by laypeople against both persons and property because their basic sense of fairness was violated by the behavior of the justice system. While it might be reasonable to suppose that the underlying psychological mechanisms that lead to verbal behavior may also lead to similar patterns of non-verbal behavior, this study did not include any data on the extent to which they may do so, or on which specific non-verbal behaviors might be affected. These important questions are a matter for future research to clarify.

The findings from each of the other equations in the statistical model serve to fill in the details of what separates the terms justice and fairness. This study has already addressed the high relatedness of the terms, such that when something is considered unjust, it is likely to also be considered unfair. Although this finding holds for both females and males, stories that are authored by females score higher on this study's unfairness scale than stories authored by males. In isolation this sex difference may simply be due to extraneous variables, or rater error, however, in combination with other

sex differences, (e.g. stories with high equity violations tend to be reported by more males than females.) this study provides evidence for both Kohlberg and Gilligan's moral development theories (if not their measurement systems) – with some adjustment to the vocabulary used by both.

Historically, Kohlberg's theory was developed first, and so will be addressed first here. Kohlberg's moral development measure relies heavily on reasoning that has often been characterized as having a "law and order", or "justice orientation" (Gilligan, 1982). In the lower levels, proscribed behaviors are deterred by punishment and fear of authority, while higher levels emphasize the need for all members of a society to follow laws. Gilligan developed her moral development measure because she felt that females were scoring lower than males on Kmdl due to its lack of emphasis on relationship maintenance and societal responsibility. Her model was developed to address the difference. However, both Kohlberg and Gilligan's moral development measures follow a "stage" model similar to Piaget's. Higher scores on both scales are achieved as participants move from concrete thought processes to abstractions. Both move from self-interest to "betterment of society". Over the years, various studies have shown conflicting results for sex differences on both scales. Even in this study, the most robust conclusion that can be made from the two moral development measures is that they are predictive of each other. In other words, regardless of sex, higher scores on one measure will lead to higher scores on the other. The conclusion is that the scales themselves are similar to the point of being indistinguishable. This is not to say however, that Kohlberg or Gilligan were wrong in their underlying theories, but that their scales are not

adequately capturing the theorized sex differences that may be present. The possible ramifications of this evidence to the moral development literature will be discussed in the following section.

Next, it is important to address the relative weight that individuals assign to procedures when discussing justice and fairness because of the wealth of research that has been devoted to this topic since Thibaut and Walker (1975) first coined the term “Procedural Justice”. Thibaut and Walker used justice and fairness interchangeably, however, Lind and Tyler (1988) become more specific with the terms. [See section title “Procedural Justice as Fairness” (p. 30)]. The findings suggest that procedure *is not* its own type of justice or fairness, but is an integral part of the overall concepts of both unfairness and injustice. Another specific example that was the assertion of Sales and Shuman (2005) that fairness can be defined as using proper procedure and outcome as justice. This study’s findings refute this statement. “Proper” or “correct” procedures help us decide what is fair and what is just, but do not define one term over the other.

However, when discussing the relative weight of the importance of outcome to justice and fairness, note the finding that outcome accounts for more of the variance in both fairness and justice than procedure. The predictors, unfairness and injustice account for approximately 18.8% of the total variance in procedure, when used as the criterion variable, but the predictors unfairness and injustice account for 38.6% of the variance in outcome, when used as the criterion variable. This finding suggests that while procedure is a component of fairness and justice, outcome is the more salient concept in judgments of the terms, which refutes some of the stronger claims made about the importance of

procedural justice. Moreover, unfairness has a much higher outcome component ( $\beta=.609$ ) when compared to the procedure component ( $\beta=.295$ ).

Outside of the legal realm, this finding essentially confirms that while good procedures may play a role in soothing a poor outcome for a particular person, most individuals place more emphasis on a positive outcome. This could easily pose problems for any authority figure who must decide when one party gets something that another party does not. For example, parents may give a bike to an older sibling, reasoning with the younger sibling that he is too young. Regardless of the good procedure, the younger sibling is likely to see the decision as unfair because of the unequal outcome. Parties interested in increasing actual or perceived fairness or justice, in the justice system may run into similar problems. Because the United States justice system is definitively adversarial, winners and losers abound. If legal actors are truly concerned with the increasing the justice that is administered, the current emphasis on increasing procedural justice may be misplaced. The individuals in this study show more interest in reaching a good or proper outcome than the procedures involved in reaching those decisions. Policy makers and justice administrators may be better served devising ways that both parties in a dispute feel that the *outcome* served his or her own interests.

As discussed above, equity is a concept borne of the distributive justice legacy. It is characterized by proportional balance between what a person puts in and gets out of a particular system. In this study, both outcome and procedure influence a person's sense of equity. From a distributive justice point of view, this finding makes sense because the process by which distributive decisions are made affects what any single person receives

from society. Examples of equity disputes can be seen in arguments over social security, welfare, and even fee-for-service business.

The equity dimension was also causally influenced by the injustice prompt, but not by unfairness prompt. This finding creates a further distinction between fairness and justice, but more importantly it has implications for the legal system. In order to score highly on the injustice measure, the story generated had to have legalistic content. Therefore, the analysis suggests that in the justice system, individuals are highly concerned with proportionality. Stories that are rated high on legal content appear to be driven more by “proper” inputs and outputs than by equal treatment for equal circumstances. Again, this makes sense when considering the heavy emphasis placed on “appropriate” punishments for crimes, and the constitutionally protected avoidance from “cruel and unusual” punishment.

Equality is another term that is borrowed from distributive justice and is a fundamental tenet of the criminal justice system. While equity is concerned with work to reward relationships, equality is the sense that all members of the society will be treated in the same way. This study shows that unfairness is significantly correlated to equality, while injustice is not. The implication of this correlation is that the participants in this study are more concerned with equality in everyday life, but not as concerned with equality inside of the legal arena. According to these results, high ratings on unfairness influence high ratings on violations of equality. Again, this makes perfect sense; if things are not equal, they are unfair. However, low ratings on the injustice scale influence high ratings on equality violations. Put another way, the participants in this study indicate that

when the law or legal actors are involved, there are lower ratings of equality violations. It appears that the expectation of equal treatment under the law is being met, at least in the minds of this study's participants.

In total, these findings suggest that violations of equality fall primarily in the domain of unfairness, but not injustice. Life in the "real world", separate from the legal realm, is a place where people "ought to be" treated equally, and when they are not, the situation is considered unfair. In direct contrast, violations of equity are a main component of injustice, but not unfairness. For those concerned with the legal system, this study provides evidence that individuals feel that people "ought to be" treated based on their individual efforts, and when they are not, it is not just. This is a problem for a criminal justice system that is set up to treat similar crimes similarly, without regard to individual differences.

#### Limitations of the current analysis

This study suffers from many of the same limitations in generalizability that applies to almost any study or experiment using college undergraduates as a population. The main characteristics of this sample was that they were young, well-educated, usually upper-middle class, mostly Caucasian and skewed toward females. As such, the results of this study cannot extend to all members of society, nor should it. It is my hope that future research in this area will allow sampling from other segments of society that allow for specific areas of generalization.

The participant was instructed by the prompt to write a story, and explain what made it unfair or unjust. This explanation is the basis for the content that was rated on

the eight dimensions. Therefore, we are moving from the hypothesis that the content is causally related to the fairness or justice condition, which then triggers the psychological module most related to that dimension. For this study, I hypothesized that what is internal to the story itself is the cause of the dimensional relationships, not what is internal to the individual.

While it is possible that the characteristics of the individual (their traits) may influence the stories that they choose to write (stories about discrimination or violations of human dignity), this study did not code for specific story narrative. Additionally, only the first story from each participant was rated. The fundamental problem being that the characteristics of the story were confounded with the characteristics of the individual. Other than the sex of the writer, it was not possible to draw conclusions about what types of people write what types of stories, whether the content of one story influenced the content of other stories written by the same person, or whether a particular story was written due to the characteristics of the individual or the subsections of society to which that individual belongs. These investigations must wait until all stories from all participants are rated. The Conclusion section will further address how future research that includes data from other “types” of people (lawyers, small community members, etc) will make it possible to measure whether there are predictable differences between- and/or within-groups. Only with more data can a determination be made as to how much variance is actually due to the story and how much to the characteristics of the individual.

For now, it is assumed that that it is the strong situation of the prompt and not the characteristics of the participant (trait) that is driving the system of correlations for the

following reasons: the prompt will trigger some activation of a psychological mechanism in the mind that deals with each identified dimension of the story (for example, injustice or unfairness), and that these mechanisms, or “domain-specific modules”, will be triggered by each other in the theoretically-specified sequence (for example, having to think about and explain *why* a thing is unfair or unjust might trigger considerations of either equality or equity, and so on). Therefore, due to the results of this study there is some support that causal relationships do exist within the brain (as by cross-talk between neural circuits) and are integral to the exploratory nature of the study to create future SEMs.

## CONCLUSION

This study was originally designed to test the hypothesis that fairness and justice are related, yet separate terms and to understand the concepts that help make them up. The results of this study allow us support the original hypothesis with the additional knowledge that: 1) for unfairness, outcome, procedure, and equality are important, 2) for injustice, outcome, procedure, and equity are important, 3) and that there are some sex differences such that females are more concerned with fairness. As with most exploratory research, this study has opened new avenues of questioning; most importantly, why are the concepts different, for which groups are they different, and what is the best way to translate this data into viable options for policy? It is imperative to propose theories that would explain these differences, state them clearly as testable hypotheses, and develop a research program devoted testing these theories and cross-validating the current research. Further studies in this area will go beyond the original predictions that have been made here, and continue to develop causal theory as to the principles behind these findings.

### Future directions

*1. Develop and refine testable hypotheses as to WHY people may think about these terms differently.* Now that there is some evidence at hand as to how people think of the terms fairness and justice, the next logical extension of this research is to examine how differences in fairness and justice developed. First, socio-legal scholars need to develop a theory as to why people conceive of these terms differently. Once the theory is

developed, researchers can generate testable hypotheses based on these results and to refine and retest the theory.

The first hypothesis I would propose to test with the interpretations of the current data is that humans are born with an innate sense of what this study would term “fairness”. Evidence is gathering from multiple sources that “fairness” as a concept developed first, as a natural outgrowth of face-to-face relationships in small village-type communities (Figueredo, 2001). People relied on their cognitive abilities to keep track of trading relationships, interactions between individual members of the society, and reciprocity between kin groups. Life in small groups, where each individual knows all the others in the society would not necessitate formalized rules of exchange, or distribution. Neither would it be necessary to bring in a neutral, third party authority figure to settle disputes, because all members of the society would have a stake in resolving the dispute in order to maintain group cohesion. In evidence of this argument are studies of children ages 6-8 from the US and South Africa, who would rather throw away excess candy bars than distribute the extras unequally – even when given the option to keep the extra candy (Novotny, 2012). This displays a concept of equality even in very young members of society. In addition, studies of capuchin monkeys (Brosnan & de Waal, 2003), showed that after completing the same tasks, capuchin monkeys who were rewarded with cucumbers would reject their reward if they saw another monkey receive a grape (a more favored reward). The authors surmised from the animal behavior that the monkeys had a sense of “work to reward” comparisons. According to the current study, these behaviors show a sense of equity – another important component to fairness.

Based on what is known about the evolution of law, another hypothesis worth testing is that “justice”, as a concept, developed later, when populations grew and trading and other interactions were no longer feasible with only people from your own village. Third parties were needed to decide disputes, and these third parties needed to be legitimate authority figures to help make decisions. Once small groups and societies began interacting with each other regularly, different concepts of “fairness” or what “ought to be” came into conflict. Concepts of what constituted proper and improper trading, distribution, and behavior became codified. Because 90% of human societies are patriarchal, it is reasonable to assume that males, as the traditional leaders (decision-makers, judges, etc) became the arbiters of the newer concept of “justice”. Males were then expected to learn the rules to maintain and apply societal norms. While males became more rigidly trained in procedures for acceptable dispute resolution, females would have been traditionally excluded from this group, and would have continued honing their relationship skills.

Kohlberg and Gilligan’s theories would fit into a hypothesis where fairness developed first and justice second. In this study, Kmdl and Gmdl were used as an entrée into the realm of sex differences among perceptions of fairness and justice. Gilligan’s well-known argument against Kmdl was that it only measures highly structured, rule-oriented development, causing females to score lower on Kmdl. Gilligan asserted that females’ inherent differences due to environmental factors and genetic preparation for caring and motherhood caused them to value relationships over rules or rule-following. It is possible that in their attempts to create viable scales, Kohlberg, Gilligan, and

subsequent moral development researchers failed to that capture the ultimate cause of sex differences, the combined role of environmental factors and innate genetic factors. It is not that females are incapable of understanding justice as a legal institution, but rather due to generations of exclusion from the legal processes and education, a justice orientation may have been ancestrally derived mostly in males. This data confirm that there are no sex differences in the level of moral reasoning a person can reach based on Kohlberg and Gilligan's moral development scales, meaning that males and females can both reach high levels of moral development. However, this data tend to show a female preoccupation with fairness and equality, and may provide evidence of the evolution from fairness to justice.

2. *Continue rating existing data.* As with any preliminary study, the first step to more robust results will be to cross-validate the findings of this study. In the case of this study, more data points are readily available from the additional unrated stories that were collected for the present paper. This data will help address some of the limitations of the current study by providing multiple observations per participant, and allowing us to separate individual differences across stories from story differences. Individuals in the two conditions of this study produced as few as one story per person, but as many as 21 stories. By coding the existing stories in this data set, future research can begin a cross-validation of the findings of this study (to be certain that this study's findings are replicated within the experimental groups) and extend the conclusions to include individual differences across stories. As stated previously, this study collected demographic information on all participants with emphasis on their racial and ethnic

backgrounds and levels of identification with those groups. This information could be extremely useful in determining whether differences between stories are permanent and detectable characteristics of the individual.

*3. Expand samples to include groups that do not confound multiple dimensions of interest.* It is essential to sample from groups other than University psychology students (in this sample, mostly female, Caucasian, upper-middle class, English speakers, etc) in order to draw conclusions that will be useful to people in other areas of society. This study provided an important first step in understanding how a student sample conceives of fairness and justice. However, to be truly useful toward creating and testing theory, it is critical to sample from groups with specific dimensions of generalizability that are chosen for their singular ability to confirm or refute the proposed theory. In order to test these hypotheses, future studies will need to sample according to the direction in which researchers want to generalize their findings.

For example, individuals from other groups may have different expectations of what it means to be fair or just. In shaping public policy, it is important for us to examine the beliefs of legal actors such as lawyers, judges, and court administrators, who receive lengthy and detailed training in justice concepts, to discover whether their conceptions of fairness and justice are in alignment with this sample, or as I believe, are more similar to the suppositions made by the socio-legal scholars referred to in this paper. If the hypothesis holds that justice developed secondary to fairness, these individuals may share a Kohlbergian “law and order” mindset that leads them to generate stories of injustice. If differences exist between the citizen without legal training and

those involved with the justice system, those differences may need to be addressed. Is the fact that these professions are still dominated by men; is this evidence of an innate component to this mindset? Would that mindset be due to the training they received to do their jobs, or did they choose their jobs based on an underlying genetic predisposition to “law and order”? This is just one example of a dimension that could be sample for that would test the theory. Group-wise comparisons will help fill in gaps in the research.

Sampling from different cultural groups would also help provide evidence to either falsify or support these hypotheses. For example, there are still small group-based societies in North America such as groups who choose not to use a formalized justice system for dispute resolution. These groups still rely on their own standards, which may or may not match those of the groups previously discussed. If the hypothesis that justice only developed to help societies deal with increasing populations or diverse groups is correct then smaller societies may not have need of a concept of justice as it is defined in this study. If this is the case, then individuals from smaller societies should write stories with less injustice content and more unfairness content than other samples. In North America, there are isolated groups such as the Comca’ac in Mexico, or the Amish in Pennsylvania whose responses to these prompts may support such a hypothesis.

However, it may also be the case that it is not the size of the community that matters, but the relative individualism or collectivism of the culture that causes differences. If this is the case, sampling from a classically defined collectivist culture, such people from Japan or Mexico, [where there is only one word to describe the combined concepts of fairness and justice (*fukohei* in Japanese, and *justo* in Spanish)]

would provide us with data to conclude whether community size outweighs collectivism's influence. Randomly sampling from the larger United States may prove useful in characterizing a robust Structural Equations Model justice and fairness, but supporting or rejecting new hypotheses can only be accomplished by sampling from strategically selected groups.

*4. Begin confirmatory analysis of the theory derived from the exploratory data collected from the current study.* One way to begin a confirmatory analyses of these findings would be to use *a priori* theory to build a Structural Equations Model of the concepts of both fairness and justice. With more participants there will then have sufficient information to begin testing a Structural Equations Model that will graphically represent the fairness and justice constructs as well as the factors that feed into them, the path coefficients that demonstrate the strength and direction of causal influence, and possibly discover more factors that can account for more of the variance in the overall model. A graphical representation of the perceptions of the proposed samples will provide easy comparisons between groups to help us in the next step.

*5. Explore the social consequences of increasing justice and/or fairness.* Most people would agree that increasing justice and fairness are worthy goals. The difficulty is the manner in which such change is achieved. In certain circumstances, increasing justice or fairness may not be entirely desirable. For example, let us examine the development and implementation of mandatory sentencing laws such as a three strikes policy. These laws are meant to serve as deterrents to patterns of criminal behavior and were put in place by lawmakers who were tired of seeing multiple offenders on the streets

due to judicial discretion in sentencing. Once in place these sentencing mandates removed a judge's ability to punish criminals differently for similar classes of offenses. A violent felony, such as murder, was given the same weight in the three strikes laws as non-violent thefts that met felony standards. The unintended consequences of these laws were that some individuals were being punished too harshly for their crimes according to societal standards. This increase in the equality aspect of justice, without a balancing increase in the equity aspect led to societal outrage, and eventual revocation or revision of three-strikes type policies (Scheck, 2010).

It is customary for authors in the field of public policy and the law to use their research findings to make policy recommendations that they believe will assist practitioners in the area to improve their field. Due to the extremely broad scope of the content of this study, in addition to its very preliminary findings, the current results are not definitive enough for a policy recommendation. First, it would be difficult to say with any degree of certainty that these findings are true for the entire US population. Second, without any confirmatory evidence that the path models are correct, the model may still be wrong with regard to the relative weights of the path analyses. Third, without other samples to draw from, addressing any recommendations to society at large would be premature. At this point, policy recommendations would be so vague and broadly stroked as to render them meaningless.

The current study has shown that both equity and equality are important components of justice. If justice is equal treatment across the board, and fairness includes individual consideration, increasing one may come at the cost of the other.

Getting the balance wrong could create as many problems for the justice system as they are intended to solve. This research will help provide a clearer understanding of the interactions of the factors that make up fairness and justice, and may help lawmakers avoid such unintended consequences as the backlash from mandatory sentences.

APPENDIX A  
TABLES AND FIGURES

Table 1

*Predicted Correlations among Variables*

	Fairness	Justice	Procedure	Outcome	Equity	Equality	Gmdl	Kmdl	Prompt	Sex
Fairness	1									
Justice	+ (A) - (SS)	1								
Procedure	+ (A) + (SS)	+ (A) - (SS)	1							
Outcome	+ (A) - (SS)	+ (A) + (SS)	- (A) - (SS)	1						
Equity	+	+	0	+	1					
Equality	+	+	0	+	-	1				
Gmdl	0 (K) + (G)	0 (K) 0 (G)	+	-	-	+	1			
Kmdl	0 (K) 0 (G)	0 (K) + (G)	-	+	+	-	-	1		
Prompt	0 (A) - (SS)	+ (A) + (SS)	0 (A) - (SS)	0 (A) + (SS)	0	0	0 (K) - (G)	0 (K) + (G)	1	
Sex	0 (A) 0 (SS)	+ (A) 0 (SS)	0 (A) 0 (SS)	0 (A) 0 (SS)	+	-	0 (K) - (G)	0 (K) + (G)	0	1

*Note:* For Prompt: Fairness Prompt = 0, Justice Prompt = 1;  
For Sex: Female Respondents = 0, Male Respondents = 1

Table 2

*Rater Training Reliabilities*

	Intra-Rater	Story	Rater	Story/Rater	Inter Time 1	Inter Time 2
Injustice	.954	F(39,156)=17.16, p<.0001	F(4,156)=2.76, p=.030	F(156,200)=2.44, p<.0001	.910	.815
Unfairness	.848	F(39,156)=4.56, p<.0001	F(4,156)=12.98, p<.0001	F(156,200)=1.32, p=.0312	.645	.581
Procedure	.850	F(39,156)=3.90, p<.0001	F(4,156)=6.29, p=.0001	F(156,200)=1.57, p=.0013	.645	.540
Outcome	.847	F(39,156)=4.68, p<.0001	F(4,156)=8.95, p<.0001	F(156,200)=1.35, p=.0224	.649	.664
Equity	.907	F(39,156)=6.62, p<.0001	F(4,156)=5.41, p=.0004	F(156,200)=2.12, p<.0001	.731	.705
Equality	.896	F(39,156)=7.91, p<.0001	F(4,156)=5.01, p=.0008	F(156,200)=1.68, p=.0003	.759	.713
Gmdl	.837	F(39,156)=4.39, p<.0001	<b>F(4,156)=.96, p=.4286</b>	F(156,200)=1.43, p=.0091	.606	.674
Kmdl	.818	F(39,156)=4.10, p<.0001	<b>F(4,156)=.67, p=.6160</b>	<b>F(156,200)=1.27, p=.0553</b>	.580	.653

*Note:* Average Intra-rater reliability = .870

Average Inter-Rater reliability: Time 1 = .691

Average Inter-Rater reliability: Time 2 = .668

Table 3

*Analysis Reliabilities*

	Intra-Rater	Story	Rater	Story/Rater	Inter Time 1	Inter Time 2
Injustice	.992	F(156, 200)=6.23, p<.0001	F(3,200)=8.05, p<.0001	F(200,60)=3.23, p<.0001	.896	.914
Unfairness	.977	F(156, 200)=2.07, p<.0001	F(3,200)=13.84, p<.0001	F(200,60)=2.30, p<.0001	.764	.698
Procedure	.953	F(156, 200)=1.85, p<.0001	F(3,200)=21.87, p<.0001	<b>F(200,60)=1.08, p=.3711</b>	.742	.577
Outcome	.966	F(156, 200)=1.94, p<.0001	F(3,200)=31.79, p<.0001	<b>F(200,60)=1.41, p=.0602</b>	.715	.687
Equity	.974	F(156, 200)=2.37, p<.0001	F(3,200)=12.32, p<.0001	F(200,60)=1.84, p=.0033	.783	.658
Equality	.975	F(156, 200)=2.61, p<.0001	F(3,200)=9.00, p<.0001	F(200,60)=1.83, p=.0035	.805	.713
Gmdl	.965	F(156, 200)=2.15, p<.0001	F(3,200)=13.85, p<.0001	<b>F(200,60)=1.40, p=.0632</b>	.769	.619
Kmdl	.954	F(156, 200)=2.78, p<.0001	F(3,200)=10.29, p<.0001	<b>F(200,60)=.91, p=.6842</b>	.807	.743

*Note:* Average Intra-rater reliability = .970  
Average Inter-Rater reliability: Time 1 = .785  
Average Inter-Rater reliability: Time 2 = .701

Table 4

*Correlation Matrix of Criterion Variables*

	INJUSTICE	UNFAIRNESS	PROCEDURE	OUTCOME	EQUITY	EQUALITY	Gmdl	Kmdl
INJUSTICE	--	.253*	.318*	.280*	.360*	-.079	.098	.204*
UNFAIRNESS		--	.340*	.550*	.359*	.292*	.123*	.124*
PROCEDURE			--	.006	.241*	.135*	.145*	.162*
OUTCOME				--	.335*	.105	.075	.046
EQUITY					--	.145*	.032	.078
EQUALITY						--	.464*	.391*
Gmdl							--	.752*
Kmdl								--

*Note:* \* indicates a Pearson correlation coefficient that is significant at the 0.05 level (2-tailed).

Table 5

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable INJUSTICE*

Predictor	INJUSTICE	
	<i>F</i> (1, 154)	<i>p</i>
<b>TYPE</b>	<b>47.15</b>	<b>&lt;.0001</b>
SEX	1.34	.2494
TYPE*SEX	1.92	.1684

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 6

*Hierarchical Multiple Regression Analysis for Predictors the Criterion Variable UNFAIRNESS*

Predictor	UNFAIRNESS	
	<i>F</i> (1, 152)	<i>p</i>
<b>INJUSTICE</b>	<b>11.13</b>	<b>.0011</b>
TYPE	1.65	.2015
<b>SEX</b>	<b>6.94</b>	<b>.0093</b>
INJUSTICE*SEX	2.14	.1454
TYPE*SEX	0.00	.9481

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 7

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable  
PROCEDURE*

Predictor	PROCEDURE	
	<i>F</i> (1, 150)	<i>p</i>
<b>UNFAIRNESS</b>	<b>21.32</b>	<b>&lt;.0001</b>
<b>INJUSTICE</b>	<b>10.60</b>	<b>.0014</b>
TYPE	2.17	.1427
SEX	.08	.7740
UNFAIRNESS*SEX	.45	.5048
INJUSTICE*SEX	.15	.7006
TYPE*SEX	.15	.7003

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 8

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable  
OUTCOME*

Predictor	OUTCOME	
	<i>F</i> (1, 148)	<i>p</i>
PROCEDURE	.01	.9295
<b>UNFAIRNESS</b>	<b>82.04</b>	<b>&lt;.0001</b>
<b>INJUSTICE</b>	<b>9.78</b>	<b>.0021</b>
TYPE	1.03	.3114
SEX	.08	.7801
PROCEDURE*SEX	.06	.8045
UNFAIRNESS*SEX	.17	.6807
INJUSTICE*SEX	.01	.9028
TYPE*SEX	.08	.7799

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 9

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable EQUITY*

Predictor	EQUITY	
	<i>F</i> (1, 146)	<i>p</i>
<b>OUTCOME</b>	<b>22.73</b>	<b>&lt;.0001</b>
<b>PROCEDURE</b>	<b>11.57</b>	<b>.0009</b>
UNFAIRNESS	3.08	.0815
<b>INJUSTICE</b>	<b>9.03</b>	<b>.0031</b>
TYPE	2.16	.1441
<b>SEX</b>	<b>4.58</b>	<b>.0340</b>
OUTCOME*SEX	.56	.4540
PROCEDURE*SEX	.044	.5082
UNFAIRNESS*SEX	.12	.7245
INJUSTICE*SEX	.28	.5969
TYPE*SEX	1.46	.2291

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 10

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable EQUALITY*

Predictor	EQUALITY	
	<i>F</i> (1, 144)	<i>p</i>
EQUITY	3.60	.0599
OUTCOME	.61	.4364
PROCEDURE	2.01	.1580
<b>UNFAIRNESS</b>	<b>9.64</b>	<b>.0023</b>
<b>INJUSTICE</b>	<b>5.62</b>	<b>.0191</b>
TYPE	.02	.8790
SEX	.01	.9114
EQUITY*SEX	1.39	.2410
OUTCOME*SEX	.60	.4410
PROCEDURE*SEX	2.03	.1566
UNFAIRNESS*SEX	.12	.7289
INJUSTICE*SEX	.16	.6872
TYPE*SEX	.00	.9963

**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 11

*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable Gmdl*

Predictor	<i>Gmdl</i>	
	<i>F</i> (1, 142)	<i>p</i>
<b>EQUALITY</b>	<b>43.51</b>	<b>&lt;.0001</b>
EQUITY	.26	.6116
OUTCOME	.33	.5668
PROCEDURE	1.92	.1682
UNFAIRNESS	.80	.3714
INJUSTICE	3.63	.0589
TYPE	.65	.4203
SEX	.85	.3594
EQUALITY*SEX	1.25	.2648
<b>EQUITY*SEX</b>	<b>5.85</b>	<b>.0168</b>
OUTCOME*SEX	.42	.5156
PROCEDURE*SEX	.10	.7505
UNFAIRNESS*SEX	.00	.9441
INJUSTICE*SEX	.18	.6705
TYPE*SEX	.31	.5782

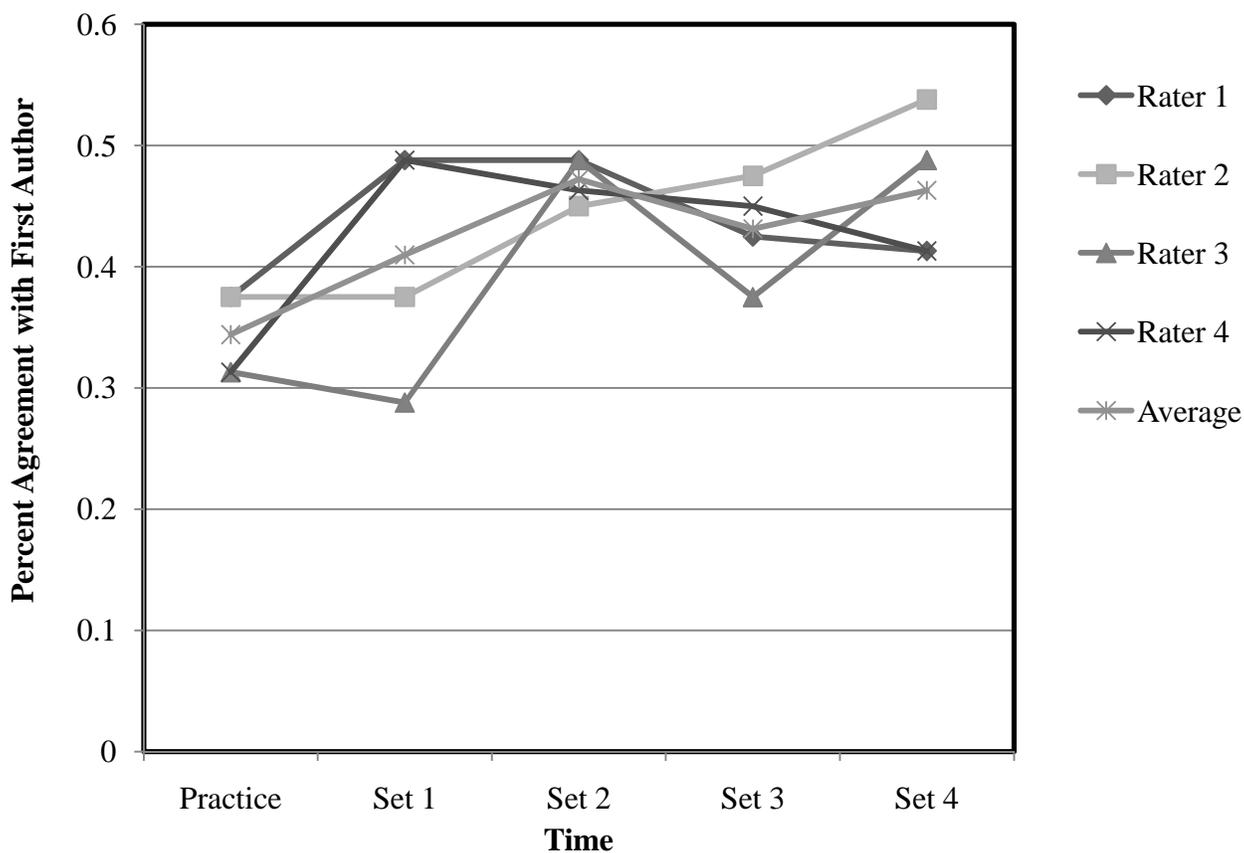
**Note:** Predictor variables that are significant in the model are highlighted in bold

Table 12

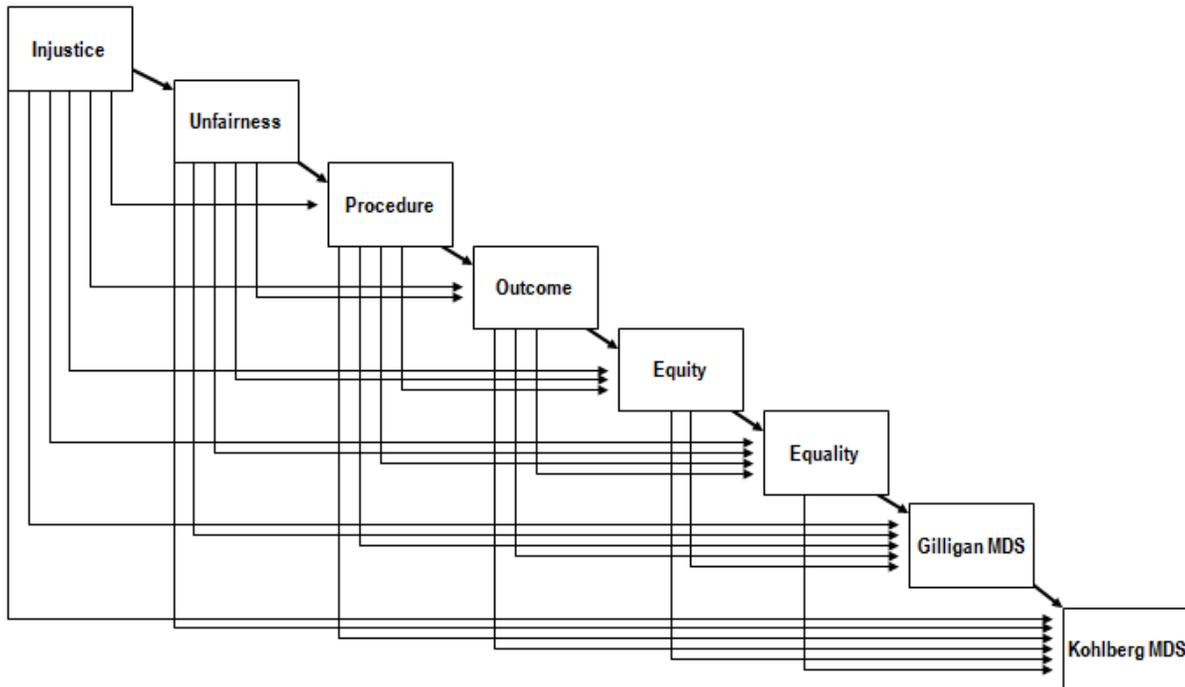
*Hierarchical Multiple Regression Analysis for Predictors of the Criterion Variable Kmdl*

Predictor	<i>Kmdl</i>	
	<i>F</i> (1, 142)	<i>p</i>
<b>Gmdl</b>	<b>88.16</b>	<b>&lt;.0001</b>
EQUALITY	.35	.3592
EQUITY	.35	.3575
OUTCOME	.39	.5352
PROCEDURE	.52	.4705
UNFAIRNESS	.04	.8461
<b>INJUSTICE</b>	<b>6.86</b>	<b>.0098</b>
TYPE	.63	.4288
SEX	.11	.7367
G*SEX	.17	.6764
EQUALITY*SEX	2.22	.1382
EQUITY*SEX	2.53	.1141
OUTCOME*SEX	.04	.8327
PROCEDURE*SEX	.66	.4170
UNFAIRNESS*SEX	.91	.3416
INJUSTICE*SEX	2.38	.1255
<b>TYPE*SEX</b>	<b>6.74</b>	<b>.0104</b>

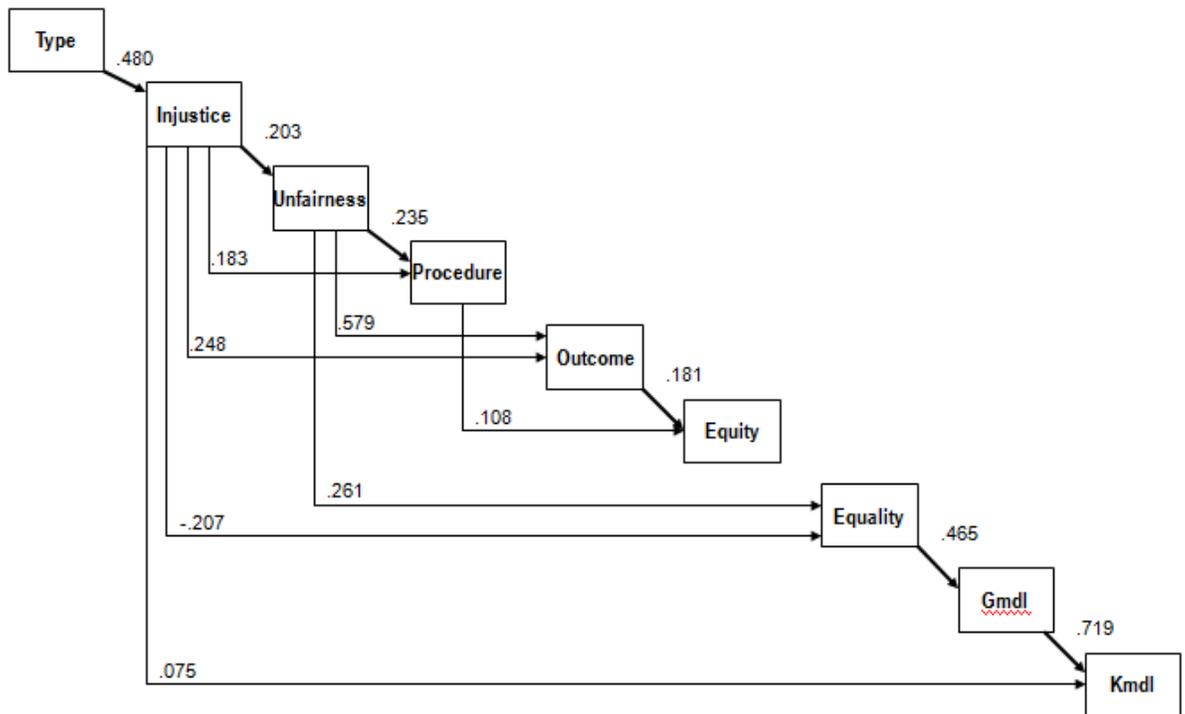
**Note:** Predictor variables that are significant in the model are highlighted in bold



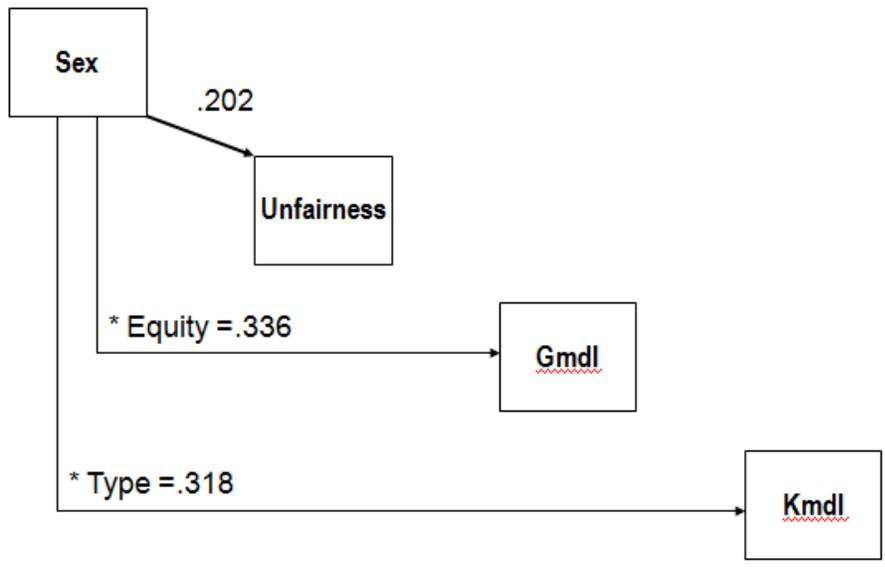
**Figure 1.** Graphical representation of the progression of rater training over time. Each rater's scores on the eight dimensions for each story were matched to the author's in an all-or-nothing comparison. Each match was coded as a 1 and each mismatch a 0. The all-or-nothing average was then computed to see how well each coder's training evolved over the training period. This all-or-nothing approach produced a highly rigorous standard for the raters and showed that over time, they became more consistent with the author's ratings.



**Figure 2.** Hypothesized Cascade Model. Cascade model of criterion variables being tested by the theoretically based hypothesized relationships. Predictor variables (not included in this figure) included the type (TYPE) of story (unfairness/injustice) the participant was asked to write about, and the sex (SEX) of the participant.



**Figure 3.** Statistically significant pathways not including sex. The results of the cascade model include both direct and indirect effects. Type of story (unfairness/injustice) directly influences ratings of Injustice, Unfairness, and Procedure. Outcome has a direct effect on Equity, and Equality has a direct effect on both Gmdl and Kmdl.



**Figure 4.** Statistically significant effects of sex. The results of the cascade model show a direct effect of sex on Unfairness, such that females produced more stories of unfairness regardless of prompt. Interaction effects were found with sex by equity for Gmdl (showing that stories high on violations of Equity also led to high ratings on Gmdl, but only for women) and sex by type for Kmdl (showing that females who received the “injustice” prompt also tended to rate highly on Kmdl) .

## APPENDIX B ORIGINAL CODING BOOKLET

The first six codes will consist of ordinal scales ranging from 0-4. For the first six scales, you can conceptualize the 0-4 codes as similar to a Likert-type scale where:

- 0 = this characteristic or descriptor is NOT FOUND in the story
- 1 = this characteristic or descriptor IS PRESENT in the story, BUT NOT SIGNIFICANTLY
- 2 = this characteristic or descriptor IS PRESENT IN A SIGNIFICANT way in the story
- 3 = this characteristic or descriptor IS AN ESSENTIAL component of the story
- 4 = this characteristic or descriptor COMPLETELY CAPTURES content of the story

### **Dimension 1: INJUSTICE**

Definition: Any and all situations that require the use of formalized institutions, such as laws, legal professionals, and law enforcement, in orders to create a sense of balance and reciprocity among members of a given society.

- A rating of 0 on this dimension would indicate that there is no mention of the law, law enforcement, legal professionals, or courts in the story.
- A rating of 4 on this dimension would indicate that the main focus of the story is on some violation of a law that required the intervention of a legal professional or institution in order to restore balance to the situation or to otherwise rectify the situation.
- Rating 4 Example: The target is sued by another person, either correctly or incorrectly and must go through the court system to resolve the issue.

### **Dimension 2: UNFAIRNESS**

Definition: Any and all situations that include injustice as described above, or describe violations of equality, equity, or the sense that a procedure is right, proper, or correct, or that an outcome is right, proper, or correct.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice, equality, equity, and that all procedures and outcomes were right, proper, and correct.
- A rating of 4 on this dimension would indicate that the main focus of the story is a violation of justice as described above, a violation of equality, equity, or the feeling that some procedure or outcome is wrong, improper, or incorrect.
- Rating 4 Example: The writer describes a situation where a young child dies.

### **Dimension 3: PROCEDURE**

**Definition:** Any and all situations that discusses the who, what, how, and why of a decision or process that leads to unfairness or injustice as described above.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice or fairness dependent upon how a decision was reached, by whom, or under what circumstances the decision or process took place.
- A rating of 4 on this dimension would indicate that the main focus of the story is a violation of the writer's sense of justice or fairness dependent upon how a decision was reached, by whom, or under what circumstances the decision or process took place.
- Rating 4 Example: The writer describes how their parents made an arbitrary decision about how to punish the target without listening to her side of the story.

#### **Dimension 4: OUTCOME**

**Definition:** Any and all situations that describe unfair or unjust gains, losses, punishments, rewards, or conclusions to a dispute.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice or fairness due to gains or losses.
- A rating of 4 on this dimension would indicate that the main focus of the story is a gain or loss that the writer feels is a violation of their sense of fairness or justice.
- Rating 4 Example: The writer describes how another person received a grade of A on an assignment that the person cheated on.

#### **Dimension 5: EQUITY**

**Definition:** Any and all situations that describe unfair or unjust proportions of work to reward or inputs to outputs.

- A rating of 0 on this dimension would indicate there is no violation of the writer feels that the amount of reward is equal to the amount of effort, or that the punishment fits the crime.
- A rating of 4 on this dimension would indicate that the main focus of the story is a sense of unfairness or injustice due to unequal rewards to effort or that a punishment does not fit the crime.
- Rating 4 Example: The writer describes a murderer who does not receive any jail time for their crime.

#### **Dimension 6: EQUALITY**

**Definition:** Any and all situations that describe unfair or unjust treatment of people (individuals or groups) when things should be equal

- A rating of 0 on this dimension would indicate that the writer feels that everyone is being treated fairly, and no one is being given any advantage over anyone else.
- A rating of 4 on this dimension would indicate that the main focus of the story is the unfair or unjust treatment of a group or an individual based on factors that should not be taken into account.
- Rating 4 Example: The writer describes oppression of an ethnic minority group by a government.

**Please note that Dimensions 7 & 8 have their own rating scales that are different from those above. Please follow the subsequent guidelines for Dimensions 7 & 8.**

### **Dimension 7: K**

Score 1 if the writer believes people should behave in a manner to avoid punishment, people should defer to authority without question.

Score 2 if the writer believes that people should behave in a manner to gain reward, people are trying to satisfy their own needs.

In 1 & 2, the writer expresses the opinion that right and wrong are defined in terms of the stimulus-response conditioning that is the basis of his or her culture's values, morals, and laws.

Score 3 if the writer believes that people should behave in order to gain approval and avoid disapproval. People should conform to stereotypical roles.

Score 4 if the writer believes that people should behave in order to do one's duty and avoid guilt. Conforming to roles benefits greater society.

In 3 & 4, the writer uses basic rules help maintain good relationships between people. In this stage, the individual accepts these rules, even if they go against their best-interests because he or she is able to reason that these rules lead to order in the society.

Score 5 if the writer abstractly considers that there are agreed upon rights in every society which supersede laws. Focus is on the culture's definitions of personal rights.

Score 6 if the writer has developed their own personal moral standards which could be applied universally.

The writer's ability to reason abstractly allows them to think in terms of a morality that would apply to anyone, anywhere. A person is able to understand that there is no absolute right or wrong when one considers all the possible factors involved in a moral dilemma, then all moral systems are relative, and a personal moral standard is as good as any other

### **Dimension 8: G**

Score 1 if the writer's main concern is that someone is trying to ensure their own survival.

Score 2 if the writer's main concern is that by trying to care for him- or herself, the individual is being selfish.

Score 3 if the writer seems to believe that good behavior means caring for others, frequently equated with self-sacrifice. With 3, the writer shows a sense of responsibility to others. Here, the writer sees self sacrifice is equivalent to goodness and selflessness is an admirable trait.

Score 4 if the writer's main concern is the balance between caring for the self and others. Conflict arises as the target searches for equilibrium.

Score 5 if the writer's focus is on the dynamics of relationships. The target takes pains to eliminate the tension between self and others.

Score 6 if care is extended beyond personal relationships to a general recognition of the interdependence of self and all others, accompanied by a universal condemnation of exploitation and hurt.

In 5 & 6, the writer realizes that he or she is worth the same consideration he or she would give any other person. Now, the focus is no longer sacrifice, but is based on the principle of nonviolence: that one does not hurt others or the self (meaning that it is acceptable to be a little selfish again) which includes not harming feelings or relationships between other persons.

APPENDIX C  
CODING BOOKLET – FINAL REVISION

**0** = this characteristic or descriptor is **NOT FOUND** in the story

**1** = this characteristic or descriptor **IS PRESENT** in the story, **BUT NOT SIGNIFICANT**

**2** = this characteristic or descriptor **IS PRESENT IN A SIGNIFICANT** way in the story

**3** = this characteristic or descriptor **IS AN ESSENTIAL** component of the story

**4** = this characteristic or descriptor **COMPLETELY CAPTURES** content of the story

**Dimension 1: INJUSTICE**

Definition: Any and all situations that require the use of formalized institutions, such as laws, legal professionals, government agencies, and law enforcement, in order to create a sense of balance and reciprocity among members of a given society.

- A rating of 0 on this dimension would indicate that there is no mention of the law, law enforcement, legal professionals, government agencies, or courts in the story.
- A rating of 4 on this dimension would indicate that the main focus of the story is on some violation of a law that required the intervention of a legal professional or institution in order to restore balance to the situation or to otherwise rectify the situation.
- Rating 4 Examples:
- The target is sued by another person, either correctly or incorrectly and must go through the court system to resolve the issue.
- An innocent man is sent to jail
- Someone is pulled over by the police for no apparent reason

**Dimension 2: UNFAIRNESS**

Definition: Any and all situations that include injustice as described above, or describe violations of equality, equity. Also includes the sense that a procedure or that an outcome is wrong, improper, or incorrect for some reason. Distinction: This category is **NOT** about the severity of or how unfair something is, but *whether* there is a violation of the author's sense of right and wrong.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice, equality, equity, and that all procedures and outcomes were right, proper, and correct.
- A rating of 4 on this dimension would indicate that the main focus of the story is a violation of justice as described above, a violation of equality, equity, or the feeling that some procedure or outcome is wrong, improper, or incorrect.
- Rating 4 Example:

- The writer describes a situation where a young child dies.
- Someone who gets into a school, or a club, a job because of an advantage such as being rich or pretty

### **Dimension 3: PROCEDURE**

Definition: Any and all situations that discuss the how and why of a decision or process.

Distinction: In this category, dealing with arbitrary rules or discriminatory treatment may also be a result, but here the focus is on the *process*. When the focus is on the *outcome*, dimension 4 should receive the higher score.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice or fairness dependent upon HOW a decision was reached, by whom, or under what circumstances the decision or process took place.
- A rating of 4 on this dimension would indicate that the main focus of the story is a violation of the writer's sense of justice or fairness dependent upon HOW a decision was reached, by whom, or under what circumstances the decision or process took place.
- Rating 4 Example:
- The target is punished by parents who did not listen to her side of the story.
- The target was fired based on the accusations of another, and no one investigated the claims.
- The rule or law itself is the problem, not the particular outcome.

### **Dimension 4: OUTCOME**

Definition: Any and all situations that describe unfair or unjust gains, losses, punishments, rewards, or *conclusions* to a dispute or life event. Distinction: Here, the main element of the story has to do with the way things turned out – not the process.

- A rating of 0 on this dimension would indicate there is no violation of the writer's sense of justice or fairness due to gains or losses.
- A rating of 4 on this dimension would indicate that the main focus of the story is a gain or loss that the writer feels is a violation of their sense of fairness or justice.
- Rating 4 Example:
- The writer describes how another person received a grade of A on an assignment that the person cheated on.
- A drunk driver killing an innocent
- Someone is kicked out of a dorm for drinking one beer

### **Dimension 5: EQUITY**

Definition: Any and all situations that describe unfair or unjust proportions of work to reward or inputs to outputs. Violations could include situations where “No Work is Rewarded”, “Hard Work is Not Rewarded”, “The Reward is Disproportionate to the Work”, “Wrongful Behavior is Punished Disproportionately to the Act and Intent”, “Wrongful Behavior is Not Punished”, “Wrongful Behavior Rewarded”, “Good Behavior is Not Rewarded”, or “Good Behavior is (Innocence) Punished”. Violations occur due to the author’s internal sense of proportion.

- A rating of 0 on this dimension would indicate there is no violation of the writer feels that the amount of reward is equal to the amount of effort, or that the punishment fits the crime.
- A rating of 4 on this dimension would indicate that the main focus of the story is a sense of unfairness or injustice due to unequal rewards to effort or that a punishment does not fit the crime.
- Rating 4 Example:
- The writer describes a murderer who does not receive any jail time for their crime.
- Being good gets no attention, or seems to bring no immediate benefit

#### **Dimension 6: EQUALITY**

Definition: Any and all situations that describe unfair or unjust treatment of people (individuals or groups) when *things should be equal*. Here, the outcome, or treatment has nothing to do with a person’s work or non-work behavior, but with some other factor that should be irrelevant. Discrimination stories would generally fall into this area.

- A rating of 0 on this dimension would indicate that the writer feels that everyone is being treated fairly, and no one is being given any advantage over anyone else.
- A rating of 4 on this dimension would indicate that the main focus of the story is the unfair or unjust treatment of a group or an individual based on factors that should not be taken into account.
- Rating 4 Example:
- The writer describes oppression of an ethnic minority group by a government.
- More attractive people are treated better or flattered

**Please note that Dimensions 7 & 8 have their own rating scales. Please follow the subsequent guidelines for Dimensions 7 & 8. These scales show a progression of thought from less to more sophisticated. It is best to begin with 1 and work your way through the list sequentially until you find the “most appropriate” level of thought or consideration.**

#### **Dimension 7: K**

In 1 & 2, the writer defines right and wrong in terms of the stimulus-response conditioning that is the basis of his or her culture's values, morals, and laws.

Score 1 if the writer believes people should behave in a manner to avoid punishment, or that people should defer to authority without question. (Obedience)

Score 2 if the writer believes that people should behave in a manner to gain reward, or approval. Here, people are trying to satisfy their own needs and wants. (Self-Interest)

In 3 & 4, the writer uses basic rules to help maintain good relationships between people. In this stage, the individual accepts these rules, even if they go against his or her best-interests, because he or she is able to reason that these rules lead to order in the society.

Score 3 if the writer believes that people should behave in a way that is expected of them. People should conform to stereotypical roles. (Conformity)

Score 4 if the writer believes that people should behave according to one's duty and to avoid guilt. The author believes that conforming to roles benefits greater society. (Law and Order)

In 5 & 6, the writer's ability to reason abstractly allows them to think in terms of a morality that would apply to anyone, anywhere. A person is able to understand that there is no absolute right or wrong when one considers all the possible factors involved in a moral dilemma, then all moral systems are relative, and a personal moral standard is as good as any other

Score 5 if the writer abstractly considers that there are agreed upon rights in every society which supersede laws. Focus is on the culture's definitions of personal rights. (Human rights)

Score 6 if the writer has developed their own personal moral standards which could be applied universally. This would be the most sophisticated level of reasoning about an issue or violation. (Universal human ethics)

**EXAMPLE:**

**The writer describes a situation in which the target's children are hungry. The target cannot borrow any more money and, so decides to steal in order to get what his/her children need.**

**Score 1) The target should not steal because it is against the law and he/she could get into trouble. Or, the writer is sorry that the target is in this position, but offers no way to resolve the situation.**

**Score 2) The target should steal the food because it will save his family. Or the target should not steal the food because then he/she will go to jail and will hurt his/her family more. (Short-term)**

**Score 3) The target should steal the food because it is his/her job to provide for his/her family. He or she is out of legal options and must therefore look to his/her higher purpose as a parent than “good citizen”.**

**Score 4) The target must resist the urge to steal because if he/she did, then anyone could, which would lead to a breakdown in society. Or if he/she does steal, the target should take their punishment, because it will preserve the role of order in the society.**

**Score 5) The target should steal because everyone has a right to life and happiness, regardless of the law. (Broad perspective, not just short-term benefit)**

**Score 6) The target should never steal because it takes away from the society, regardless of whether the act is illegal or not. Or, each life is equally significant. As humans we should preserve life. Writer may suggest a government program or weigh a human life against the cost of whatever was stolen as justification.**

#### **Dimension 8: G**

In 1 & 2, one learns only to care for one’s self. No one else is taken into consideration. (Self-Oriented) (Benefits to me-Costs to others)

Score 1 if the writer’s main concern is that someone is trying to ensure their own benefit. The writer does not describe *why* this is wrong, merely that it is. (Benefit yourself with no consideration for others)

Score 2 if the writer’s main concern is that by trying to benefit for him- or herself, an individual in the story is being selfish. To be selfish is deemed wrong or bad. (Benefitting the self is selfish. MUST acknowledge “selfishness”, “greedy”, “self-centered”)

In 3 & 4, one internalizes norms about helping or benefitting for others and tends to neglect one’s own needs or wants. Here, the writer sees self sacrifice as equivalent to goodness and selflessness is an admirable trait. (Other-Oriented) (Benefits to Others-Costs to Self)

Score 3 if the writer seems to believe that good behavior means caring for others, frequently equated with self-sacrifice. With 3, the target’s actions show a sense of responsibility to others. The author feels that people *should* sacrifice for others. (All benefit to others, all costs to self. “Pure Altruism”)

Score 4 if the writer acknowledges that there is a conflict because the target wants to do for others, but realizes the cost may be too high. The target searches for equilibrium; focus is on the balance between caring for the self and others. (Benefit to others, but costs now acknowledged)

In 5 & 6, the writer realizes that he or she is worth the same consideration as they would give to any other person. Here, the target learns to balance caring for self with caring for

others. The focus is no longer sacrifice, but that one does not hurt others OR the self (meaning that it is acceptable to be a little selfish again) which includes not harming feelings or relationships between other persons. (Universal-Oriented) (All Benefits to All, All Costs to All)

Score 5 if the writer's focus is on balancing the needs of others with the needs of the self. The target takes pains to eliminate the tension between self and others. (Balance of cost and benefit ratio for the individual. "Mutual Altruism")

Score 6 if care is extended beyond personal relationships to a general recognition of the interdependence of self and all others, accompanied by a universal condemnation of exploitation and hurt. This would be the most sophisticated level of reasoning about an issue or violation. (Balance of cost and benefit ratio for humanity)

### **EXAMPLE**

**A student describes a situation in which she studies hard, and a friend does not. He gets a higher grade in the class than she does.**

**Score 1 if the student writes that the friend has a history of cheating.**

**Score 2 if the student writes that the friend was selfish in cheating because it ruins things for hardworking students like herself.**

**Score 3 if the author feels that the friend should have included her in cheating so that she could also benefit and get a good grade. Or, he should not have cheated and sacrificed his time like she did to get a good grade.**

**Score 4 if the author acknowledges that he had good reasons for cheating, but she feels conflicted because he still should have sacrificed the time and studied.**

**Score 5 if she feels his actions have damaged their relationship as friends because he benefitted at her expense.**

**Score 6 if she feels that his cheating hurt the whole class and himself because he did not learn the material.**

## REFERENCES

- Alvarado, B.I., Sales, B.D., & Figueredo, A.J. (2006). *A quantitative analysis of Hispanic and Caucasian perceptions of unfairness*. Unpublished Master's thesis, Department of Psychology, The University of Arizona, Tucson, Arizona.
- Auerbach, J. S. (1983). *Justice Without Law?* New York: Oxford University Press.
- Brosnan, S. F., & De Waal, F. B. (January 01, 2003). Monkeys reject unequal pay. *Nature*, 425, 6955, 297-9.
- Emery, R. E. (2012) *Renegotiating Family Relationships* (2<sup>nd</sup> ed.). New York: Guilford Press.
- English, P.W. & Sales, B.D. (2005). *More than the Law: Behavioral and Social Facts in Legal Decision Making*. Washington, D.C.: American Psychological Association.
- Figueredo, A.J. (2001). Blame, retribution and deterrence among both perpetrators and survivors of male violence against women. *Virginia Journal of Social Policy & the Law*, 8(1), 219-251.
- Figueredo, A.J., & Gorsuch, R.I. (2007). Assortative mating in the jewel wasp: 2. Sequential canonical analysis as an exploratory form of path analysis. *Journal of the Arizona-Nevada Academy of Science*, 39(2), 59-64.
- Figueredo, A.J., & Sage, R.M.S. (2007). Assortative mating in the jewel wasp: 1. Female matching of eye-color genotype, not host-feeding phenotype. *Journal of the Arizona-Nevada Academy of Science*, 39(2), 51-57.
- Figueredo, A.J., Tal, I.R., McNeill, P., & Guillén, A. (2004). Farmers, herders, and fishers: The ecology of revenge. *Evolution and Human Behavior*, 25(5), 336-353.

- Finkel, N. J. (2001). *Not Fair! The Typology of Commonsense Justice*. Washington, DC: American Psychological Association.
- Finkel, N. J. (1995). *Commonsense Justice: Jurors' Notions of the Law*. Cambridge, MA: Harvard University Press.
- Fleischacker, S. (2004). *A Short History of Distributive Justice*. Cambridge, MA: Harvard University Press.
- Gilligan, C. (1982). *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Gilligan, C., & Attanucci, J. (1988). Two moral orientations. In C. Gilligan, J. Ward, J. Taylor, & B. Bardige (Eds.), *Mapping the Moral Domain: A Contribution of Women's Thinking to Psychological Theory and Education* (pp. 73-86). Cambridge, MA: Harvard University Press.
- Hegtvædt, K., & Cook, K. (2001). Distributive justice: Recent theoretical developments And applications. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of Justice Research in Law*. Boston: Kluwer Academic/Plenum Publishers.
- Kohlberg, L. (1981). *Philosophy of Moral Development*. San Francisco: Harper & Row.
- Kohlberg, L., Levine, C., & Hwer, A. (1983). Moral Stages: A Current Formulation and a Response to Critics. In J.A. Meacham (Ed.), *Contributions to Human Development (Vol. 10)*. New York: Karger.
- Kitzmann, K. M., & Emery, R.E. (1993). Procedural justice and parents' satisfaction in a field study of child custody dispute resolution. *Law and Human Behavior*, 17(5), 553-567.

- Klotkin, J. & Friedman, D. (1993, February). The Los Angeles riots: Causes, myths and solutions. Retrieved from Democratic Leadership Council:  
[www.dlc.org/documents/LA\\_riots.pdf](http://www.dlc.org/documents/LA_riots.pdf).
- Lerner, M. J. (1980). *Belief in a Just World: A Fundamental Delusion*. New York: Plenum Press.
- Lind, E. A., Kurtz, S. Mustane, L., Walker, L. & Thibaut, J.W. (1980). Procedure and outcome effects on reactions to adjudicated resolution of conflicts of interests. *Journal of Personality and Social Psychology*, 39 (4), 643-653.
- Lind, E. A., & Tyler, T. R. (1988). *The Social Psychology of Procedural Justice*. New York: Plenum Press.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18(3), 302-318.
- Makkai, T., & Braithwaite, J. (1996). Procedural justice and regulatory compliance. *Law and Human Behavior*, 20(1), 83-98.
- Morrison, B., & Ahmed, E. (2006). Restorative justice and civil society: Emerging practice, theory, and evidence. *Journal of Social Issues*, 62(2), 209-215.
- Novotny, A. (2012, March). Fairness comes first in children's minds. *gradPSYCH*, p.13.
- Rawls, J. (Ed.). (2001). *Justice as Fairness: A Restatement*. Cambridge: The Belknap Press of Harvard University Press.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: The Belknap Press.
- Roemer, J. E. (1996). *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press.

- Sales, B. D., & Shuman, D. W. (2005). *Experts in Court: Reconciling Law, Science, and Professional Knowledge*. Washington, DC: American Psychological Association.
- SASInstitute Inc. (2011). *Base SAS ® 9.3 Procedures Guide: Statistical Procedures*. Cary, NC: SAS Institute Inc.
- Scheck, O. (2010, April 1). New York's 'Three Strikes Law' Ruled Unconstitutional. Retrieved from DNAinfo.com:  
<http://www.dnainfo.com/newyork/20100401/manhattan/new-yorks-three-strikes-law-ruled-unconstitutional>.
- Thibaut, J. W., & Walker, L. (1975). *Procedural Justice: A Psychological Analysis*. New York: L. Erlbaum Associates.
- Tyler, T. R. (1990). *Why people obey the law*. New Haven: Yale University Press.
- Tyler, T. R., & Lind, E. A. (2001). Procedural Justice. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of Justice Research in Law* (pp. 65-92). New York: Plenum Publishers.
- van der Toorn, Tyler, T.R., & Jost, J.T. (2010). More than fair: Outcome dependence, system justification, and the perceived legitimacy of authority figures. *Journal of Experimental Social Psychology*, 47, 127-138.
- Vidmar, N. (2001). Retribution and Revenge. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of Justice Research in Law* (pp. 31-64). New York: Plenum Publishers.
- Walker, S. (2001). *Sense and Nonsense about Crime and Drugs: A Policy Guide* (5<sup>th</sup> ed.). Belmont, CA: Wadsworth Thomson Learning.