

GENDER AND MORALITY-BASED JUDGMENTS IN AND OUT OF THE COURTROOM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Vivian Naomi Rotenstein

August 2021

© 2021 Vivian Naomi Rotenstein

GENDER AND MORALITY-BASED JUDGMENTS IN AND OUT OF THE COURTROOM

Vivian Naomi Rotenstein, Ph.D.

Cornell University 2021

In and outside of legal settings, judgments and decisions are made on a daily basis based on an individual's gender or cultural group. Legally, gender can affect jurors' perceptions of case facts, their deliberation behavior, and final verdicts. Gender is also especially relevant in sexual misconduct cases, influencing a victim's decision to report an incident or pursue legal action, or shaping jurors' and judges' perceptions of culpability and forgiveness. In a non-legal context, decisions we make in response to morally ambiguous situations can be influenced by cultural ideals. This dissertation examines the role of juror gender in the prosecution of sexual misconduct cases, and in civil jury dynamics and decisions. It also examines the role of culture in moral decision-making. The first paper of this thesis explores the phenomenon of Himpathy – or excessive sympathy afforded onto male perpetrators of sexual misconduct – in the context of sexual harassment. It finds that sexist attitudes against women are associated with heightened trust, forgiveness, and sympathy towards a male perpetrator, and lower trust and sympathy towards a female victim. The second paper explores the role of juror gender in shaping deliberation dynamics and liability and damage award decisions. Juror gender did not affect rates of interruptions or mentions of numeric information during a deliberation, or individual damage award estimates. The final paper explores the role of culture and the presence of a loved one in solving moral quandaries. It provides evidence that Chinese participants are more likely to save a set of strangers over their loved one in the face of death, but that this preference does not hold for situations involving emotional or economic losses.

BIOGRAPHICAL SKETCH

Vivian Rotenstein, a Southern California native, graduated with a Bachelor of Arts degree in Psychology from the University of Southern California in 2016. She received her Master's degree in Development Psychology at Cornell University in 2018. Within her PhD, she has concentrations in Law, Psychology, and Human Development, and Social and Personality Psychology. Her research often employs a mixed-methods approach, and explores areas which combine tenets of developmental and social psychology to inform legal issues. Developmentally, she has conducted research on children's conception and detection of lying. Within social psychology and law, she conducts research on implicit and explicit forgiveness attitudes towards formerly incarcerated individuals, and biases (by race, gender, etc.) within criminal and civil juror decision making and voir dire for capital punishment cases. Having completed an internship at the American Institutes for Research, she is especially passionate about translational research that can enact community change and impact social or legal policies.

Dedicated to Michaela Rotenstein.

ACKNOWLEDGMENTS

Research truly takes a village. Firstly, I would like to thank the Cornell Human Ecology Scholarships and Fellowship Fund for providing me with research funding. The experiment described in Chapter 2 of this dissertation was also funded by a National Science Foundation grant SES-1536238: “Quantitative Judgments in Law: Studies of Damage Award Decision Making” , to Valerie P. Hans and Valerie F. Reyna, and by Cornell Law School. Thank you to all the research participants, both virtual and in-person, who volunteered to participate in this dissertation work.

I would like to thank my incredible committee members: Stephen Ceci, and Melissa Ferguson and Valerie Hans. Thank you so much for all of your guidance, wisdom, and support over the last five years. You have been such fantastic, inspiring mentors, and consistently remind me of my passion for research. Thank you also to Florio Arguillas from the Cornell Institute for Social and Economic Research and Lynn Johnson from the Cornell Statistical Consulting Unit, for their assistance with qualitative and quantitative analyses within this dissertation.

I would like to thank past and current graduate students of the Child Witness and Cognition Lab – Kayla Burd, Caisa Royer, Amelia Hritz, and I-An (Amy) Su for being brilliant, academic inspirations, wonderful colleagues, and compassionate and supportive friends. Additionally, a hearty thank you to all the undergraduate research assistants who assisted with this dissertation work (and other projects in the Child Witness and Cognition Lab); mentoring you all over the years has been so rewarding, and your hard work and commitment is so appreciated. I must also thank my dad, Lisa, Michael, Cynthia, and Matt for their unrelenting support and care over the years. Lastly, to my mom, Michaela, whom I miss so much every day, thank you for pushing me to be my best self and always believing in me.

TABLE OF CONTENTS

Biographical Sketch	iii
Table Of Contents	vi
Chapter 1: Introduction	1
Chapter 2: Himpathy In and Out of the Courtroom: Gender Differences in Empathy Towards Defendants	3
Chapter 3: Womens On the Jury: The Role of Gender in Civil Juries	38
Chapter 4: How Much Is Your Loved One Worth? Moral Decision Making Across Cultures....	65
Appendix B	95
Appendix C	102
Appendix D	106

CHAPTER 1: INTRODUCTION

In and out of the courtroom, we make judgments and decisions on a daily basis based on an individual's gender or cultural group. Within the law, across cases in both the criminal and civil domains, gender is an especially pertinent factor in shaping jurors' perceptions of cases, deliberation behavior, and their final verdicts. Gender is also meaningful in cases of sexual misconduct: influencing a victim's decision to report their case and/or pursue legal action, how the perpetrator of the misconduct may be viewed by the public, and within the courtroom, influencing how much blame or forgiveness judges and jurors may place on the defendant and victim. Outside of the courtroom and in a non-legal context, we may frequently face morally ambiguous situations, and can rely on culturally driven norms or values to inform how to act or what to decide in such instances. This dissertation presents empirical work exploring the role of gender in and out of the courtroom, specifically in the contexts of civil jury deliberations and the prosecution of sexual misconduct cases. It also presents work examining the role culture plays in how we make moral decisions.

In the first paper of the dissertation, entitled "Himpathy In and Out of the Courtroom: Gender Differences in Empathy Towards Defendants", I explore the phenomenon of Himpathy, or exaggerated empathy towards a male perpetrator of sexual misconduct, in how sexual harassment cases are perceived. I provide evidence demonstrating that Himpathy may not be best conceptualized as a form of sexism that only arises with particular perpetrator-victim characteristics present. Rather, it is better conceived as a facet of sexism at large, irrespective of the characteristics of the perpetrator or victim.

In the second paper of the dissertation, entitled "Women on the Jury: The Role of Gender in Civil Juries", I employ a mixed-methods approach (drawing on qualitative deliberation data

and quantitative data of proposed damage awards) to examine how juror gender may influence deliberation dynamics and jurors' individual assessments of liability and damages. I provide evidence demonstrating that jurors' proposed damage awards, and the degree of variability in those awards, are influenced by the amount of guidance jurors are given in deriving damages (the experimental condition they are exposed to), but no confirmatory evidence is found for differential rates of interruptions or average damage awards proposed by a juror's gender.

Lastly, in the third and final paper of the dissertation, entitled "How Much is Your Loved One Worth? Moral Decision Making Across Cultures", I expand on previous work of cross-cultural moral decision making by evaluating moral reasoning between Chinese and American participants in hypothetical as well as more real-life scenarios, while also adding a much more personal element to the moral decision-making calculus. I provide evidence showing that in a hypothetical moral dilemma involving the life of your loved one or a set of strangers at stake (due to a runaway speeding trolley car), Chinese participants are more likely to save strangers' lives than that of their loved ones, compared to their American counterparts. This cross-cultural difference does not hold, however, for other types of moral dilemmas involving emotional or economic loss.

CHAPTER 2:

HIMPATHY IN AND OUT OF THE COURTROOM: GENDER DIFFERENCES IN EMPATHY TOWARDS DEFENDANTS

Introduction

Empathy involves the ability to understand and share in the emotional experience of another individual, which can be instigated in response to either overt or covert emotional cues (Cuff et al., 2014; Eisenberg & Strayer, 1988; Posick et al., 2014; Riess, 2017). Heightened empathic concern can serve as conduit for greater overall pro-social, altruistic, and compassionate behavior towards others (Bagozzi & Moore, 1994; Cuff et al., 2014; Hoffman, 2008; Riess, 2017). Females have consistently been shown to express higher levels of empathy and its associated components (e.g. perspective-taking, empathic concern, personal distress) compared to men, as evidenced by self and peer-reports, and longitudinal assessments (Karniol et al., 1998; Toussaint & Webb, 2005; Vicenta Mestre et al., 2009).

The Pitfalls of Empathy

Despite its benefits, there are instances in which the expression of empathy can go awry, such that its expression is skewed towards only particular sub-groups of individuals, disadvantaging and silencing other voices. For example, the concept of empathic anger suggests that given our innate desire for fairness and reciprocity, when we are faced with instances in which these values are violated (e.g. a victim was unfairly wronged by a perpetrator of a crime), instead of merely feeling empathetic distress toward the victim, our empathy can become anger or aggression. Further, our empathic anger can motivate us to right the wrong, by, for example, helping the victim and punishing the perpetrator (Hoffman, 2008; Vitaglione & Barnett, 2003). Empathic anger can also, in some instances, lead us to hold pro-defendant and anti-victim

attitudes. The philosophical concept of Himpathy elucidates contexts in which such attitudes may be held.

“Himpathy”: Its Definition and Possible Presence In & Out of the Courtroom

In 2018, public hearings regarding Brett Kavanaugh’s nomination to the Supreme Court were held, during which sexual allegations against Kavanaugh were also discussed. In the wake of these hearings, as well in light of the 2015 case *People vs. Brock Turner*, Cornell philosophy professor Kate Manne coined the term “Himpathy” in a New York Times opinion piece she authored. She also elaborates on this concept in her book *Down Girl, the Logic of Misogyny* (2017). According to Manne, the term “Himpathy” serves to describe the exaggerated and disproportionate amount of forgiveness and empathy afforded to racial-majority, heterosexual men in positions of power in the setting of transgressions. Himpathy is especially pertinent, in response to cases of sexual assault or intimate partner violence (Manne, 2017, 2018).

Additionally, Himpathy is perpetuated as the conversation surrounding such transgressions place their focus solely on the male perpetrator, such as the details of their offense and how this transgression has led to their drastic fall from grace. In doing so, such a narrative effectively erases the voice of female victims, negating the validity and severity of their experience, and instead, serves to pique sympathy towards the male transgressor (Barger, 2018; Illing, 2018; Manne, 2018; Zeilkova et al., 2018).

Manne further argues that Himpathy is perpetuated by the patriarchal culture we live in, and at a psychological level, may be driven by an in-group preference and out-group bias, a phenomenon documented robustly across multiple contexts (Illing, 2018). The concept of Himpathy suggests that men may be especially prone to supporting a male perpetrator amidst contrary evidence to his good character and reputation, while derogating and diminishing the

female victim's voice.

The Relevance of Himpathy in Legal Proceedings

Himpathy brings to light the biased fashion in which empathy and forgiveness towards sexual offenders, especially white, high socioeconomic status or otherwise privileged males, may be manifested by the lay public. However, an unanswered set of questions remain: if and how the concept of Himpathy manifests itself in pre-trial, prosecutorial decisions to press charges and bring the offender's case to trial, and within individual jurors' judgments (i.e. perceptions of the defendant's culpability, overall likeability, etc.).

The Effect of Defendant & Victim Gender on Juror Attitudes Within Criminal Cases

How might the role of a defendant and victim's gender shape juror's case attitudes and final sentencing decisions? Prior work has examined this question in the domain of capital punishment cases, sexual assault cases, and other types of violent crimes.

In terms of victim gender effects, within prior work on capital punishment cases, empirical evidence has supported the notion of a "female victim effect," such that homicides involving a female victim have the highest likelihood of seeking and ultimately imposing the death penalty (Holcomb et al., 2004; Royer et al., 2014). A similar effect was found in work by Curry, Lee, & Rodriguez (2004), such that male perpetrators who victimized females received the longest sentences (Curry et al., 2004). Additionally, in their analysis of homicide cases from the Bureau of Justice Statistics' 1996 dataset, Glaeser & Sacerdote (2003) found that vehicular homicides involving females received 60% longer sentences compared to those involving male victims (Glaeser & Sacerdote, 2003). Based on this past work, it seems that greater empathy is not necessarily afforded to male compared to female perpetrators. If anything, female defendants are afforded greater leniency within decisions to prosecute. Additionally, the gender of the

victim (rather than the defendant) seems to be more influential in sentencing decisions of capital punishment cases, and in one analysis of violent crime data described above, male perpetrators who victimized a female were most harshly punished.

Given this, might the case type matter in shaping judgments of a defendant based on their gender? In cases depicting sexual assault however, the presence of Himpathy is still not apparent: different experimental studies have shown that mock jurors are more likely to propose harsher, longer sentences, and provide more guilt ratings towards male compared to female defendants accused of sexual assault (Cox & Kopkin, 2016; Pozzulo et al., 2010; Quas et al., 2002).

The degree of victim blaming, and attributions of responsibility granted to the victim within a rape case – facets also strongly associated with Himpathy – do seem to differ by the gender of both the juror and the rape victim. Ford, Liwag-McLamb, & Foley (1998), found that within a sample of college students, the degree of victim blaming was contingent on the victim's sex and sexual orientation, such that heterosexual female victims and homosexual male victims were perceived as more responsible (Ford et al., 1998). Additionally, Pozzulo, Dempsey, Mader, & Allen (2010) found that in response to a fictional sexual assault case, female compared to male jurors assigned more credibility and believability to the victim (irrespective of the victim's gender); male jurors were also more likely to believe that the victim desired and caused the sexual assault, compared to female jurors. Although overall male defendants were given higher guilt ratings than female defendants, there were differences in defendant judgments by juror gender. That is, male compared to female jurors perceived the defendant as more credible and believable, while females held more pro-victim attitudes (Pozzulo et al., 2010).

More recent work by Dodson et al. (2020) presented participants with vignettes of a

female victim bringing forth sexual assault allegations and the male perpetrator denying the allegations. The authors next evaluated participants' judgments of the victim's blameworthiness, how angry they felt toward the victim, their levels of sympathy toward the perpetrator, and their degree of ascription to moral values upholding authority and ingroup loyalty (known as binding moral values). Their results upheld the notion of Himpathy; namely, the greater participants' endorsement of binding moral values, the more they blamed the victim and held them responsible. This relationship was driven by participants' anger toward the female victim and sympathy towards the male perpetrator (Dodson et al., 2020).

Given the above empirical work, it is likely that even in the absence of an effect of Himpathy in terms of evaluating empathy towards criminal offenders, facets of Himpathy can be revealed in terms of evaluating participants' perceptions of victim blameworthiness and responsibility following a sexual assault allegation.

The Role of Sexual Assault Defendant and Victim Gender in Pre-Trial Decisions

Beyond juror perceptions of both defendants and victims, how does the gender of the victim and defendant affect pre-trial decisions, such as the decision to press charges on the offender and bring the case to trial? A large body of work examining pre-trial prosecutorial decisions in various criminal cases presents evidence contrary to what Himpathy may predict.

This work demonstrates that following a grand jury indictment, prosecutors are less likely to decide to continue prosecuting a female defendants, compared to a male defendant (Albonetti, 1986). It additionally demonstrates no substantial difference by a defendant's gender on the likelihood of a charge reduction (Bishop & Frazier, 1984). Another analysis of pretrial decision data from Los Angeles County showed that female defendants of all racial groups were more likely to have their charges rejected or dismissed compared to their male counterparts (Spohn et

al., 1987).

Regarding the case processing of sexual assault cases specifically, prior work has found that the prestige level of the defendant (Benedict & Klein, 1997), the defendant and victim's race (see the sexual stratification hypothesis, Kelley et al., 2021; O'Neal et al., 2019; Spohn & Spears, 1996; Walsh, 1987), and the relationship between the victim and the defendant (Kelley et al., 2021; Spohn & Holleran, 2001) can all influence rates of arrest and filing charges against the defendant.

In their analysis of sexual assault allegations against male college-level and professional athletes, Benedict & Klein (1997) found that as professional athletes, male defendants holding this level of prestige were more likely to be arrested and indicted (though also less likely to ultimately be convicted), compared to the Bureau of Justice Statistics' national felony prosecution rates. Thus, contrary to what Himpathy would predict, males in a high-prestige position in this analysis did not receive preferential treatment in terms of lower pre-trial decisions to prosecute them.

Sexual Harassment Cases

Within the context of sexual harassment cases in the civil domain, a large body of work has examined individual differences in if, and to what extent, individuals perceive certain behaviors to constitute harassment. These perceptions may vary by the respondent's gender (Madan & Nalla, 2015; Rotundo et al., 2001), type of legal standard used - the reasonable person versus the reasonable woman standard (Blumenthal, 1998; Hurt et al., 1999), the extent of one's ascription towards sexist attitudes (Russell & Trigg, 2004; Wiener & Hurt, 2000), as well as the presence of prototypical #MeToo features of sexual harassment within the situation: a superior male perpetrator, repeatedly initiating sexual contact in a private setting (Kessler et al., 2020).

Beyond classifying behaviors as harassing (or not), prior research has illuminated the role of the timing of a victim's harassment accusations on subsequent victim and perpetrator assessments (Balogh et al., 2003; Lucarini et al., 2020; von Sikorski & Saumer, 2021). Work by von Sikorski & Saumer (2020), for example, found that the longer the victim's delay in making an accusation of harassment, the more participants were likely to ascribe negative motives and blame onto the harassment victim. Work by Balogh et al. (2003) and Lucarini et al. (2020) illustrated a similar pattern of findings.

Much less work, however, has examined the role of gender – of the defendant, plaintiff, or judge assigned to the case - on the decision to bring a sexual harassment suit to trial to be prosecuted. The current work seeks to fill this gap in the literature, while also assessing the role of the defendant's prestige level and the timing of the victim's reporting on the decision to prosecute.

The Current Study: Research Questions

The proposed work seeks to answer if Himpathy exists within sexual harassment cases, specifically within situations involving the creation of a hostile workplace environment. Specifically, we examine whether respondents have heightened empathy for male, high-powered perpetrators of sexual harassment in a workplace setting. If so, does this heightened empathy manifest itself in lowering the likelihood of pre-trial decisions to bring a harassment suit against the defendant to trial? Related to the victim of the harassment, may Himpathy manifest itself in terms of more negative attitudes towards this victim? And how may a victim's gender or the timing of their accusation following the harassment affect subsequent attitudes toward them? Additionally, is the presence of Himpathy impacted by characteristics of the assessors themselves, such as their gender, age, political orientation, the moral values they ascribe to, or

their attitudes about sexism and gender roles? By answering these questions, we hope to illuminate gender biases that may be at play within multiple prosecution stages of sexual misconduct cases, amplifying victims' voices and ensuring more equitable prosecution.

The Current Study: Hypotheses

First, across hypotheses, Himpathy will be conceptualized as an overall preferential bias towards male perpetrators of harassment and negativity towards a female harassment victim. These attitudes will be expressed by participants (details of whom will be delineated in the Methods section) who are presented with various hypothetical scenarios depicting workplace sexual harassment.

With this conceptualization in mind, it is first hypothesized that Himpathy will be most prevalent in situations in which the perpetrator of the harassment is a male boss (in a position of prestige and power within a workplace) and the female victim ultimately reports the harassment one year after it takes place (i.e. a significant timing delay in reporting). The effect of Himpathy will be manifested in terms of the following hypothesized attitudes participants may hold toward the perpetrator and victim of the harassment.

Regarding the perpetrator, we hypothesize that participants will be more likely to believe: a) there is insufficient evidence to pursue civil action against the perpetrator (a proxy for pretrial decisions), b) rate the perpetrator as more likeable and trustworthy, less blameworthy, and c) express greater sympathy and forgiveness towards the perpetrator.

In contrast, we hypothesize that participants will perceive the victim as being less trustworthy, more blameworthy, and will express less sympathy towards them.

This set of hypotheses aligns with the philosophical definition of Himpathy as being

perpetuated by high-power, male perpetrators of sexual misconduct against female victims, and is in line with prior work demonstrating that the longer a victim delays reporting the harassment, the more the negative the views towards them tend to be (see Balogh et al., 2003; Lucarini et al., 2020; von Sikorski & Saumer, 2020). However, these hypotheses are in contradiction to aforementioned work demonstrating that men are not always afforded greater empathy in sexual assault or other types of cases (e.g. Cox & Kopkin, 2016; Curry et al., 2004). Additionally, in terms of prestige and pretrial decisions, these hypotheses would contrast with work showing that highly privileged male accusers of assault do not necessarily experience lower rates of arrest and indictment (see Benedict & Klein, 1997).

Methodology

Participants

A total of 216 online participants were recruited via Academic Prolific for an online survey study. Subjects were compensated \$2.17 for their participation. The survey was administered via Qualtrics. Of the 216 participants, 16 did not meet the study's inclusion criteria (13 did not finish the study, two participants did not provide real Prolific ID's, and one person did not consent to the study), such that 200 participants were ultimately included in the analysis.

Participants' ages ranged from 18 to 56 ($M_{\text{age}} = 25.48$, $SD = 7.21$), and included 112 males (56.28%), 84 females (42.21%), one participant who identified as non-binary (0.5%), one participant who did provide an answer to the gender inquiry (0.5%), and two participants who selected that they preferred to not answer the inquiry (0.50%). The sample was predominantly White (72.50%), 11.50% Hispanic, 5.5% Multiracial, 4% Asian / Pacific Islander, 4% Other, 2% Black or African American, and one participant (0.50%) did not provide their race.

Design

The study conformed to a 2 (Defendant and Victim Gender: Male / Female, Female / Male) x 2 (Timing of Victim's Reporting: 2 weeks, 1 year) x 2 (Perpetrator's prestige status relative to the victim: the Victim's Boss, the Victim's Colleague) fully randomized between-subjects design. Participants were randomly assigned to receive information pertaining to one of the eight resulting conditions:

Condition 1: Female Perpetrator, Perpetrator is the Male Victim's Colleague, 1 year Timing of Victim's Reporting (N = 24)

Condition 2: Female Perpetrator, Perpetrator is the Male Victim's Colleague, 2-week Timing of Victim's Reporting (N = 22)

Condition 3: Female Perpetrator, Perpetrator is the Male Victim's Work Superior, 1-year Timing of Victim's Reporting (N = 27)

Condition 4: Female Perpetrator, Perpetrator is the Male Victim's Work Superior, 2-week Timing of Victim's Reporting (N = 25)

Condition 5: Male Perpetrator, Perpetrator is the Female Victim's Work Superior, 1-year Timing of Victim's Reporting (N = 27)

Condition 6: Male Perpetrator, Perpetrator is the Female Victim's Work Superior, 2-week Timing of Victim's Reporting (N = 24)

Condition 7: Male Perpetrator, Perpetrator is the Female Victim's Colleague, 1-year Timing of Victim's Reporting (N = 26)

Condition 8: Male Perpetrator, Perpetrator is the Female Victim's Colleague, 2-week Timing of Victim's Reporting (N = 25)

The content of each condition is further delineated in the Procedure, and a sampling of the exact wording of each basic scenario-type can be found in Appendix A.

Procedure

Participants read an informed consent form, which included a warning of the sensitive nature of the topic (sexual harassment allegations), indicated that the study would take about 30 minutes to complete, and that they would be compensated \$2.71 for their participation.

Upon consenting to participate in the study, participants were randomly assigned to one of eight scenarios (in alignment with the between-subjects design described above) depicting a sexual harassment, hostile work environment situation between two individuals at a company. The scenarios depict a male (Marcus) or female (Lauren) employee at a consulting firm, who experiences sexual harassment (conceptualized in these scenarios as repeated sexual commentary on one's body, repeatedly being asked out on dates despite persistent rejections) by their male or female work superior (conceptualized here as the victim's project manager) or colleague. The plaintiff reports the events to the company's Human Resources department, but they do not take any remedial action. In light of the company's inaction, and the hostile and unpleasant work environment generated by the harassment situation, the plaintiff leaves the company, in spite of the income they know they will lose from their departure. Later, either two weeks or one year after the harassment initially started, the plaintiff sues the defendant and the company for pain and suffering and back wages (lost income during their departure). To substantiate the case facts, participants also read perpetrator and victim testimony statements. Participants were told these deposition statements were collected by the attorney, to better understand each party's perspective in the situation. These statements were consistent in content (with the exception of the perpetrator and victim's gender, and the timing delay of the victim's reporting) across scenarios.

As a manipulation check, participants were asked to indicate (depending on the scenario

they received) the extent to which they agreed with the notion (1 Strongly Disagree – 5 Strongly Agree) that the perpetrator they read about had a higher authority status than the victim, and that the victim had a significant delay between the incident of harassment and their lawsuit filing.

Next, participants read the federal definition of sexual harassment, as outlined by the Code of Federal Regulations. The definition read as follows:

"(a) Harassment on the basis of sex is a violation of section 703 of title VII. 1 Unwelcome sexual advances, requests for sexual favors, and other verbal or physical conduct of a sexual nature constitute sexual harassment when (1) submission to such conduct is made either explicitly or implicitly a term or condition of an individual's employment, (2) submission to or rejection of such conduct by an individual is used as the basis for employment decisions affecting such individual, or (3) such conduct has the purpose or effect of unreasonably interfering with an individual's work performance or creating an intimidating, hostile, or offensive working environment."
(16 Code of Federal Regulations Section 1604.11).

Participants were asked to imagine they were the judge overseeing this case and whether, based on the aforementioned federal definition, they believed that there were sufficient grounds to sue the defendant (either Marcus or Lauren, based on the condition they received) and pursue legal action, or that they believed there were insufficient grounds to sue and the case should be dismissed. Next, participants were asked a set of questions assessing their attitudes towards the perpetrator and victim. These included four questions assessing the defendant on 7-point scales of their likability (1 Very Unlikable – 7 Very Likable), bad-good (1 Very Bad – 7 Very Good), uncaring – caring (1 Very Uncaring – 7 Very Caring), and cruel-kind (1 Very Cruel – 7 Very Kind). Participants' ratings on these questions were averaged to generate a global liking score for the defendant. They were then asked to assess how believable the situation they read was (1 Not at All Believable – 7 Very Believable), how trustworthy they felt the perpetrator and victim were (1 Not at All Trustworthy – 7 Very Trustworthy), how blameworthy and responsible the perpetrator and victim each were for their actions (1 Not at All – 5 A Great Deal), to what extent

they would forgive the perpetrator for their actions (1 Wouldn't forgive them at all – 11 Would forgive them completely), and to what extent they felt sympathetic to the victim and perpetrator in the situation they read about (1 Not at All – 5 A Great Deal).

They were additionally asked to provide an open-ended explanation for the blameworthiness and responsibility rating they provided for both the perpetrator and victim. These qualitative responses were later thematically analyzed by a team of research assistants. Lastly, participants completed a set of attitudinal questionnaires to assess their behavioral and emotional forgiveness attitudes towards the defendant, baseline empathic tendencies, sexist attitudes, moral values, and their demographic information was collected.

Materials

A) Decisional and Emotional Forgiveness Scales

To substantiate the 11-point forgiveness rating, participants were asked to provide and to assess their forgiveness attitudes in more detail via the Decisional and Emotional Forgiveness scales, developed by Everett Worthington and colleagues. The Decisional Forgiveness scale is composed of eight 5-point items (1 Strongly Agree – 5 Strongly Disagree), has high internal reliability of $\alpha = 0.82$, and assesses behavioral changes within the forgiver towards a transgressor. For example, in the current study participants were asked the extent to which, if they were the victim, they would not talk to the defendant, intentionally hurt the defendant, seek revenge, etc. Due to an inputting error, the 5th item in this scale was omitted from the survey. The Emotional Forgiveness scale is also composed of eight 5-point items (1 Strongly Agree – 5 Strongly Disagree), has high internal reliability of $\alpha = 0.83$, and assesses a forgiver's positive and negative emotions towards a transgressor (Worthington et al., 2007). In the current study,

participants were asked to imagine they were the victim, and to indicate the extent to which they would feel bitter, upset, care, love, etc. towards the defendant, following the situation they read about.

B) Moral Foundations Questionnaire

Driven by prior work showing that ascription towards the moral dimensions of in-group loyalty and authority (Dodson et al., 2020) can affect reactions to sexual misconduct situations, participants completed the Moral Foundations Questionnaire, or MFQ-30. This questionnaire assesses five dimensions of moral intuitions: Harm/Care, Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, and Purity/Sanctity (Graham et al., 2011).

C) Toronto Empathy Questionnaire

The current study sought to examine how a respondent's baseline, more trait versus state-like empathic tendencies – i.e. empathy irrespective of the situation at hand - could affect their reactions to the depicted sexual misconduct.. This was assessed via the Toronto Empathy Questionnaire (Spreng et al., 2009) was administered to participants. This questionnaire has high internal reliability (Cronbach's $\alpha = 0.85 - 0.87$) and consists of 16 items regarding one's degree of agreement (on a five-point scale, from 0 = Never to 4 = Always) with the extent to which they feel or act in emotion (versus cognitive)-based empathic ways. For example, sample items include: "*I have tender, concerned feelings for those less fortunate than me.*" or "*I become irritated when someone else cries.*"

D) Ambivalent Sexism Inventory (ASI)

The Ambivalent Sexism Inventory, or ASI, was developed by Glick & Fiske in 1996 as a means to assess two sub-components of sexist attitudes towards women: benevolent sexism, a set of attitudes and behaviors that are sexist in that they conform to traditional stereotypes about

women that may seem benevolent to the perceiver, but can be damaging to the female recipient of such stereotypical, male-dominated rhetoric. They contrast this with hostile sexism, which is conceptualized as more overt, outwardly negative and damaging statements or beliefs about women. The inventory overall has high internal reliability, Cronbach's $\alpha = 0.85 - 0.87$ (Glick & Fiske, 1996).

Results

Manipulation Check

A series of independent samples t-tests were run to assess respondents' average degree of agreement to the two manipulation check questions (dependent variable), based on the independent variables of the victim's reporting status (2 weeks vs. 1 year), and the perpetrator's prestige level (a boss or colleague). Each t-test confirmed that the experimental manipulation worked. That is, participants in conditions depicting the victim reporting their harassment 1 year following the initial incident ($N = 104$) had significantly higher average agreement with the statement that the victim had a significant reporting delay ($M = 4.07$, $SD = 1.09$), compared to participants in conditions depicting a 2-week reporting delay ($N = 96$), ($M = 2.79$, $SD = 1.32$); $t(184.702) = 7.41$, $p < 0.001$, 95% CI: 0.94 – 1.65.

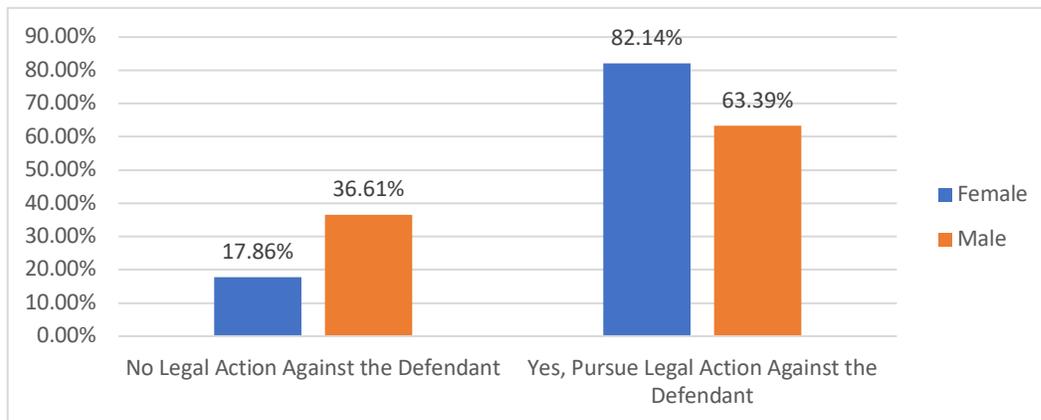
Similarly, participants in conditions depicting the perpetrator as the victim's boss ($N = 103$) had significantly higher agreement with the statement that the perpetrator had a higher authority status than the victim ($M = 4.5$, $SD = 1.11$) compared to those in conditions depicting the perpetrator as the victim's colleague ($N = 97$), ($M = 2.49$, $SD = 1.41$); $t(183.2) = 11.08$, $p < 0.001$, 95% CI: 1.64 – 2.36.

Effect of Condition and Participant Gender on Participants' Decision to Pursue Legal Action

A binary logistic regression was run to whether participants' gender or the condition they

were exposed to (varying the perpetrator's gender, prestige status, and the victim's reporting timing) significantly predicted their decisions to pursue legal action against the defendant they read about. Across conditions (N = 196), 140 participants (71.4%) indicated they felt there was sufficient evidence to pursue legal action against the defendant in question (Lauren or Marcus), while 56 participants (28.6) indicated that there was insufficient evidence to pursue legal action. Logistic regression results indicated that none of the experimental manipulations: perpetrator gender ($\beta = 0.29$, SE = 0.33, Wald (1) χ^2 , 0.76, $p = 0.38$; Odds Ratio = 1.33), prestige status ($\beta = -0.04$, SE = 0.33, Wald (1) χ^2 , 0.02 $p = 0.90$; Odds Ratio = 0.96), or the victim's reporting timing ($\beta = -1.02$, SE = 0.32, Wald (1) χ^2 , 0.20, $p = 0.66$; Odds Ratio = 1.15) significantly predicted decisions to pursue legal action against the defendant. However, participants' gender emerged as a significant predictor of decisions to pursue legal action ($\beta = -1.02$, SE = 0.35, Wald (1) χ^2 , 8.30, $p = 0.004$; Odds Ratio = 0.36). Specifically, it was found that the odds of male participants deciding to pursue legal action against a defendant was 0.36 times *lower* than that of female participants. Targeted chi-square tests of independence were run to disaggregate the significant effect of participant gender on decisions to pursue legal against a defendant ($\chi^2 (1, N = 196) = 8.27$, $p = 0.004$). As illustrated in Figure 2.1, the proportion of female participants who indicated that they would pursue legal action against a defendant (82.1%, N = 69) was significantly different than the proportion of males who did so (63.4%, N = 71).

Figure 2.1 Proportion of Males vs. Females who Indicated Legal Action Against A Defendant



Within Conditions 5-8 (N = 102), in which Marcus was depicted as the perpetrator, and his prestige status and the female victim's reporting timing were varied, 28 participants (27.5%) indicated that there was insufficient evidence to pursue legal action against Marcus, while 74 (72.5%) participants indicated there was sufficient evidence to pursue legal action.

A series of binary logistic regression analyses were conducted to assess the effect of condition (varying defendant and victim gender, the defendant's prestige status, and the victim's reporting delay) and participants' gender on the dichotomous decision to pursue legal action against the defendant or not. Decisions to pursue legal action were coded as 1, and decisions against pursuing legal action were coded as 0.

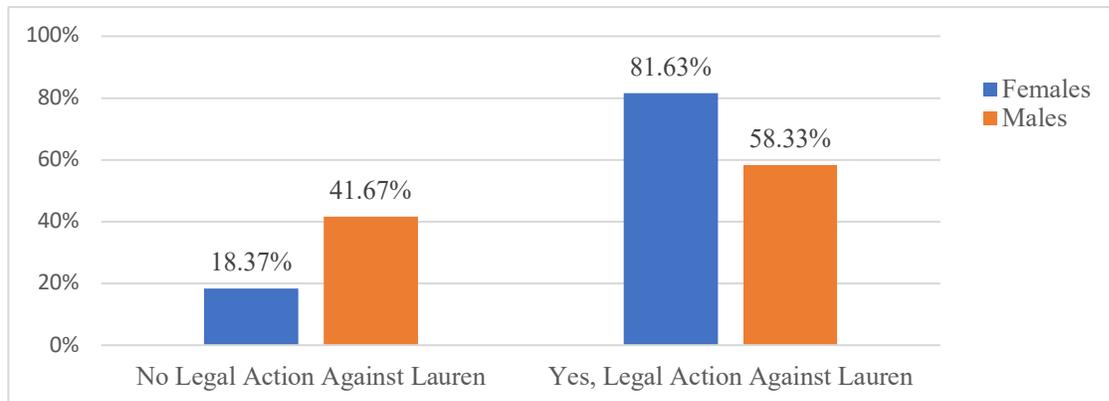
A logistic regression model (with the aforementioned predictors) was run for participants who were in the experimental manipulation that prompted them to answer whether legal action should be pursued against Marcus or not, as well as for participants who answered the same question regarding Lauren. Due to their small cell counts (1 participant who identified as non-binary, 2 who preferred to not answer, and 1 participant who did not answer), these 4 participants were excluded from the analyses.

However, logistic regression results indicated that participants' gender ($\beta = -0.852$, $SE = 0.53$, $Wald(1) \chi^2 = 2.64$, $p = 0.1$; Odds Ratio = 0.43) the perpetrator's prestige status ($\beta = -0.07$, $SE = 0.46$, $Wald(1) \chi^2 = 0.02$, $p = 0.89$; Odds Ratio = 0.94), nor the victim's reporting timing ($\beta = 0.34$, $SE = 0.46$, $Wald(1) \chi^2 = 0.53$, $p = 0.47$; Odds Ratio = 1.40) significantly predicted participants' likelihood of deciding to pursue legal action against Marcus. An additional exploratory binary logistic regression was conducted to examine whether attitudes towards the defendant and victim – specifically, the believability of the situation, perceptions of blameworthiness and trustworthiness, and general liking of the perpetrator – significantly predicted participants' decision to pursue legal action against Marcus. However, none of these variables were determined to be significant predictors.

Within Conditions 1-4 ($N = 98$), in which Lauren was depicted as the perpetrator, and her prestige status and the male victim's reporting timing were varied, 29 participants (29.6%) indicated that there was insufficient evidence to pursue legal action against Marcus, while 69 participants (70.4%) participants indicated there was sufficient evidence to pursue legal action. Binary logistic regression analyses indicated that neither the female perpetrator's prestige status nor the victim's reporting timing were significant predictors. However, participant gender emerged as a significant predictor; $\beta = -1.15$, $SE = 0.47$, $Wald(1) \chi^2 = 5.93$, $p = 0.02$; Odds Ratio = 0.32; 95% CI: [0.13 – 0.80]. Specifically, it was found that the odds of male participants deciding to pursue legal action against Lauren was 0.32 times *lower* than that of female participants. Targeted chi-square tests of independence were run to disaggregate the significant effect of participant gender on decisions to pursue legal against Lauren $\{\chi^2(1, N = 97) = 6.28, p = 0.01\}$. As illustrated in Figure 2.2, the proportion of females who indicated that they would pursue legal action against Lauren (81.68%, $N = 49$) was significantly different than the

proportion of males who indicated legal action (58.33%, N = 48).

Figure 2.2 Proportion of Males vs. Females who Indicated Legal Action Against Lauren (or not)



An additional exploratory binary logistic regression was conducted to examine whether attitudes towards the defendant and victim – specifically, the believability of the situation, perceptions of blameworthiness and trustworthiness, and general liking of the perpetrator – significantly predicted participants’ decision to pursue legal action against Lauren. All variables were found to be non-significant predictors, though ratings of the perpetrator’s trustworthiness approached significance ($\beta = -0.55$, $SE = 0.30$, Wald (1) χ^2 , 3.35, $p = 0.067$; Odds Ratio = 0.58). That is, with every one-unit increase in trustworthiness ratings, the odds of pursuing legal action against Lauren *decreased* by 0.58.

Situation Believability Across Conditions

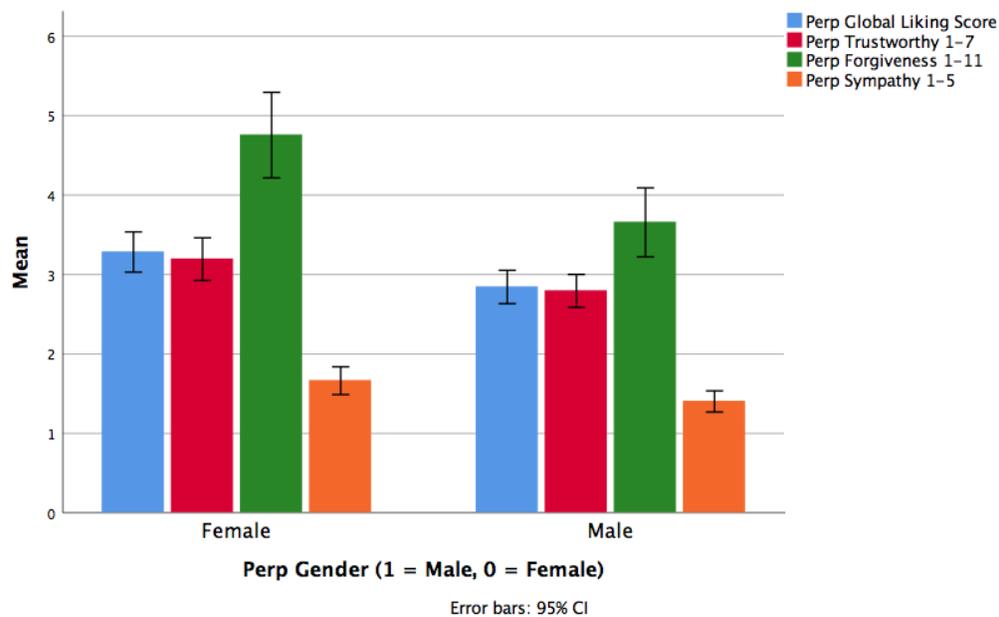
Across all conditions, participants’ average rating of the believability of the situation they read was 5.82 (SD = 1.11, Min = 1, Max = 7). A one-way ANOVA was conducted to assess the effect of condition on participants’ ratings of situation believability. However, the main effect of condition was not significant; $F(1,192) = 0.68$, $p = 0.67$.

Effects of Condition, Baseline Empathy on Perpetrator and Victim-Related Ratings:

Perpetrator Ratings

A multivariate analysis of covariance (MANCOVA) was conducted to examine the effect of perpetrator gender, victim reporting timing, and perpetrator prestige level (entered as fixed variables) on the combined dependent variables of average ratings of perpetrator forgiveness, global liking scores, trustworthiness, and ratings of sympathy towards the perpetrator, while controlling for participants' baseline empathy scores (as assessed via the Toronto Empathy Questionnaire, entered as a covariate). A main effect of perpetrator gender was found $\{F(4,188) = 3.29, p = 0.012, \text{Wilks' } \Lambda = 0.94, \eta_p^2 = 0.07\}$. Pairwise comparisons revealed that when the perpetrator in the sexual harassment scenario was a female (and the victim was a male) compared to a male perpetrator (and female victim), participants were significantly more likely to trust ($M_{\text{femaleperp}} = 3.19, SE_{\text{femaleperp}} = 0.12; M_{\text{maleperp}} = 2.78, SE_{\text{maleperp}} = 0.12; p = 0.02$), have a higher global liking of ($M_{\text{femaleperp}} = 3.29, SE_{\text{femaleperp}} = 0.12; M_{\text{maleperp}} = 2.84, SE_{\text{maleperp}} = 0.12; p = 0.008$), forgive ($M_{\text{femaleperp}} = 4.76, SE_{\text{femaleperp}} = 0.25; M_{\text{maleperp}} = 3.64, SE_{\text{maleperp}} = 0.24; p = 0.001$), and sympathize with the perpetrator ($M_{\text{femaleperp}} = 1.67, SE_{\text{femaleperp}} = 0.08; M_{\text{maleperp}} = 1.40, SE_{\text{maleperp}} = 0.08; p = 0.013$). Figure 2.3 displays this main effect graphically.

Figure 2.3 Average Participant Ratings of Perpetrator Liking, Trustworthiness, Forgiveness, and Sympathy, by Perpetrator Gender



No other significant main effects - by victim reporting timing, perpetrator prestige, or participant baseline empathy scores - were found, and the three-way interaction between perpetrator gender, victim reporting timing, and perpetrator prestige was non-significant $\{F(16, 574.99) = 0.94, p = 0.52, \text{Wilks' } \Lambda = 0.92, \eta_p^2 = 0.02\}$.

A similarly structured MANCOVA model was run for Conditions 5-8, which depicted a male perpetrator of sexual harassment against a female victim. Here, the same perpetrator-relevant dependent variables were used, perpetrator prestige and victim reporting timing were entered as fixed factors, and Hostile and Benevolent sexism scores were added as additional covariates (along with baseline empathy scores). Results indicated no significant main effects of perpetrator prestige or victim reporting timing (controlling for sexism and baseline empathy scores), However, a main effect of Hostile Sexism emerged $\{F(4,93) = 8.39, p < 0.001, \text{Wilks' } \Lambda = 0.74, \eta_p^2 = 0.27\}$. That is, for every one unit increase in participants' hostile sexism (against women, as assessed by the Ambivalent Sexism Inventory), their trustworthiness of the male perpetrator increased by 0.03 ($t = 3.45, p = 0.001$), global liking increased by 0.04 ($t = 4.34, p < 0.001$), forgiveness increased by 0.08 ($t = 4.42, p < 0.001$), and sympathy towards the perpetrator increased by 0.02 ($t = 3.86, p < 0.001$).

Victim Ratings

The same model of MANCOVA (same fixed factors and covariate) were run regarding victim judgments, using average victim sympathy and victim trustworthiness as the dependent variables. However, no significant main effects of either victim reporting timing $\{F(2, 194) = 0.27, p = 0.77, \text{Wilks' } \Lambda = 0.997, \eta_p^2 = 0.003\}$, perpetrator prestige $\{F(2, 194) = 0.72, p = 0.49, \text{Wilks' } \Lambda = 0.99, \eta_p^2 = 0.007\}$, perpetrator gender $\{F(2, 194) = 0.48, p = 0.62, \text{Wilks' } \Lambda = 0.995,$

$\eta_p^2 = 0.005$ }, or baseline empathy scores $\{F(2, 194) = 0.53, p = 0.59, \text{Wilks' } \Lambda = 0.995, \eta_p^2 = 0.005\}$ were found.

A similarly structured MANCOVA model was run for Conditions 5-8 (depicting a male perpetrator and female victim). The same victim-relevant dependent variables were used in this model, perpetrator prestige and victim reporting timing were entered as fixed factors, and Hostile and Benevolent sexism scores were added as additional covariates (along with baseline empathy scores). The results of this model mirrored those for perpetrator-related judgments; no significant main effects of perpetrator prestige or victim reporting timing (controlling for sexism and baseline empathy scores) were found. However, a main effect of Hostile Sexism emerged $\{F(2, 95) = 20.57, p < 0.001, \text{Wilks' } \Lambda = 0.70, \eta_p^2 = 0.30\}$; for every one unit increase in participants' hostile sexism (against women, as assessed by the Ambivalent Sexism Inventory), their trustworthiness of the female victim decreased by 0.04 ($b = -0.04, t = -4.76, p < 0.001$), and their sympathy toward the victim also decreased by 0.05 ($b = -0.05, t = -6.23, p < 0.001$).

Ambivalent Sexism Scores, Across Conditions

An independent-samples t-test revealed that across conditions, participants' average ambivalent sexism scores significantly differed by their gender; $t(194) = -7.07, p < 0.001$; 95% CI: $[-13.69 - -7.72]$. Men scored higher than women, on average, on both the hostile ($M_{\text{male}} = 23.85, SD_{\text{male}} = 10.69, M_{\text{female}} = 13.14, SD_{\text{female}} = 10.22$), and benevolent sexism inventories ($M_{\text{male}} = 19.72, SD_{\text{male}} = 8.76, M_{\text{female}} = 14.36, SD_{\text{female}} = 7.77$).

Perpetrator and Victim Blameworthiness

Participants' average blameworthiness rating for the perpetrator, across conditions, was 3.85 (out of 5), $SD = 1.02$. A one-way ANOVA was conducted to examine the impact of the experimental manipulations (perpetrator gender, victim reporting timing, and perpetrator

prestige) on ratings of perpetrator blameworthiness. No significant main effects or interaction effects emerged.

Participants' average blameworthiness rating for the victim, across conditions, was 1.67 (out of 5), SD = 1.02. Moreover, a one-way ANOVA indicated that victim blameworthiness ratings did not significantly differ based on any of the experimental manipulations.

The Role of Moral Foundations on Victim Blaming

Different assumptions inherent in linear regression were violated (bias identified via casewise diagnostics, assumptions of linearity and homoscedasticity) when initially assessing the predictive role of moral foundations on perpetrator and victim blaming (across conditions), so the regression was re-run using the more robust method of bootstrapping. This linear regression indicated that the greater a participant's endorsement of moral values involving purity and sanctity, the *less* attribution of blame they placed on the perpetrator (across conditions) { $b = -0.016$, $p = 0.02$, 95% CI: $-0.046 - 0.012$ }. Regarding victim blaming, bootstrapped-regression results indicated that the greater a participant's endorsement of moral values of authority and respect { $b = 0.04$, $p = 0.036$; 95% CI: $[0.011 - 0.061]$ }, the greater their victim blaming. Additionally, the greater a participant's endorsement of the moral values of harm and care, the *less* their victim blaming { $b = -0.048$, $p = 0.048$, 95% CI: $[-0.087 - -0.017]$ }.

Thematic Coding of Participants' Open-Ended Responses

A team of research assistants thematically analyzed participants' open-ended responses to their perceptions of the perpetrator and victim's blameworthiness and responsibility, in the situation they read about. A list of the most common themes identified within these responses, as well as illustrative quotes for each theme, can be found in Table 2.1.

Table 2.1 Qualitative Themes from Participants' Explanations of Perpetrator and Victim Blameworthiness and Responsibility

Perpetrator Blameworthiness Theme	Illustrative Quote	Victim Blameworthiness Theme	Illustrative Quote
No Proof	"There were no witnesses, so there is no way to prove responsibility or blameworthy on her actions"	No proof	"Same goes for the victim in this case. We don't know anything about her, nor do we have any proof of the sexual harassment. So we can't really judge her."
HR is more responsible	"Her actions (and the inactions of the HR) are the only reason everything happened."	HR is more responsible	"HR department of the company should deal with the matter earlier"
Responsible for your own actions	"Every normal, fully functional adult is always responsible for their actions, regardless of the situation."	Reporting delay	"If the facts had been so serious she would not have had to wait a year to make the complaint"
Power Dynamic/Perpetrator was in a position of authority	"In my opinion, the perpetrator of this situation abused of her power and superiority to intimidate the victim. She made inappropriate comments about him, making him feel {uncomfortable} to the point of him quitting."	Gender stereotypes make victim more believable	"I imagine the victim as a shy man, that has to work with an expansive woman. I think he knows that he is not victim of a cruel harassment, that's why it took a year to go to a lawyer."
Allegations seem overblown/exaggerated	"Some of the allegations seem overblown."	Never blame victim/never their fault	"You can't blame a victim for their actions. Unless it is proven that she is lying about what transpired you have to trust the person who is experiencing the harassment."
Perpetrator knew victim was uncomfortable/sexual harassment was	"Even if he was trying to be nice, the victim told him multiple times that she doesn't like the way	Victim was clear in expressing discomfort/telling them to stop	"He told her to stop, it's not his fault if she continued"

Perpetrator Blameworthiness Theme	Illustrative Quote	Victim Blameworthiness Theme	Illustrative Quote
intentional	he talks to her. There is no one else to blame for his actions."		
Perpetrator is responsible based on victim's loss of job/distress	"He made unpleasant comments that led Lauren to quit her job."	Victim could have done more (ex: expressing more, collecting more evidence)	"He could have asked for help from his colleagues or once again from HR."

Discussion

This study aimed to examine the presence of Himpathy, or exaggerated sympathy towards a male perpetrator of sexual misconduct against a female victim, in the context of pre-trial decisions to bring a case to trial and perpetrator and victim-related attitudes. However, affirmative evidence of Himpathy was scant in the current data.

Null Effects of Himpathy: Pre-Trial Decisions to Bring Sexual Harassment Case to Trial

We hypothesized that in situations depicting a male perpetrator of harassment and female victim, Himpathy would be manifested in terms of decisions to prosecute the male perpetrator (Marcus) being lowest in instances in which he is in a position of prestige, and his female victim took one year from the incident to report it. However, surprisingly, neither the perpetrator's prestige status nor the victim's timing delay emerged as significant predictors of the decision to prosecute. The same null effects of perpetrator prestige and victim timing delay held for situations depicting a female perpetrator (Lauren), though interestingly, female participants were more likely to state that Lauren's case should move to trial than male participants. Evidence of such in-group bias (females holding more punitive attitudes against a female perpetrator than males) was unexpected and contrary to initial hypotheses. This in-group bias should be further investigated in future work.

Perpetrator and Victim Ratings

Lack of Himpathy effects (i.e. positive, preferential treatment towards a male perpetrator) can also be revealed through participants' rating of the perpetrator and victim. When participants were presented with situations depicting a female (vs male) harassment perpetrator (irrespective of the perpetrator's prestige status or the male victim's reporting timing), they were more likely to trust, forgive, sympathize, and generally like the perpetrator.

The lack of preferential, more lenient treatment towards a male perpetrator of sexual harassment revealed in these analyses is in alignment with previous work in the sexual assault literature indicating that male perpetrators of sexual misconduct (especially those involving female victims) are often treated the most harshly in terms of both pretrial decisions (Benedict & Klein, 1997) and verdict and sentencing decisions (Cox & Kopkin, 2016; Pozzulo et al., 2010; Quas et al., 2002). However, it also brings to light the notion that Himpathy may be most clearly illuminated when it is conceptualized as another form of sexism at large, rather than a type of sexism that only arises in particular instances (i.e. involving a certain type of perpetrator only). Some of the affirmative evidence of Himpathy found in this current work speaks to this notion.

Affirmative Evidence of Himpathy: The Role of Sexist Attitudes

The current work found that in conditions depicting a male perpetrator of sexual assault against a female victim, participants' sexist attitudes, not specifics of the scenarios they were exposed to, were particularly influential in shaping their attitudes towards the perpetrator and victim. The greater a participant's ascription to hostile sexist attitudes, the more likely they were to trust, globally like, forgive, and sympathize with the male perpetrator, while also being less likely to trust and sympathize with the female victim. As aforementioned, this relationship indicates that Himpathy may be best conceptualized as a particular form of sexism in the context

of sexual misconduct, rather than a phenomenon only exhibited when certain perpetrator characteristics are in place.

Situation Believability

Specific hypotheses were not generated regarding participants' degree of belief in the harassment situation they read about. However, evidence that participants might be less likely to believe a harassment depicting a female perpetrator and male victim could be an indication of participants upholding male rape myths, or holding incorrect, harmful beliefs that instances of male-victim rape are rare, or that "men can't be raped" (Walfield, 2021). Participants' average ratings of situation believability did not vary by the condition they were exposed to, and in general was quite high across conditions (5.82 out of 7), indicating that participants did not think certain situations of harassment any less probable than others. It cannot, however, be completely discounted that certain participants may ascribe more strongly to certain rape myths than others (despite their believability ratings); future studies should thus incorporate measures of rape myth acceptance, to gain a broader understanding of participants' likelihood of holding false, harmful beliefs about sexual misconduct, irrespective of the specific misconduct situation they may be assessing.

Limitations and Future Directions

The thematic analyses conducted of participants' open-ended explanations of the blameworthiness and responsibility ratings they gave for the perpetrator and victim not only provide insight into participants' decision-making process, but also provide evidence of where future studies can improve and expand on the current work. For example, many participants indicated that there was simply not enough proof to place blame on either the perpetrator or the victim for their actions. While this did not necessarily translate into a lower likelihood of

deciding to pursue legal action against the defendant (about 70% of participants said they would pursue legal action against either Marcus or Lauren), it does indicate that future studies can explore how altering the evidence within these harassment scenarios (to include more witnesses to the harassment or more extreme forms of harassment perhaps) can alter participants' judgments. Future studies should also assess participants' perceptions of the blameworthiness of the harasser's employer and the harasser, to more cleanly disaggregate the blame participants may hold for both parties.

REFERENCES

- ALBONETTI, C. A. (1986). CRIMINALITY, PROSECUTORIAL SCREENING, AND UNCERTAINTY: TOWARD A THEORY OF DISCRETIONARY DECISION MAKING IN FELONY CASE PROCESSINGS. *Criminology*, 24(4), 623–644.
<https://doi.org/10.1111/j.1745-9125.1986.tb01505.x>
- Bagozzi, R. P., & Moore, D. J. (1994). Public Service Advertisements: Emotions and Empathy Guide Prosocial Behavior. *Journal of Marketing*, 58(1), 56–70.
<https://doi.org/10.1177/002224299405800105>
- Balogh, D. W., Kite, M. E., Pickel, K. L., Canel, D., & Schroeder, J. (2003). The Effects of Delayed Report and Motive for Reporting on Perceptions of Sexual Harassment. In *Sex Roles* (Vol. 48).
- Barger, L. (2018). *What Came Before #MeToo: The Himpathy That Shaped Misogyny*. Yes! Magazine. <https://www.yesmagazine.org/issue/affordable-housing/2018/07/10/what-came-before-me-too-the-himpathy-that-shaped-misogyny/>
- Benedict, J., & Klein, A. (1997). Arrest and conviction rates for athletes accused of sexual assault. *Sociology of Sport Journal*, 14(1), 86–94. <https://doi.org/10.1123/ssj.14.1.86>
- Bishop, D. M., & Frazier, C. E. (1984). The Effects of Gender on Charge Reduction. *Sociological Quarterly*, 25(3), 385–396. <https://doi.org/10.1111/j.1533-8525.1984.tb00198.x>
- Blumenthal, J. A. (1998). The reasonable woman standard: A meta-analytic review of gender differences in perceptions of sexual harassment. *Law and Human Behavior*, 22(1), 33–57.
<https://doi.org/10.1023/A:1025724721559>
- Cox, J., & Kopkin, M. R. (2016). Defendant and Victim Sex, Sexism, and Decision Making in an

- Ambiguous Assault Case. *Women and Criminal Justice*, 26(5), 381–393.
<https://doi.org/10.1080/08974454.2016.1167153>
- Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2014). Empathy: A review of the concept. *Emotion Review*, 8(2), 144–153. <https://doi.org/10.1177/1754073914558466>
- Curry, T. R., Lee, G., & Rodriguez, S. F. (2004). Does victim gender increase sentence severity? Further explorations of gender dynamics and sentencing outcomes. *Crime and Delinquency*, 50(3), 319–343. <https://doi.org/10.1177/0011128703256265>
- Dodson, S., Goodwin, R., Chambers, M. K., Graham, J., & Diekmann, K. (2020). Moral Foundations, Empathy, and Attitudes Toward Sexual Misconduct Claims. *Academy of Management Proceedings*. <https://doi.org/10.5465/ambpp.2020.14553abstract>
- Eisenberg, N., & Strayer, J. (1988). Empathy and Its Development. Edited by N. Eisenberg and J. Strayer. (Pp. 406; illustrated; £32.50). Cambridge University Press: Cambridge, 1988.
Psychological Medicine, 18(3), 776–776. <https://doi.org/10.1017/s0033291700008667>
- Ford, T., Liwag-McLamb, M., & Foley, L. A. (1998). Perceptions of Rape Based on Sex & Sexual Orientation of Victim. *Journal of Social Behavior and Personality*, 13(2), 253–263.
- Glaeser, E. L., & Sacerdote, B. (2003). Sentencing in Homicide Cases and the Role of Vengeance. In *Journal of Legal Studies* (Vol. 32, Issue 2, pp. 363–382).
<https://doi.org/10.1086/374707>
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism. In *Journal of Personality and Social Psychology* Copyright (Vol. 70, Issue 3). American Psychological Association, Inc.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology*, 101(2), 366–385.

<https://doi.org/10.1037/a0021847>

Hoffman, M. (2008). Empathy and Prosocial Behavior. In *Handbook of emotions* (pp. 440–455).

<https://doi.org/10.1016/B978-012370877-9.00056-6>

Holcomb, J. E., Williams, M. R., & Demuth, S. (2004). White female victims and death penalty disparity research. *Justice Quarterly*, *21*(4), 877–902.

<https://doi.org/10.1080/07418820400096021>

Hurt, L. E., Wiener, R. L., Russell, B. L., & Mannen, R. K. (1999). Gender differences in evaluating social-sexual conduct in the workplace. *Behavioral Sciences & the Law*, *17*(4), 413–433. [https://doi.org/10.1002/\(SICI\)1099-0798\(199910/12\)17:4<413::AID-](https://doi.org/10.1002/(SICI)1099-0798(199910/12)17:4<413::AID-BSL364>3.0.CO;2-B)

[BSL364>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1099-0798(199910/12)17:4<413::AID-BSL364>3.0.CO;2-B)

Illing, S. (2018). *Brett Kavanaugh hearing: why powerful men get more sympathy than women*.

Vox. <https://www.vox.com/2018/9/27/17887210/brett-kavanaugh-christine-blasey-ford-hearing-kate-manne>

Karniol, R., Gabay, R., Ochion, Y., & Harari, Y. (1998). Is gender or gender-role orientation a better predictor of empathy in adolescence? *Sex Roles*, *39*(1–2), 45–59.

<https://doi.org/10.1023/a:1018825732154>

Kelley, S. M., Fleming, J. C., Acquaviva, B. L., Meeker, K. A., & O'neal, E. N. (2021). The Sexual Stratification Hypothesis and Prosecuting Sexual Assault: Is the Decision to File Charges Influenced by the Victim-Suspect Racial-Ethnic Dyad? *Crime & Delinquency*, *67*(8), 1165–1194. <https://doi.org/10.1177/0011128721991821>

Kessler, A. M., Kennair, L. E. O., Grøntvedt, T. V., Bjørkheim, I., Drejer, I., & Bendixen, M. (2020). The Effect of Prototypical #MeToo Features on the Perception of Social-Sexual Behavior as Sexual Harassment. *Sexuality and Culture*, *24*(5), 1271–1291.

<https://doi.org/10.1007/s12119-019-09675-7>

Lucarini, A., Suitner, C., Brown, R., Craig, M. A., Knowles, E. D., & Salvador Casara, B. G. (2020). The #MeTooLate Effect: Victim blame and trust denial for sexual harassment not immediately reported. *Personality and Individual Differences*, 167, 110240.

<https://doi.org/10.1016/j.paid.2020.110240>

Madan, M., & Nalla, M. K. (2015). Sexual Harassment in Public Spaces: Examining Gender Differences in Perceived Seriousness and Victimization. *International Criminal Justice Review*, 26(2), 80–97. <https://doi.org/10.1177/1057567716639093>

Manne, K. (2017). Down girl: The logic of misogyny. In *Down Girl: The Logic of Misogyny*. <https://doi.org/10.1093/oso/9780190604981.001.0001>

Manne, K. (2018). *Brett Kavanaugh and America's 'Himpathy' Reckoning*. The New York Times. <https://www.nytimes.com/2018/09/26/opinion/brett-kavanaugh-hearing-himpathy.html>

O'Neal, E. N., Beckman, L. O., & Spohn, C. (2019). The Sexual Stratification Hypothesis: Is the Decision to Arrest Influenced by the Victim/Suspect Racial/Ethnic Dyad? *Journal of Interpersonal Violence*, 34(6), 1287–1310. <https://doi.org/10.1177/0886260516651093>

Posick, C., Rocque, M., & Rafter, N. (2014). More than a feeling: Integrating empathy into the study of lawmaking, lawbreaking, and reactions to lawbreaking. *International Journal of Offender Therapy and Comparative Criminology*, 58(1), 5–26.

<https://doi.org/10.1177/0306624X12465411>

Pozzulo, J. D., Dempsey, J., Maeder, E., & Allen, L. (2010). The Effects of Victim Gender, Defendant Gender, and Defendant Age on Juror Decision Making. *Criminal Justice and Behavior*, 37(1), 47–63. <https://doi.org/10.1177/0093854809344173>

- Quas, J. A., Bottoms, B. L., Haegerich, T. M., & Nysse-Carris, K. L. (2002). Effects of victim, defendant, and juror gender on decisions in child sexual assault cases. *Journal of Applied Social Psychology, 32*(10), 1993–2021. <https://doi.org/10.1111/j.1559-1816.2002.tb02061.x>
- Riess, H. (2017). The Science of Empathy. *Journal of Patient Experience, 4*(2), 74–77. <https://doi.org/10.1177/2374373517699267>
- Rotundo, M., Nguyen, D. H., & Sackett, P. R. (2001). A meta-analytic review of gender differences in perceptions of sexual harassment. In *Journal of Applied Psychology* (Vol. 86, Issue 5, pp. 914–922). <https://doi.org/10.1037/0021-9010.86.5.914>
- Royer, C., Hritz, A., Hans, V., Eisenberg, T., Wells, M., Blume, J., & Johnson, S. (2014). Victim gender and the death penalty. *Umkc, 82*(429). https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/umkc82§ion=21
- Russell, B. L., & Trigg, K. Y. (2004). Tolerance of sexual harassment: An examination of gender differences, ambivalent sexism, social dominance, and gender roles. *Sex Roles, 50*(7–8), 565–573. <https://doi.org/10.1023/b:sers.0000023075.32252.fd>
- SPOHN, C., GRUHL, J., & WELCH, S. (1987). THE IMPACT OF THE ETHNICITY AND GENDER OF DEFENDANTS ON THE DECISION TO REJECT OR DISMISS FELONY CHARGES. *Criminology, 25*(1), 175–192. <https://doi.org/10.1111/j.1745-9125.1987.tb00794.x>
- Spohn, C., & Holleran, D. (2001). Prosecuting sexual assault: A comparison of charging decisions in sexual assault cases involving strangers, acquaintances, and intimate partners. *Justice Quarterly, 18*(3), 652–688. <https://doi.org/10.1080/07418820100095051>
- Spohn, C., & Spears, J. (1996). The effect of offender and victim characteristics on sexual assault case processing decisions. *Justice Quarterly, 13*(4), 649–679.

<https://doi.org/10.1080/07418829600093141>

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment, 91*(1), 62–71.

<https://doi.org/10.1080/00223890802484381>

Toussaint, L., & Webb, J. (2005). Gender Differences in the Relationship Between Rumination and Forgiveness. *The Journal of Social Psychology, 145*(6), 673–685.

<https://doi.org/10.16719/j.cnki.1671-6981.2012.05.031>

Vicenta Mestre, M., Samper, P., Dolores Frías, M., & María Tur, A. (2009). Are Women More Empathetic than Men? A Longitudinal Study in Adolescence. *The Spanish Journal of Psychology Copyright, 12*(1), 76–83. <http://roderic.uv.es/handle/10550/2341>

Vitaglione, G. D., & Barnett, M. A. (2003). Assessing a New Dimension of Empathy: Empathic Anger as a Predictor of Helping and Punishing Desires. *Motivation and Emotion, 27*(4), 301–325. <https://doi.org/10.1023/A:1026231622102>

von Sikorski, C., & Saumer, M. (2020). Sexual Harassment in Politics. News about Victims' Delayed Sexual Harassment Accusations and Effects on Victim Blaming: A Mediation Model. *Mass Communication and Society, 1–29*.

<https://doi.org/10.1080/15205436.2020.1769136>

Walfield, S. M. (2021). “Men Cannot Be Raped”: Correlates of Male Rape Myth Acceptance. *Journal of Interpersonal Violence, 36*, 13–14. <https://doi.org/10.1177/0886260518817777>

WALSH, A. (1987). THE SEXUAL STRATIFICATION HYPOTHESIS AND SEXUAL ASSAULT IN LIGHT OF THE CHANGING CONCEPTIONS OF RACE. *Criminology, 25*(1), 153–174. <https://doi.org/10.1111/j.1745-9125.1987.tb00793.x>

Wiener, R. L., & Hurt, L. E. (2000). How do people evaluate social sexual conduct at work? A psycholegal model. *Journal of Applied Psychology*, 85(1), 75–85.

<https://doi.org/10.1037/0021-9010.85.1.75>

Worthington, E. L. J., Hook, J. N., Utsey, S. O., & Williams, J. (2007). Decisional and emotional forgiveness: Conceptualization and development of self-report measures.pdf. *Positive Psychology Summit*, 1–39.

Zeilkova, J., Ramirez, K., & Lipps, J. (2018). Harrassment charges: Enough Himpathy. *Science*, 361(6403), 655–655. <https://repositorio.comillas.edu/rest/bitstreams/190385/retrieve>

CHAPTER 3:

WOMEN ON THE JURY: THE ROLE OF GENDER IN CIVIL JURIES

Introduction

Gender plays an important role within our justice system: shaping perceptions of attorneys, expert witnesses, victims, and defendants, and influencing jury decision making dynamics and ultimate verdicts. Across civil and criminal systems, gender has been shown to play an important role in a jury's deliberation dynamics. Significant differences have been identified in male versus female jurors' rates of participation, the quality and content of their participation within the deliberation, and the degree of interruptions that may exist between jurors. Individual-level dynamics such as the expression – or lack thereof – of minority opinions or the expression of emotion, and group-level decision-making dynamics - conformity and group think effects, interruptions between jurors - have been shown to undergird these gender differences in deliberation. However, might there be elements to deliberations in civil cases involving compensatory damage awards (e.g. personal injury or other tort cases) that engender unique differences in how male versus female jurors behave? Furthermore, beyond the deliberation room, might juror gender matter in how jurors individually assess (either pre or post deliberation) defendants' liability and determining damages? This chapter will present a combination of qualitative and quantitative data to address these questions.

Gender Differences in Jury Deliberation Behavior, Across Civil and Criminal Systems

Across both the criminal and civil systems, a wide body of metanalytic and experimental work has identified systematic differences in how male versus female jurors act during a deliberation. For example, early mock jury deliberation work by Strodbeck & Mann in 1956 found that men tended to “pro-act” more than women, by initiating discussion aimed at solving a

problem or task within the deliberation, while women tended to react to others' commentary (Strodtbeck & Mann, 1956). Later work mock jury work by Nemeth, Endicott & Wachtler in 1976 showed that in response to a hypothetical murder trial presented to them, male and female jurors did not differ in their initial verdicts, but gender disparities became visible during deliberation. Men generally were the recipients of more communication than their female counterparts, and because of greater communication aimed at them, men in these experiments were also more likely to be interrupted (by both male and female jurors) than female jurors (Nemeth et al., 1976).

Despite Nemeth et al.'s 1976 work showing that men were more likely to be interrupted due underlying differences in how much men and women spoke during deliberation, other work has shown that men are also more likely to instigate interruptions against other jurors (Marder, 1987). Further, male jurors have been shown to make more statements overall, and within their communication, were more likely to offer ideas and suggestions based on the case-evidence than female jurors (Hastie et al., 1983; Marder, 1987; Strodtbeck & Mann, 1956)

In terms of more subtle exhibitions of dominance (i.e. beyond the content or frequency of their speech), male jurors have been shown to be more likely to sit at the head of the deliberating table, and therefore, be more likely to be selected as the foreperson of the jury (Hastie et al., 1983; Marder, 1987; Nemeth et al., 1976). Further, Nemeth et al.'s 1976 work also found that generally, male jurors were more likely to be perceived by their fellow jurors as being independent, strong, influential, and possessive of leadership qualities. In contrast, mock jury deliberation work by Salerno, Peter-Hagene, & Jay (2017) found that females who used anger in expressing their minority holdout opinion were subsequently perceived as being less effective or influential communicators, compared to their male counterparts (Salerno et al., 2017).

Juror Gender within Civil Juries

The aforementioned body of work illustrated a pattern of male dominance (and backlash against women who attempted to go against the norm) within deliberations, within both criminal and civil juries. However, may there be something unique to civil juries in tort cases involving liability and damage award decisions, that either amplifies, dampens, or alters these trends?

In terms of liability decisions, juror gender does not seem to be a robust or reliable predictor of liability decisions (Nagel & Weitzman, 1972; Vinson et al., 2008). Rather, case relevant factors such as the defendant's intentionality and plaintiff injuries seem to be more influential (Greene et al., 1999). A juror's belief in a litigation crisis - the notion that plaintiffs are too quick to sue, are just greedy for money, etc.- also influences verdict, such that heightened belief in such a crisis has been found to increase jurors' likelihood of finding for the defendant (Vinson et al., 2008). However, holding such anti-litigation attitudes has not been shown to differ by a juror's gender (Hans & Lofquist, 1994; Neubauer & Meinhold, 1994).

Regarding damage award decisions, the role of juror gender in these assessments has been mixed. Many studies show no impact of juror gender (Vidmar, 1993; Vidmar & Rice, 2008; Vinson et al., 2008), while older data from the Jury Verdict Research Corporation indicates that in male-dominated juries, male plaintiffs were awarded significantly above the average amount typically given for the case-type at hand, while female plaintiffs were given significantly below the average (Nagel & Weitzman, 1972). Thus, the gender composition of the civil jury may play a role in how damage awards are given out.

However beyond juror gender, what seems to be the most influential in shaping damage awards is simply the amount that may be requested by the plaintiff's, especially if that amount is grounded in trial evidence and does not appear extreme, in light of the evidence (Chapman &

Bornstein, 1996; Hastie et al., 1999). This aligns with the well-known phenomenon of anchoring in civil damage determinations.

Juror Gender and Lost Income Assessments within Damage Award Decision-Making

One area in which both juror and plaintiff gender may be especially pertinent is in calculating damages for a plaintiff's lost wages. That is, if jurors are calculating lost wages awards based on estimates of a plaintiff's income, female plaintiffs would arguably be disadvantaged, given that in 2020, men's median salary was approximately 19% higher than that of women (Barroso & Brown, 2021). Empirical work by Sherri Lamb (1996) and Goodman et al. (1991) supports this phenomena (Goodman et al., 1991; Lamb, 1996). Further, irrespective of the gender of the plaintiff, if a female juror is using her own income (which given the aforementioned pay gap, would presumably be lower than her male counterpart's) as an anchor by which to calculate damages for the plaintiffs, her estimates could be, on average, lower than that of her male counterparts.

Juror Gender and Numeracy During a Deliberation

Though an understudied area, prior work has shown the importance of numeracy, or one's general adeptness at dealing with numbers by themselves and in context, in damage award decision-making (Helm et al., 2017). For example, work by Helm, Hans, Reyna, and Reed (2020) indicated that jurors who were more numerate subsequently proposed damage awards that were less variable, as well as more commensurate with the plaintiff's pain and suffering (Helm et al., 2020). However, examining how numeracy may play out in a deliberation (i.e. if and how it is brought up in jurors' discussion of damages) has yet to be concretely explored. The current work seeks to fill in this research gap.

Current Study

The current study aimed to examine, through a mock jury paradigm involving both qualitative and quantitative data, the role of juror gender in civil jury deliberations, and decisions of liability and damage award decision-making.

Regarding deliberation behavior, it was postulated that male jurors would instigate a significantly higher number of interruptions compared to women, and that those interruptions would tend to be against women, rather than men. This would align with aforementioned trends in interruption rates by juror gender, exhibited in both criminal and civil juries (Marder, 1987). We additionally hypothesized that male jurors would be more likely to propose the first damage award number during the deliberation. This gender effect would be in alignment with prior work showing that men were more likely to offer ideas and suggestions more than their female counterparts in a deliberation (Hastie et al., 1983; Marder, 1987; Strodtbeck & Mann, 1956), and that in the non-legal context of economic negotiations, men are more likely to initiate the negotiation with an aggressive offer (Mazei et al., 2015). Related to damage decisions within the deliberation, we hypothesized that the notion of gender differences in numeracy will be encapsulated by male jurors being more likely to be selected as the “number person” in a deliberation (the juror elected to make all calculations), as well as make a greater number of statements regarding their comfort with math, compared to their female counterparts.

Outside of the deliberation, we hypothesized that there would be no gender differences for jurors’ liability determinations. This hypothesis is shaped by previous work showing that other factors outside of juror gender seem to be more influential in shaping liability decisions (e.g. Greene et al., 1999; Vinson et al., 2008). Regarding damage awards, we postulated that influenced by their own possible lower income, female jurors will provide lower damage award

amounts (especially for lost wages) to the plaintiffs as compared to their male counterparts.

Methodology

Participants

Participants were recruited as part of a larger body of work being conducted to examine how civil jurors decide on liability and damage-awards in a personal injury case. This work is funded by the National Science Foundation. Participants were recruited for an in-person jury deliberation study held at Cornell Law School from October 2019 to March 2020. Upon the onset of the global coronavirus pandemic, this study transitioned to a virtual format. Data from the in-person iteration of this study are presented in the current work, while data collection for the virtual iteration of this study is currently ongoing and not presented in the current manuscript.

Approximately 142 participants, across twenty-four total juries, completed all tasks associated with the study. Per our inclusion criteria, all participants were above eighteen years old, were United States citizens, and were community members from the Ithaca area or surrounding areas (e.g. Syracuse, Rochester, Binghamton, etc.). Participants were recruited via newspaper advertisements in the Ithaca Times, local listservs such as the Tompkins County listserv, word of mouth, flyer advertisements posted throughout the Ithaca area, and online advertisements on local social media pages.

Participants' ages ranged from 18 to 87 years old ($M_{\text{age}} = 51.12$ years old, $SD = 20.42$ years). The sample included 56 men (39.44% of the total sample), 84 women (59.15% of the total sample). Two participants did not provide their gender (1.41% of the total sample). The sample was predominantly White (85.21% of the total sample), 6.34% Black, 2.11% Chinese, 0.70% Asian Indian, 0.70% Mixed Race, and 4.93% of participants indicated some other race (e.g. White, Caucasian & Middle Eastern, or the participant preferred not to answer).

Additionally, 94.37% of the sample was not of Hispanic, Latino, or Spanish origin, while 2.11% of the sample was of Mexican, Mexican American, or Chicano origin, 0.70% was Puerto Rican, 0.70% was Cuban, and 2.11% indicated another Hispanic, Latino, or Spanish origin.

Out of the 142 total participants, 24 total juries were crafted; 22 of those juries contained six jurors each, and two out of the 24 juries were made up of five jurors.

Procedure

Participants arrived at Cornell Law School and were brought to a mock jury room by a set of research assistants. They completed an informed consent form, and upon consenting to participate in the study, were given a questionnaire to fill out with various demographic questions – a form akin to voir dire, or jury selection. However, unless a participant did not fit our inclusion criterion or did not consent to participate, no participants were eliminated based on their answers to this demographic, voir-dire equivalent questionnaire.

Upon completing these forms, jurors were randomly assigned to a six-person jury. Participants were not placed in a jury together if they were related (i.e. husband and wife, siblings, etc.), or if they otherwise indicated that they knew one another (i.e. neighbors, friends, etc.). We consistently aimed to develop juries made up of six individuals. If there were fewer or more than six jurors in attendance, these individuals were considered alternate jurors, and would complete the same study procedure, but on an individual basis and without deliberating with other jurors as a group. Juries were all crafted with the goal of gender equity (an equal mix of men and women in the jury).

After random assignment, each jury was escorted into another mock jury room by a research assistant. These rooms were set up with a number placard (from one to six) at each seat of a table, and participants were instructed to sit anywhere they liked. All proceedings within this

mock jury room were recorded for data analysis purposes via a livestream recording and a tripod video camera.

Once seated, the jurors watched a one-hour long, filmed mock trial video. Participants were permitted to take hand-written notes while watching. The trial video depicted a hypothetical civil lawsuit case involving a traffic accident, in which the plaintiffs, Susan Kessler and Ashley Franklin, were suing the defendant, Jeffrey Burnside, for damages for the injuries they incurred as a result of the accident. Susan Kessler's injuries were more severe, and lasted longer, compared to those of Ashley Franklin.

Each jury was randomly assigned to one of three conditions, in which the content of the opening and closing statements (namely, the quality and quantity of guidance the plaintiff's attorney provided to the jury, in regard to how to determine the defendant's liability and decide on damage awards), was varied. These three conditions included a *control*, *gist-only*, and *gist plus numeric* condition.

Specifically, in the *control condition* [$N = 7$ juries; $N = 42$ jurors], the plaintiff's attorney provided very minimal guidance, stating to jurors:

"I ask you to find Mr. Burnside and Macklin Furniture Company liable and compensate these victims. Although the medical bills have already been paid and are not at issue in this case, my clients need to be compensated for their lost wages and pain and suffering. Members of the jury, let me advise you about the calculation of an amount for pain and suffering. No one can tell you how to arrive at an amount of money for pain and suffering. I cannot tell you. The judge cannot. No one can. You just have to use your own good common sense and your best judgment."

In the *gist-only condition* [$N = 7$ juries; $N = 40$ jurors], the plaintiffs' attorney provided jurors with a general sense of the length and severity of each of the plaintiff's injuries, as well as a general framework to use in calculating their pain and suffering damage awards (to imagine what one hour of pain and suffering might be worth to each plaintiff). Specifically, the attorney

stated:

“Members of the jury, let me advise you about determining an amount for pain and suffering. Please think about how much 1 hour of pain and suffering is worth to each of my clients. Remember, Ashley lost four months of her life on brain rest, and took two years to recover. Ashley’s injuries were bad, but Susan’s were of a completely different magnitude. In addition to her physical injuries, Susan suffered a loss so tremendous that she’s unable to work or engage in the most basic social interactions. She can’t even remember her kids’ ages. By her own admission, Susan is not normal. What is that type of loss worth for an hour? Susan will suffer this loss every hour of every day, day after day, week after week, month after month, year after year. For the rest of her life. Please, use your best sense to compensate the plaintiffs: Ashley Franklin for her two years of pain and suffering, and Susan Kessler, whose pain and suffering began two years ago, and will last the rest of her life. Thank you.”

In the *gist plus numeric condition* [$N = 10$ juries; $N = 60$ jurors,], the plaintiff’s attorney replicated the information in the gist-only condition. The attorney also added specific dollar amounts for what the metric of an hour of pain and suffering might be worth (a minimum of \$10 per hour), as well as what specific dollar amounts that would amount to, using the hour metric, for each plaintiff (a minimum of \$175,200 for the less severely injured plaintiff Ashley Franklin, and a minimum of \$4,204,800 for the more severely injured plaintiff Susan Kessler).

Specifically, the attorney stated:

“Members of the jury, let me advise you about determining an amount for pain and suffering. Please think about how much one hour of pain and suffering is worth to each of my clients. Remember, Ashley lost four months of her life on brain rest, and took two years to recover. Ashley’s injuries were bad, but Susan’s were of a completely different magnitude. In addition to her physical injuries, Susan suffered a loss so tremendous that she’s unable to work or engage in the most basic social interactions. She can’t even remember her kids’ ages. By her own admission, Susan is not normal. What is that type of loss worth for an hour? Susan will suffer this loss every hour of every day, day after day, week after week, month after month, year after year. For the rest of her life. Please, use your best sense to compensate the plaintiffs: Ashley Franklin for her two years of pain and suffering, and Susan Kessler, whose pain and suffering began two years ago, and will last the rest of her life. Surely, this suffering is at least ten dollars an hour minimum. Therefore we request that you reward Ashley Franklin a minimum of one hundred seventy five thousand two hundred dollars, and Susan Kessler a minimum of four million two hundred four thousand and eight hundred dollars. Thank you.”

Upon watching the trial video, the jurors completed an individual verdict questionnaire, in which they were asked to provide their verdict (liability decisions) and other case-related judgments (delineated further in the *Materials* section). Jurors were permitted to use their notes as a basic phone calculator for this questionnaire, if needed. Next, participants took a 15-minute break and then came back to the mock jury room, having been instructed to return to their same seats as before, for deliberation. The juror at the seat labeled “Juror 1” was announced to be the randomly selected foreperson and was instructed to complete the group verdict form (details of which are delineated in the *Materials* section). Jurors were instructed to deliberate until they reached a unanimous verdict (no hung juries were allowed). Upon completion of their deliberation, jurors were asked to complete a final, post-deliberation individual questionnaire, regarding their case perceptions (e.g. their degree of agreement with their jury’s group decisions, their assessment of the case facts, etc.) and a variety of individual difference measures (e.g. their views on lawsuits, concussions, their objective and subjective numeracy, etc.). Jurors were instructed that they could not have their notes nor any type of calculator out for this questionnaire. After completion of this final survey, jurors were escorted to a debriefing room, where they were thanked for their participation, compensated (via a \$50 check), and debriefed (i.e. the intent of the study was revealed to them, and participants expressed any questions or comments they had about their study experience).

Materials

A) Individual Verdict Questionnaire

After watching the trial video, all jurors were asked to complete an individual verdict questionnaire. Specifically, this questionnaire asked them to provide a) their liability decisions for the defendant Jeffrey Burnside and Macklin Furniture Company, the company he was

employed by at the time of the accident, the specific monetary damage amount they would award Susan Kessler and Ashley Franklin for their a) lost income, and b) pain and suffering, c) their categorization of their proposed damage award as being either *nil, low, medium, or high*, and d) their confidence in their proposed damage awards, from 1 (Not at All Confident) to 7 (Extremely Confident).

B) Group Verdict Form

Upon completing their deliberation, the foreperson of the jury (always the participant who chose to sit at the seat for Juror #1) was instructed to complete the group verdict form, based on the jury's decisions. Namely, in this form, the foreperson was asked to complete the jury's decisions on A) the liability of the Macklin Furniture Company, the employer of the defendant Jeffrey Burnside (with participants being instructed that they were finished with the form if they selected "No" for this question), B) liability of the defendant, Jeffrey Burnside, and their damage award decisions for the two plaintiffs – Susan Kessler and Ashley Franklin, subdivided into a C) lost wages damage award and D) pain and suffering damage award awarded to each plaintiff.

C) Post-Deliberation Questionnaire

Following their deliberation, jurors were given a final questionnaire which included a wide variety of case-related judgments and individual difference measures such as subjective & objective numeracy, a juror's approximate individual and household yearly income range, a juror's views towards lawsuits, their degree of agreement with their jury's verdict and proposed damage award, and their belief in a civil litigation crisis.

Results

Qualitative Coding Schema

The content of deliberations from all 24 juries were transcribed by a team of trained research assistants. Interruptions between jurors were also denoted and quantified. Transcripts were prepared such that all jurors were only identified by their juror number (i.e. any information about their gender was removed), and were further de-identified by renaming Jury numbers into their alphabet equivalent (i.e. Jury 1 became Jury A).

In order to qualitatively code the deliberations, an extensive coding schema, composed of approximately 56 variables, was constructed. This coding schema was devised based on both a-priori variables of interest, as well as data-driven variables of interest that emerged during data collection. The schema includes, for example, codes related to a) both plaintiffs' damage awards for pain and suffering and lost wages, identifying jurors who proposed the first pain and suffering damage award number during deliberation, identifying jurors who mentioned the damage award anchor amounts provided by the plaintiff's attorney, b) commentary on the plaintiffs' injuries, such as the extent to which the accident / each plaintiff's injuries affected their social life, family life, and their spouses, c) comparing the two plaintiff's damage awards or injuries, d) discussion around the defendant and/or furniture company's liability, or e) identifying jurors who were the designated "number person", in charge of making calculations for the jury. A detailed list of all variables included in this coding schema, with example quotes from participants, can be found in Appendix B.

A team of four undergraduate research assistants were trained in applying this coding schema to all 24 deliberation transcripts using the qualitative software NVivo 12 Plus. All coders were blind to the condition (control, gist-only, or gist-plus numeric) that each jury was exposed to, as well as to all demographic features of the jurors.

One of the juries within the sample, Jury B, was coded by all four coders. The coding of

the resulting six pairs of coders was compared, and the pair of coders with the highest initial interrater reliability was selected for further analysis. Disagreements in coding between these two coders were resolved through discussion. The final interrater reliability between two of the four coders for Jury B was high: Cohen's kappa coefficient, $M = 0.97$, Range: 0.71 (Ashley Franklin Award Numbers → Ashley Franklin Lost Wages Anchor) – 1.0 (Susan Kessler Commentary → Susan Kessler Spouse). Exact Cohen's Kappa coefficient scores and Agreement percentages between the two coders for all variables can be seen in Appendix C.

Frequency of Qualitative Variables

The frequency breakdown of each code (by how many juries it was identified within, as well as how many times each code was used across the 24 juries) is displayed in Figure 3.1 and 3.2, respectively. A total of 53 qualitative variables were assessed. Three variables - Length Defendant, Length Kessler, and Length Franklin were omitted from analyses – as they were not ultimately included in the qualitative coding. As indicated by Figure 3.1, the notion of either plaintiff being responsible for their injuries (i.e. statements alluding to the fact that they should have known better, or they that they were somehow also in the wrong) only arose in two juries, while statements regarding the defendant's responsibility, commentary on the plaintiffs' damage awards, the nature of Susan Kessler's injuries (how bad, long, and interfering they were), and descriptions of juror's own personal experiences informing their opinions of the case arose in all 24 juries.

Figure 3.1: Coverage of Qualitative Variables by Number of Juries

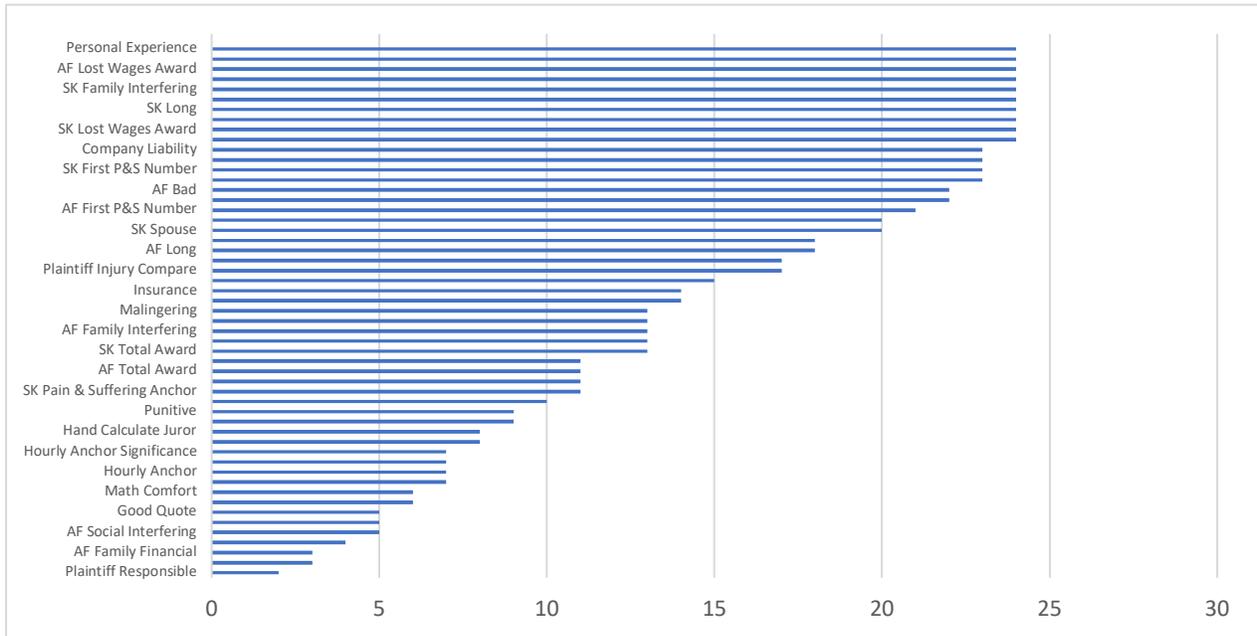
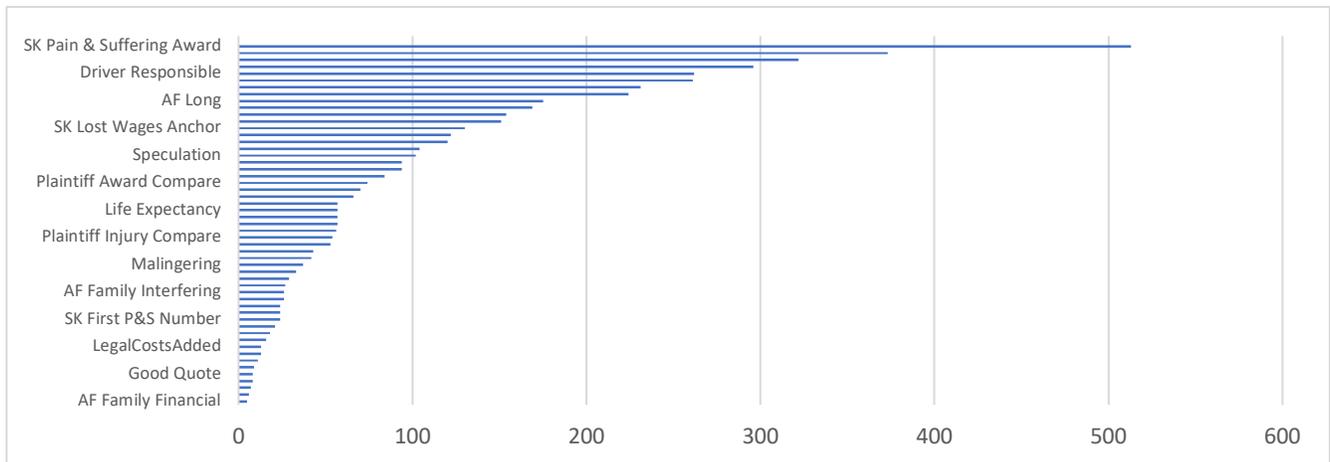


Figure 3.2: Coverage of Qualitative Variables by Number of Instances Coded

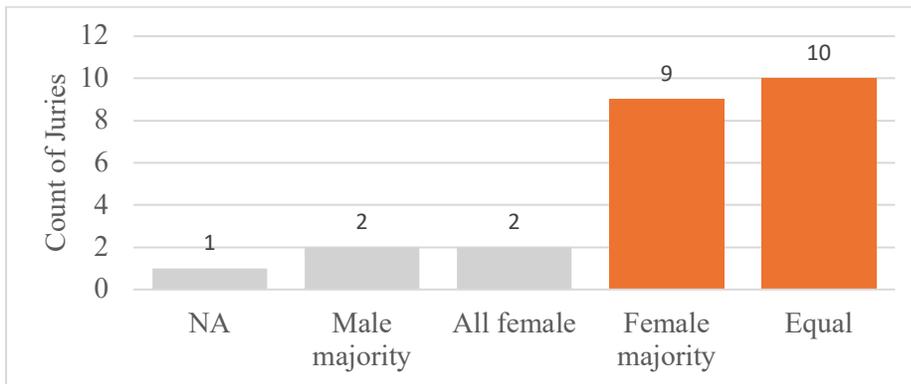


As shown in Figure 3.2, across all juries, statements regarding Susan Kessler’s Pain and Suffering Award arose most frequently (513 instances coded), while statements regarding the impact that Ashley Franklin’s injuries would have on her family’s financial well-being were only coded five times.

Gender Breakdown of Deliberating Juries

Across the 24 deliberating juries, two juries were composed of all females, ten juries were composed of an equal number of females and males (three females and three males), nine juries were composed of a female-majority (three or more jurors in a five-person jury, four or more jurors in a six-person jury), two juries included a male majority (four or more jurors in a six-person jury), and one jury could not be clearly categorized for its gender composition. A visual representation of this gender categorization is in Figure 3.3. However, due to the small and uneven cell sizes of this gender categorization, analyses were conducted comparing the deliberation behavior of male versus female jurors at large, irrespective of the gender makeup of the jury.

Figure 3.3. Gender Categorization of Deliberating Mock Juries



Rates of Interruptions, by Juror Gender and Condition

Across all 24 deliberating juries, the instance of an interruption between one or more deliberating jurors was identified 5,217 total times ($M_{\text{interruption}} = 7.25$ interruptions instigated per juror, $SD = 8.16$), with a minimum of 0 interruptions instigated by a juror and a maximum of 71 interruptions instigated by one juror. Disaggregating interruptions by jury, there were an average of 217.38 interruptions per jury ($SD = 118.44$, $Min = 37$, $Max = 533$).

Examining interruptions by juror gender across 24 juries more closely, an independent

samples t-test did not reveal any significant differences in the average number of interruptions instigated by male (M = 6.98, SD = 8.63) versus female jurors (M = 7.74, SD = 7.93) during a deliberation; $t(557.48) = -1.17, p = 0.24$.

A one-way ANOVA did not find any significant difference in the average number of interruptions generated by condition {control (M = 155.29, SD = 97.37), gist-only (M = 222, SD = 77.75), gist plus numeric (M = 257.6, SD = 143.83)}; $F(2, 21) = 1.63, p = 0.22$.

Juror Gender and Award Calculation-Related Codes

The rates by which male versus female jurors proposed the first pain & suffering award number (First P&S Number code), mentioned comfort levels with math (Math Comfort code), or spearheaded calculations of a damage award during deliberation (Number Person code) were calculated. The resulting frequencies, and the proportions of males and females who mentioned each code, are summarized in Table 3.1.

Table 3.1 Frequencies and Proportions of Personal Experience, First Pain & Suffering Number, Math Comfort, and Number Person Codes, by Juror Gender

Qualitative Code	Males (N = 53)	Females (N = 81)
Personal Experience	20 (37.73%)	44 (54.32%)
Ashley Franklin First P&S Number	10 (18.87%)	11 (13.58%)
Susan Kessler First P&S Number	9 (16.98%)	13 (16.05%)
Math Comfort	0 (0%)	5 (6.17%)
Number Person Juror	1 (1.89%)	5 (6.17%)

A series of chi-square tests of independence and Fisher’s exact tests (due to some cell counts being less than 5) were run to calculate whether the proportion of females versus males

who made mention of the aforementioned codes significantly differed from one another. Results revealed that the relationship between juror gender and mentions of personal experience approached significance $\{X^2(1, N = 134) = 3.53, p = 0.06\}$, while the relationship between juror gender and mentioning and first pain and suffering damage award for Ashley Franklin $\{X^2(1, N = 134) = 0.68, p = 0.41\}$ or for Susan Kessler $\{X^2(1, N = 134) = 0.02, p = 0.89\}$ did not reach significance. Additionally, juror gender was not significantly associated with either mentions of being comfortable with math $\{\text{Fisher's exact test}, p = 0.16\}$ or mentions of being the designated “number person” to make calculations during the deliberation $\{\text{Fisher's exact test}, p = 0.41\}$.

An additional comparison was run to examine the frequency of mentions of the aforementioned codes (numeric-related codes and personal experience) by gender, and within each experimental condition. A summary of these frequencies, and the proportion of males and females who mentioned each code, can be found in Table 3.2. Given the low cell counts of each of these codes, these frequencies were not further disaggregated across the 24 juries.

Table 3.2 Frequencies and Proportions of Personal Experience, First Pain & Suffering Number, Math Comfort, and Number Person Codes, by Juror Gender and Condition

Qualitative Code	Condition = 1, Control (N = 41)		Condition = 2, Gist (N = 41)		Condition = 3, Gist Plus Numeric (N = 54)	
	Males (N = 18)	Females (N = 23)	Males (N = 13)	Females (N = 27)	Males (N = 22)	Females (N = 31)
Personal Experience	5 (27.8%)	10 (43.5%)	6 (46.2%)	21 (77.8%)	9 (40.9%)	13 (41.9%)
Math Comfort	0 (0%)	0 (0%)	0 (0%)	2 (7.4%)	0 (0%)	3 (9.7%)
Number Person Juror	0 (0%)	1 (4.3%)	0 (0%)	1 (3.7%)	1 (4.5%)	3 (9.7%)
SK First P&S Number	3 (16.7%)	4 (17.4%)	1 (7.7%)	5 (18.5%)	5 (22.7%)	4 (12.9%)
AF First P&S Number	1 (5.6%)	5 (21.7%)	3 (23.1%)	4 (14.8%)	6 (27.3%)	2 (6.5%)

A series of chi-square tests of independence and Fisher's exact tests (due to some cell counts being less than 5) were run to calculate whether the proportion of females versus males who made mention of the aforementioned codes, within each condition, significantly differed from one another. Results revealed that there was a non-significant association between a juror's gender and their mention of any of the above codes (i.e. personal experience, proposing the first pain and suffering damage award number for Susan Kessler or Ashley Franklin, math comfort, or being the designated number person), with each experimental condition, all p s > 0.054 .

Relationship Between Juror Gender, Condition, and Defendant Liability

Across all juries, jurors overwhelmingly found the defendant liable $\{N_{\text{liability}} = 136$ (95.8%), $N_{\text{notliable}} = 6$ (4.2%)}. However, a binary logistic regression (with decisions of liability coded as 1, and not liable coded as 0) indicated that juror gender did not significantly predict the odds of jurors' decision of liability ($\beta = -0.30$, $SE = 0.88$, Wald (1) χ^2 , 0.12, $p = 0.73$; Odds Ratio = 0.74).

The Role of Juror Gender and Condition on Damage Award Confidence

Participants provided separate damage award estimates for Susan Kessler and Ashley Franklin's lost wages and their pain and suffering. Jurors who did not clearly indicate their numeric damage awards in their responses (i.e. they stated that they did not know how to calculate an amount, provided a formula for calculating an amount but did not provide the final amount, provided a verbal explanation for how they would calculate the damages but did not provide an actual numeric amount) were excluded from the analyses. With these exclusions in place, the descriptive statistics (number of awards being assessed, minimum and maximum

award amount, average and median damage award, standard deviation) for each type of award for each plaintiff can be viewed in Table 3.2.

Table 3.2: Descriptive Statistics of Lost Wages & Pain & Suffering Damage Awards

	N	Minimum	Maximum	Mean	Median	Std. Deviation
Susan Kessler Lost Wages	126	\$100	\$1,000,000,000	\$10,083,379.37	\$515,000	89,886,359.30
Susan Kessler Pain & Suffering	132	\$0	\$1,000,000,000	\$9,547,698.48	\$1,000,000	86,947,469.21
Ashley Franklin Lost Wages	130	\$0	\$1,000,000,000	\$7,846,647.69	\$100,000	87,693,040.77
Ashley Franklin Pain & Suffering	127	\$0	\$1,000,000,000	\$8,222,219.69	\$100,000	88,722,686.09

Jurors' average confidence (assessed on a 1-7 scale) in the awards they proposed for Susan Kessler was 5.06 (SD = 1.55), compared to 4.86 for Ashley Franklin's awards (SD = 1.4).

A MANOVA was conducted to assess the effect of condition and juror gender on award confidence (confidence measures for each plaintiff, entered as the two dependent variables).

However, jurors' award confidence did not significantly differ by either juror gender { $F(2, 132) = 0.42, p = 0.66, \text{Wilks' } \Lambda = 0.99, \eta_p^2 = 0.006$ } or condition { $F(4, 264) = 1.68, p = 0.15, \text{Wilks' } \Lambda = 0.95, \eta_p^2 = 0.025$ }.

The Role of Juror Gender and Condition on Proposed Damage Award Amounts

The distribution of all damage awards (for pain and suffering and lost wages, for each plaintiff) by juror gender and condition were first assessed for normality. However, an examination of different numerical and visual outputs (i.e. histograms, normal Q-Q plots and box plots, skewness and kurtosis z-values, and Shapiro Wilk-test p-values) indicated that the distribution of all damage awards were non-normal: all skewness and kurtosis z-values were beyond +/- 1.96, all Shapiro Wilk test p-values were below 0.05, and an examination of

histograms and normal-QQ plots indicated violations of normality. Thus, for all subsequent analyses, participants' proposed damage awards were log-transformed.

With these transformations in place, a series of four mixed effect linear models were conducted to assess the effect of a juror's gender and the experimental condition they were exposed to on (natural log-transformed) damage awards proposed. Gender and experimental condition were entered as fixed effects, while jury was entered as a random effect. One model was constructed for each damage award proposed (entered as the dependent variable).

When assessing damage awards proposed for Susan Kessler's Lost Wages, Ashley Franklin's Lost Wages, and Ashley Franklin's Pain and Suffering, no main effects of either juror gender or condition were discovered (all p s > 0.14), and the interaction between juror gender and condition was non-significant (p s > 0.18). When examining awards proposed for Susan Kessler's Pain and Suffering, however, a significant main effect of condition emerged, $F(2,121) = 3.80$, $p = 0.025$. Follow-up pairwise comparisons using Fisher's LSD test indicated that jurors provided significantly higher (natural log-transformed) damage awards in the gist plus numeric Condition 3 ($M = 6.01$, $SE = 0.11$, 95% CI: [5.79 – 6.23]) compared to the control Condition 1 ($M = 5.52$, $SE = 0.14$, 95% CI: [5.24 – 5.79]). No other significant main effects (of juror gender) or interactions (between juror gender and condition) emerged.

To examine the effect of condition on natural log damage awards more closely, an additional one-way ANOVA was conducted. A main effect of condition on the natural log of damage awards proposed for Susan Kessler's Lost Wages ($F(2,121) = 3.32$, $p = 0.04$, $\eta^2 = 0.052$) emerged. Bonferroni post-hoc comparisons indicated that jurors provided higher (natural log-transformed) damage awards for Susan Kessler's Lost Wages in the gist plus numeric condition, Condition 3 ($M = 5.77$, $SE = 0.09$, 95% CI: [5.60 – 5.95]) compared to the control condition,

Condition 1 ($M = 5.36$, $SE = 0.16$, 95% CI: [5.04 – 5.69]).

Regarding the natural log of damage awards provided for Susan Kessler's Pain and Suffering, Levene's test for the homogeneity of variance revealed that the variance of these natural log damage awards significantly differed by condition (Levene's statistic = 4.05, $p = 0.02$; $SD_{\text{Condition1(control)}} = 1.15$, $SD_{\text{Condition2(gist-only)}} = 0.65$, $SD_{\text{Condition3(gist-plusnumeric)}} = 0.63$). As perhaps could have been expected given the limited amount of guidance given to jurors in the control condition, the variability of their (natural log-transformed) damage awards was greatest in this condition, compared to Conditions 2 and 3 which contained more guidance. Follow up analyses using the more robust, conservative Welch's ANOVA (used instead of the traditional ANOVA F-test due to the violation of the homogeneity of variance assumption) did not reveal any main effect of condition on natural log damage awards for Susan Kessler's Pain and Suffering; $F_{2,67} = 2.65$, $p = 0.08$.

Discussion

The following study aimed to examine the role that juror gender plays in civil juries, through qualitative analyses of jurors' deliberation behavior, and quantitative analyses of jurors' liability and damage award decisions.

Interruptions

We hypothesized that on average, male jurors would instigate more interruptions within a deliberation compared to females. However, results from the current study yielded no significant difference in average interruptions instigated by juror gender.

We also hypothesized that case-relevant factors (such as the amount of guidance jurors are given to come up with their damage award, which was manipulated in the current study) would engender differences in interruption rates. We expected that juries in the control condition would

have experienced significantly more interruptions than those in the gist plus numeric condition, due to the difficulty (and subsequent potential conflict among jurors) that having such limited guidance (in the control condition) would create. However, no significant differences in interruption rates were found by condition. Future work should more closely examine the tone and surrounding context of the interruption. Prior empirical work on the coding of interruptions can provide insight into why it may be meaningful to do so. In their examination of how male versus females coded for interruptions within a deliberation, Okamoto et al. (2002) indicated that in general, the more severe or serious an intrusion of speech, the more likely coders were to code the intrusion as an interruption. However, this relationship was less accentuated for female vs. male coders. The authors argue that this finding may show that females are simply taking more from the broader situational context into account in evaluating the intrusive statement (e.g. the interrupter's intent, power position, etc.), which might be causing them to not interpret an interruption as such. Though interruptions in the deliberations of the current work are identified through an automated process (and transcripts of the deliberations are set up in such a way that interruptions are demarcated in a standard way across juries), the above work highlights the importance of taking the broader context of interruptions into account (e.g. the tone and intention behind the interruption) beyond just calculating their frequency.

Juror Gender and Numeracy-Related Statements and Behaviors During Deliberation

We hypothesized that male jurors would be more likely to be the first to propose a pain and suffering damage award number. However, data from the current study showed that juror gender was unrelated to propositions of the first pain and suffering number, assessed both across conditions and within each experimental condition. This lack of effect could be explained by both the difficulty of the task at hand (coming up with a quantity for something as subjective as

pain and suffering) such that both males nor females were challenged by boldly declare a first idea, or by jurors all announcing a potential damage award number simultaneously, such that it was difficult to discern whose number was truly first. Additionally, jurors may have chosen a more equitable approach to deciding a damage award, in which each juror proposed their own number and the conversation flowed from these number in aggregate, as opposed to from one sole numeric proposal.

We also hypothesized that in the context of a deliberation, male jurors would exhibit their possible greater numeracy ability by being more likely to be selected as the “number person” within the jury, and also making a greater number of statements alluding to their comfort with math. (See longitudinal work by (Arora & Pawlowski, 2017) indicating men tended to have higher mathematical literacy than females, persisting over time). However, results from the current study indicated that selection of jurors as a number person, and expression of math comfort, were both unrelated to juror gender. These null findings could be driven by generally low instances of these concepts arising within a deliberation; together, these two qualitative codes were only identified 11 times within the total sample. While male jurors perhaps may indeed feel more comfortable with math than their female counterparts, this tendency may not be best captured during a jury deliberation.

Juror Gender and Liability and Damage Award Decisions

As predicted, decisions of liability for the defendant Jeffrey Burnside did not significantly differ by juror gender. Regarding damage award calculations, we hypothesized that female jurors would have, on average, lower damage award estimates compared to male jurors. This hypothesis was not upheld however, as natural log-transformed damage award estimates did not significantly differ by juror gender.

Future Directions and Implications

Though the current work did not identify differences by a juror's gender in specific deliberation behaviors, liability decisions, or average damage awards proposed, there are many potential avenues for future work on this set of questions. For example, in order to more accurately assess how jurors calculate lost income damages for plaintiffs of different genders, future studies should include a male and female plaintiff within the case facts. Additionally, regarding deliberation behavior, future work can more closely examine the severity (e.g. tone, intention) behind interruptions instigated by a juror, beyond interruption frequency. More targeted analyses can also be conducted to examine how numeracy in males versus females may affect their deliberation behavior (in terms of to what extent mentions of math comfort or being the designated number person arise) as well as their damage award estimates.

Ultimately, through this work, we hope to identify facets unique to civil jury deliberation (i.e. discussion about numeracy, proposals of damage awards) that can then engender differences in deliberation behavior and statements by a juror's gender.

REFERENCES

- Arora, A., & Pawlowski, E. (2017). *Examining Gender Differences in the Mathematical Literacy of 15-Year-Olds and the Numeracy Skills of the Age Cohorts as Adults*.
- Barroso, A., & Brown, A. (2021). *Gender pay gap in U.S. held steady in 2020*. Pew Research Center. <https://www.pewresearch.org/fact-tank/2021/05/25/gender-pay-gap-facts/>
- Chapman, G., & Bornstein, B. (1996). *The More You Ask for, the More you Get: Anchoring in Personal Injury Verdicts* (pp. 519–540).
- Goodman, J., Loftus, E. F., Miller, M., & Greene, E. (1991). Money, Sex, and Death: Gender Bias in Wrongful Death Damage Awards. *Law & Society Review*, 25(2), 263.
<https://doi.org/10.2307/3053799>
- Greene, E., Johns, M., & Bowman, J. (1999). The effects of injury severity on jury negligence decisions. *Law and Human Behavior*, 23(6), 675–693.
<https://doi.org/10.1023/A:1022341522714>
- Hans, V. P., & Lofquist, W. S. (1994). Perceptions of civil justice: The litigation crisis attitudes of civil jurors. In *Behavioral Sciences & the Law* (Vol. 12, Issue 2).
<https://doi.org/10.1002/bsl.2370120207>
- Hastie, R., Penrod, S. D., & Pennington, N. (1983). Inside the Jury. In *Inside the Jury*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674865945>
- Hastie, R., Schkade, D. A., & Payne, J. W. (1999). Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards. In *Law and Human Behavior* (Vol. 23, Issue 4).
- Helm, R. K., Hans, V. P., & Reyna, V. F. (2017). Trial by Numbers. In *Cornell Journal of Law and Public Policy* (Vol. 27).

<https://scholarship.law.cornell.edu/cjlpp> Available at: <https://scholarship.law.cornell.edu/cjlpp/vol27/iss1/3>

- Helm, R. K., Hans, V. P., Reyna, V. F., & Reed, K. (2020). Numeracy in the jury box: Numerical ability, meaningful anchors, and damage award decision making. *Applied Cognitive Psychology*, 34(2), 434–448. <https://doi.org/10.1002/acp.3629>
- Lamb, S. (1996). *TOWARD GENDER-NEUTRAL DATA FOR ADJUDICATING LOST FUTURE EARNING DAMAGES : AN EVIDENTIARY PERSPECTIVE*.
- Marder, N. S. (1987). *Gender Dynamics and Jury Deliberations*.
- Mazei, J., Hüffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., & Hertel, G. (2015). A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological Bulletin*, 141(1), 85–104. <https://doi.org/10.1037/a0038184>
- Nagel, S., & Weitzman, L. (1972). Sex and the Unbiased Jury. *Judicature*, 56(3), 108–111. <https://heinonline.org/HOL/Page?handle=hein.journals/judica56&id=110&div=30&collection=journals>
- Nemeth, C., Endicott, J., & Wachtler, J. (1976). *From the ' 50s to the ' 70s : Women in Jury Deliberations* Author (s): Charlan Nemeth , Jeffrey Endicott and Joel Wachtler Published by : American Sociological Association Stable URL : <http://www.jstor.com/stable/3033495> *Women in Jury Deliberations* *. 39(4), 293–304.
- Neubauer, D., & Meinhold, S. S. (1994). Too quick to sue? Public perceptions of the litigation explosion. *Justice System Journal*, 16(3), 1–14. <https://doi.org/10.1080/23277556.1994.10871179>
- Salerno, J. M., Peter-Hagene, L. C., & Jay, A. C. V. (2017). Women and African Americans are less influential when they express anger during group decision making. *Group Processes &*

Intergroup Relations, 1–23. <https://doi.org/10.1177/1368430217702967>

Strodtbeck, F. L., & Mann, R. D. (1956). Sex Role Differentiation in Jury Deliberations.

Sociometry, 19(1), 3. <https://doi.org/10.2307/2786099>

Vidmar, N. (1993). Empirical evidence on the deep pockets hypothesis: jury awards for pain and suffering in medical malpractice cases. *Duke Law Journal*, 43(2), 217–266.

<https://doi.org/10.2307/1372792>

Vidmar, N., & Rice, J. J. (2008). Assessments of Noneconomic Damage Awards in Medical Negligence : A Comparison of Jurors with Legal Professionals. *Iowa Law Review*, 78.

<https://heinonline.org/HOL/Page?handle=hein.journals/ilr78&id=895&div=29&collection=journals>

Vinson, K. V., Costanzo, M. A., & Berger, D. E. (2008). Predictors of verdict and punitive damages in high-stakes civil litigation. *Behavioral Sciences and the Law*, 26(2), 167–186.

<https://doi.org/10.1002/bsl.807>

CHAPTER 4:

HOW MUCH IS YOUR LOVED ONE WORTH? MORAL DECISION MAKING ACROSS CULTURES

Introduction

Moral judgments abound in our day to day lives: whether inside a courtroom making assessments of a defendant's culpability, a victim's credibility, deciding on a defendant's sentence, or outside the legal setting, deciding how to act most ethically when placed with a possible moral quandary. Philosophers and moral psychologists have long assessed the schemas one may use to make a moral judgment or decision, and for developmental psychologists, how these schemas of moral understanding and judgment may change over the course of the lifespan. Moral research has identified that our moral judgments can differ along a multitude of individual difference factors, such as political orientation, gender, and culture.

Schemas of Moral Decision Making: Deontology and Consequentialism

The application of deontological and consequentialist or utilitarian principles are two common schemas that have been argued to underpin our moral decisions. Under deontology, a principle originally developed by the philosopher Kant, our moral decisions are undergirded by a universal moral rule or reason that must be upheld or followed, regardless of the consequences of the action (e.g. "It is immoral to kill under any circumstances"). In contrast, consequentialist or utilitarian thinking (originally developed by John Stuart Mill) emphasizes that the morality of an action or judgment is based on its consequences, and such, one should make moral decisions that serve to maximize the greatest good for the greater number of people (e.g. "It's ok to kill one person, as long as the consequence may be that the most number of lives are saved as possible") (Conway & Gawronski, 2013a).

The Trolley Car (and associated) Paradigms of Assessing Moral Decision-Making

One of the most common paradigms used to assess moral decision making is the Trolley Car problem, originally developed by the English philosopher Philippa Foot. The paradigm depicts a trolley car speeding down the tracks towards five people who are tied up. However, there is a lever that would divert the trolley to one other person tied on a sidetrack of the trolley car tracks. If you push this lever, the trolley car would be diverted towards this one person and they would be killed, but the five individuals tied up on the main trolley tracks would be saved. In this paradigm, a utilitarian would hold that the five people should be saved by diverting the trolley to the side and subsequently sacrificing the one individual, as this action would maximize the greatest good for the greatest number of people. A deontologist, on the other hand, would hold that irrespective of the number of lives saved, even sacrificing one individual is moral unacceptable (Conway & Gawronski, 2013b; Thomson, 1985).

Schemas of Moral Decision Making: Haidt's Social Intuitionist Model

Alongside deontology and consequentialism, Jonathan Haidt's social intuitionist model also explains the process of moral reasoning. In this model, Haidt argues that our moral judgments are generated by fast, gut-reactions we have to moral stimuli, and that any more extensive moral reasoning happens after our initial moral judgments are made. As such, Haidt believes that moral judgments precede moral reasoning, rather than judgments being a consequence of moral reasoning. Importantly, Haidt also argues that these moral judgments are not limited to the individual, but are dependent on our social and cultural context (Haidt, 2001).

The Role of Culture in Moral Decision Making

Work by Sachdeva, Singh, & Medin (2011) describes the importance of studying culture in moral decision making, discussing that the core moral values we hold - as conceptualized by

Graham and colleagues' five Moral Foundations of Harm/Care, Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, and Purity/Sanctity – may vary by cultural group (Graham et al., 2011; Sachdeva et al., 2011). For example, some work by Haidt, Koller, and Dias (1993) indicates that individuals of varying culture and socio-economic statuses may differ on how they evaluate moral acts involving possible violations of harm/care, authority/respect, or purity/sanctity (Haidt et al., 1993).

Beyond cultural differences in ascription to certain universal moral values, additional work has examined cross-cultural differences in moral decision making involving utilitarian or deontological principles, as assessed by the Trolley Car paradigm and its iterations.

Cultural Differences (and Lack Thereof) in Responses to the Trolley Car Problem and its Iterations:

In an effort to unearth a possible set of universal moral principles (i.e. principles or utilitarianism or deontology) that may exist across cultural lines, a large body of work has examined moral decision-making cross-culturally. For example, work by O'Neill and Petrinovich (1998) presented US and Taiwanese undergraduates with 25 different iterations of the Trolley Car problem, as well as a set of questions assessing their personal beliefs on a variety of moral issues (e.g. capital punishment, abortion, etc.). Participants' responses to these dilemmas were coded for the presence of different moral dimensions such as the notion of taking action / inaction, inclusive fitness (i.e. selecting an option that would benefit your kin), or speciesism (i.e. selecting an option to benefit humans over non-humans). The authors found that across both US and Taiwanese populations, participants employed similar moral dimensions in making their decisions, especially those of speciesism, action / inaction, and inclusive fitness. They argue that their work underscores that across cultures, participants will make moral

decisions that have some evolutionary benefit and encourage reciprocal altruism (O'Neill & Petrinovich, 1998). Other work by Awad et al. (2019) examined moral decision making across 42 different countries and 70,000 participants. They found when presented with three sacrificial moral dilemmas (the traditional Trolley Car "Switch" paradigm, and the loop and footbridge versions of the paradigm), all participants across all cultures provided the same order of the acceptability of each type of sacrifice, with the Switch paradigm being most acceptable, followed by the Loop and Footbridge scenarios (Awad et al., 2019).

However, other work employing similar moral decision-making frameworks (via the Trolley Car paradigm and its iterations) has identified cultural differences in decision-making. This undercuts the idea put forth by the above work that our moral decisions are driven by the same universal principles, and that these principles are applied similarly across cultures.

For example, work by Ahlenius & Tännsjö (2012) found that Chinese and US and Russian (Western) participants responded differently to a set of Trolley Car-esque moral dilemmas (the traditional switch paradigm, the footbridge scenario, and the loop scenario). Namely, they found that across these dilemmas, Chinese participants were less likely than either sub-sample of Western participants (US or Russian) to choose the utilitarian response, that would maximize the greatest good (i.e. sacrificing one individual to save the lives of others) (Ahlenius & Tännsjö, 2012). Additional work by Gold et al. (2014) confirms this trend, finding that in response to Trolley Car-related dilemmas, Chinese participants were less likely to select the option in the dilemma that would sacrifice one life to save five, and were less likely to think that such an action would be acceptable, compared to their Western, British counterparts. In other words, Chinese participants were shown to be less utilitarian compared to Westerners. Additionally, this cultural difference was accentuated when participants were presented with

scenarios different from, and less severe than, death (i.e. scenarios involving losing prize money during a game show, or an orphan(s) losing their meal(s)). The trend found with British participants in this experiment also mirrored that found in American participants in a virtual reality version of the trolley car paradigm (Gold et al., 2014; Navarrete et al., 2012).

What could explain these cultural differences in decision making? Gold et al. (2014) argue that one cultural-specific attitude that may be driving Chinese participants' decision to not take action in these trolley car paradigms (e.g. by allowing the trolley car to sacrifice five individuals, rather than taking action to divert it and save the five individuals by sacrifice one individual instead) is their belief in fate, and allowing things to run their natural course without intervention. In their work, Chinese participants were much more likely to subscribe to this set of attitudes, compared to their British counterparts.

The Impersonality of Moral Decision-Making, Adding a Human Element

Many scholars have critiqued the Trolley Car and associated moral dilemma paradigms for their contrived, impersonal, and unrealistic manner of assessing moral decision making, by solely employing hypothetical situations involving strangers (BLOOM, 2011; Khazan, 2014). For example, in his critique of the Trolley Car and associated paradigms, and his call to add a more realistic, human element to these moral decisions, Paul Bloom argues that “what’s clear is that human morality did not evolve for the sake of dealing with strangers” (BLOOM, 2011).

In an effort to increase the realism associated with such moral dilemmas, some work has transformed the Trolley Car problem into virtual reality experiments (e.g. Navarrete et al., 2012). While the paradigm is the same in this work and thus still represents a hypothetical, quite unrealistic dilemma, participants are required to engage in a (virtual) behavior, rather than passively answering a set of questions. Other work, however, has adopted tenets of the Trolley

Car paradigm into more realistic situations involving economic or emotional loss (Gold et al., 2013, 2014).

Adding Realism to Moral Decisions via the Inclusion of Loved Ones

Beyond extending scenarios into the domain of economic or emotional loss, another facet of realism to bolster these artificial scenarios could involve adding information about a participant's actual loved one (friends, family, partner, etc.) to the moral decision-making calculus. Evidence of a lack of cultural differences in moral decisions when loved ones are at stake would align with aforementioned work by O'Neill and Petrinovich (1998), showing that across cultures, we are evolutionarily driven to save our loved ones over strangers. However, affirmative evidence of cultural differences in dilemmas involving loved ones could speak to the common dimension on which cultural differences are often defined, as collectivist versus individualist. A plethora of work in cultural psychology has examined these two constructs, but generally speaking, collectivist cultures – often found within Eastern societies such as China or Japan - tend to place an emphasis on selflessness, group loyalty, and group and social cohesion, while individualistic cultures – often found in Western society such as in Canada, the UK, or the United States – qualities such as autonomy and individual rights are upheld (Triandis, 1988, 1992, 2018).

The Current Study

The current work sought to extend previous work on cross-cultural moral decision making by evaluating moral reasoning in hypothetical as well as more real-life scenarios (specifically, emotional and economic loss), while also adding a much more personal element to the moral decision-making calculus, with the inclusion of one's loved ones at stake. In doing so, the current work expands upon simpler assessments of utilitarian thinking (e.g. would you save

five lives to sacrifice one) by asking how utilitarian principles may or may not be upheld when the livelihood of personal kin (e.g. their lives, emotional well-being, or their economic well-being) are at stake.

A series of hypotheses are postulated regarding how individuals from individualistic (United States) versus collectivistic cultures (China) will respond to moral dilemmas involving pitting their loved one against an increasing number of strangers within the Trolley Car paradigm, a dilemma involving emotional loss, and a dilemma involving economic loss.

In deciding between saving their loved one over one stranger in the Trolley Car problem, we hypothesize that participants, across cultural groups, will be more likely to choose their loved one. This finding would align with the notion that even in individualistic cultures in which family or relationship-ties may be less emphasized, when the life of a loved one is at stake, members of both individualistic and collectivistic cultures will be equally as likely to save that loved one, over a stranger.

Further, in making decisions between saving their loved one and ten, 100, or 1000 strangers, we hypothesize that Chinese participants will be more likely than Americans to choose to save to save the strangers over their loved one. We hypothesize that this choice could be driven by Chinese participants' greater propensity for collectivistic thinking – conceptualized here as a concern and care for others above and beyond their own kin, and Americans' propensity for individualistic thinking – which we conceptualize here as a narrower concern for just one's kin, regardless of the plight of other strangers.

Lastly, driven by prior work showing that cultural differences in moral decision-making held (and were even accentuated) more realistic situations of emotional and economic loss (Gold et al., 2014), we hypothesize that the aforementioned cultural differences will be replicated in

emotional and economic loss dilemmas being depicted in the current work.

Though work on cross-cultural differences in reaction time for various moral decisions has been understudied, we hypothesized that within scenarios involving saving your loved one or ten, 100, or 1,000 strangers, Chinese participants will be faster, on average, compared to their American counterparts in making these decisions. This could be explained by the “1 vs. 10 ratio” serving as the point (within these dilemmas) at which Chinese participants’ collectivistic tendencies prompt them to more quickly decide to choose the well-being of others (even strangers) over their own personal welfare (that of their loved one).

Methodology

Participants

Two different sub-samples of participants were recruited for the current study. A Chinese community sample was recruited using the online recruiting website WJX.com, and an American community sample was recruited via Amazon Mechanical Turk. A sample of Chinese and American undergraduates were also collected (using the same procedure as will be described), however, data from the two community samples are presented here, due to their greater representativeness. Participants recruited via Amazon Mechanical Turk were compensated \$1.50 for their participation, and Chinese participants recruited WJX.com were compensated the equivalent of approximately \$0.77 for participating. All participants completed an online survey crafted via Qualtrics.

After excluding participants who did not complete the study in its entirety or failed the attention check questions, the sample consisted of 151 Chinese participants and 141 American participants, for a total of 292 participants recruited across cultural groups. Detailed descriptive statistics for each sub-population can be found in Table 4.1.

Table 4.1 Descriptive Statistics for US and Chinese Participants

	US MTurk (N = 141)		Chinese WJX (N = 151)	
	N or Range	% or Mean	N or Range	% or Mean
<i>Age</i>	19-71	33.82	19-77	34.56
<i>Gender:</i>				
Male	93	66.0%	76	50.3%
Female	48	34.0%	75	49.7%
<i>Race of US Participants</i>			-----	
American Indian / Alaska Native	3	2.1%		
Asian	7	5%		
Black or African American	8	5.7%		
Hispanic / Latino	0	0%		
Multiracial	0	0%		
Native Hawaiian or Pacific Islander	1	0.7%		
Other	4	2.8%		
White	118	83.7%		
Race not specified	0	0%		

Design

The study conformed to a 3 (Moral dilemma scenario type: traditional trolley car problem, emotional dilemma, economic loss dilemma) x 4 (Number of strangers asked to sacrifice within the dilemma: 1, 10, 100, 1,000 strangers) within-subjects design. All participants were presented with twelve moral dilemmas and asked a series of follow-up questions in response to each dilemma. The aforementioned design was retained for Chinese participants, but the entirety of the survey was translated into Mandarin. The exact content of each scenario is delineated in the Procedure section.

Procedure

All participants first read an informed consent form, indicating that the study would take about 10-15 minutes to complete, and that upon completion, they would receive research credit or be compensated for their participation. Next, participants were told that they would be presented with a set of moral dilemmas, and that they would have a minimum of eight seconds to read the dilemma (this minimum was put in place to ensure that participants read the stimuli thoroughly), and five seconds to provide an answer to the questions accompanying each dilemma. The five-second time limit was put in to place to pressure participants to provide quick, intuitive responses to these situations, rather than making longer, deliberative decisions. The impetus for providing such quick judgments is in alignment with Jonathan Haidt's Social Intuitionist Model, which argues that our moral judgments are primarily driven by fast gut-reactions to moral stimuli, and that more deliberative moral reasoning is completed post-hoc, after any initial moral intuitions we may hold (Haidt, 2001). However, in an effort to collect responses to all dilemmas from all respondents, participants were also unable to move forward in the survey unless they completed the questions associated with each dilemma. As such, some participants went through multiple iterations of five-second sets to provide an answer, ultimately

taking longer than five seconds to respond. Further information regarding participant response times is delineated in the Results section.

After this introductory information, participants were presented with a slightly revised iteration of the traditional Trolley Car problem, a moral dilemma involving emotional distress, and a moral dilemma involving economic loss. Each dilemma had four different versions, involving 1, 10, 100, or 1,000 strangers being depicted in the dilemma. The general framework of each dilemma and their follow-up questions were as follows:

Revised Trolley Car Dilemma:

“Think about the person you love and care most deeply about.

Now, imagine that you see a runaway trolley car speeding towards this person, who is incapacitated on the tracks.

You are standing next to a lever which controls the direction in which the trolley car is headed.

On a side track, you see {1, 10, 100, 1,000} stranger(s), lying down tied up on the tracks. If you pull the lever, the train will be redirected to the stranger, saving the person you care about, and striking the stranger.

If you do not pull the lever, however, the stranger(s) will be saved, and the person you love will be struck by the trolley.”

Do you pull the lever – saving the person you care about and killing {1, 10, 100, 1000} strangers? (Responses: Yes, No).

Within all iterations of the revised trolley car dilemma, each scenario included an image of a trolley car, and sample images of 10, 100, and 1,000 individuals. These images were included to help facilitate participants’ understanding and imagination of the situation.

Emotional Loss scenario:

“Suppose you’re driving a van every morning that would take **1** adult to their workplace, and your child to school.

Normally you could make it to both places on time, regardless of the sequence of destinations you go to.

However, on this particular morning, the traffic is extremely heavy, and you estimate that either your child or the {1, 10, 100, 1,000} adult(s) will be late (depending on which destination you go first), and because all of you value being on time, being late would cause extreme emotional stress.”

Which destination would you go first? (Responses: Take the child to school, Take {1, 10, 100, 1,000} adult(s) to work).

Economic Loss Scenario:

“Suppose you’re a janitor who is also serving as a security guard in a gym locker room.

During your shift, you see your child putting his smartphone and laptop in a locker on the right side of the gym, and {a, 10, 100, 1,000} stranger(s) put their smartphone and laptop in a locker on the left side of the gym.

As you're cleaning, both lockers start to fall down.

You can only stop one locker from falling over, but the other locker will collapse, thus breaking all the valuable items stored inside.”

Which locker will you protect? (Responses: My child’s locker, the stranger’s / strangers’ locker).

Two attention check questions were interspersed throughout the scenario iterations.

Answers to these questions were used as exclusion criterion for all participants. Following the presentation of all 12 scenarios, participants were asked: “In a few sentences, please reflect on your decision-making process for the above scenarios. Moreover, please describe the factors you felt went into your decisions, the level of difficulty you felt was involved in making your decisions, and any other information you feel is relevant.” Participants’ responses to this question were later thematically analyzed by a team of research assistants. Lastly, participants’ demographic information was collected: US participants were asked for their age, gender, race, and political orientation, and Chinese participants were asked for their age and gender.

Results

Moral Dilemma Decisions by Cultural Group and Age

To examine the presence of cultural differences in responses to each type of moral dilemma (trolley car, emotional distress, and economic loss), a series of binary logistic regressions were conducted, Chi squared tests of independence were conducted in follow-up, where applicable. Across all scenarios, driven by the impetus in the current study to understand how participants across cultures make moral decisions when their loved ones are involved, the decision to save or select one's loved one (or their valuables, in the case of the economic loss scenario) was coded as 1, and decisions to save or select stranger(s) was coded as 0. The categorical variable of cultural group (US or Chinese) was entered as a predictor to test hypotheses related to cross-cultural differences in decision-making, and participants' age and gender were also added as predictors, for exploratory purposes.

Cross-Cultural Responses to a (Revised) Trolley Car Paradigm

A set of binary logistic regressions found that cultural group was a significant predictor of participants' decision to save their loved one in the trolley car scenario (and sacrifice strangers) rather than save one, ten, 100, or 1000 strangers. Age and gender were not significant predictors of participants' decision to save their loved one over any number of strangers (1, 10, 100, or 1,000); $p > 0.05$ for all comparisons. Detailed regression results are displayed in Table 4.2.

Table 4.2: Binary Logistic Regression Predicting Participants' Trolley Car Decision to Save their Loved One over Stranger(s) by Cultural Group

Scenario:	β	SE	Wald χ^2	p	Odds Ratio	OR 95% CI
1 stranger	1.05	0.34	9.74	0.002	2.87	1.48 – 5.54
10 strangers	0.89	0.25	12.69	< 0.001	2.42	1.49 – 3.94

100 strangers	0.84	0.26	10.89	0.001	2.32	1.41 – 3.83
1000 strangers	0.74	0.27	7.88	0.005	2.11	1.25 – 3.54

*Note: statistics are associated with the dichotomous predictor “Cultural Group”; Reference Category: US

Targeted chi-square tests of independence were conducted to disaggregate the significant effect of cultural group on trolley car decisions, and demonstrated that the proportion of US participants who chose to save their loved one over 1, 10, 100, and 1,000 strangers was significantly *higher* than that of Chinese participants. Details of the results of each chi square test for each iteration of strangers being assessed can be found in Table 4.3.

Table 4.3. Chi Square Test of Independence Between Cultural Group and Participants’ Decision to Save a Loved One over Stranger(s) in a Trolley Car Scenario

Scenario:	χ^2	df	N	p
1 stranger	10.36	1	292	0.001
10 strangers	12.89	1	292	< 0.001
100 strangers	13.02	1	292	< 0.001
1000 strangers	7.82	1	292	0.005

Cross-Cultural Responses to Emotional Loss Dilemmas

A set of binary logistic regressions found that cultural group was a significant predictor of participants’ decision to take their child to school (and make strangers late for work) rather than take one or 10 strangers to work, as depicted in an emotional loss dilemma. Exact statistics of the regression can be found in Table 4.4. However, age and gender were not significant predictors of participants’ decision to take their child to school over any number of strangers (1, 10, 100, or 1,000); all ps > 0.15.

Table 4.4: Binary Logistic Regression Predicting Participants’ Emotional Loss Dilemma

Decisions to Take their Child to School over Stranger(s) to Work, by Cultural Group

Scenario:	β	SE	Wald χ^2	p	Odds Ratio	OR 95% CI
1 stranger	-1.24	0.26	23.35	< 0.001	0.29	0.18 – 0.48
10 strangers	-0.68	0.26	6.62	0.01	0.51	0.3 – 0.85
100 strangers	-0.36	0.28	1.57	0.21	0.70	0.50 – 1.52
1000 strangers	-0.14	0.29	0.24	0.63	0.87	0.50 – 1.52

*Note: statistics are associated with the dichotomous predictor “Cultural Group”; Reference Category: US

Targeted chi-square tests of independence were run to disaggregate the significant effect of cultural group on participants’ decision to take their child to school (making strangers late to work), compared to taking 1 or 10 strangers to work. Results indicated that the proportion of US participants who indicated they would take their child to school, over taking 1 or 10 strangers to work, was significantly *lower* than that of Chinese participants. Details of the results of each Chi square test for each iteration of strangers being assessed are shown in Table 4.5. A graphical representation of this set of results can also be found in Figure 4.1 and Figure 4.2.

Table 4.5. Chi Square Test of Independence Between Cultural Group and Participants’ Decision to Take their Child to School over Stranger(s) to Work, in an Emotional Loss Dilemma

Scenario:	χ^2	df	N	p
1 stranger	26.41	1	292	< 0.001
10 strangers	6.91	1	292	0.009
100 strangers	1.47	1	292	0.23
1000 strangers	0.26	1	292	0.61

Figure 4.1. Proportion of US vs. Chinese Participants Who Would Choose to Take their Child to School over 1 Stranger(s) to Work, in an Emotional Loss Dilemma

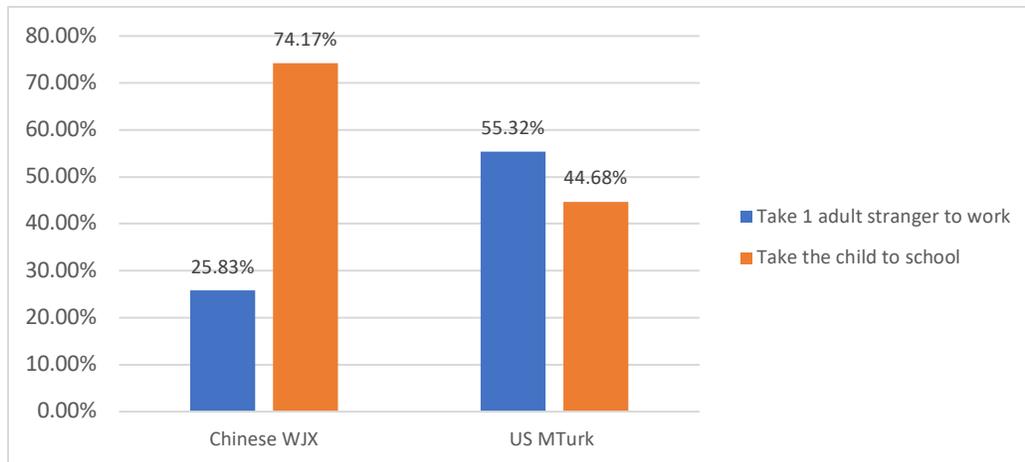
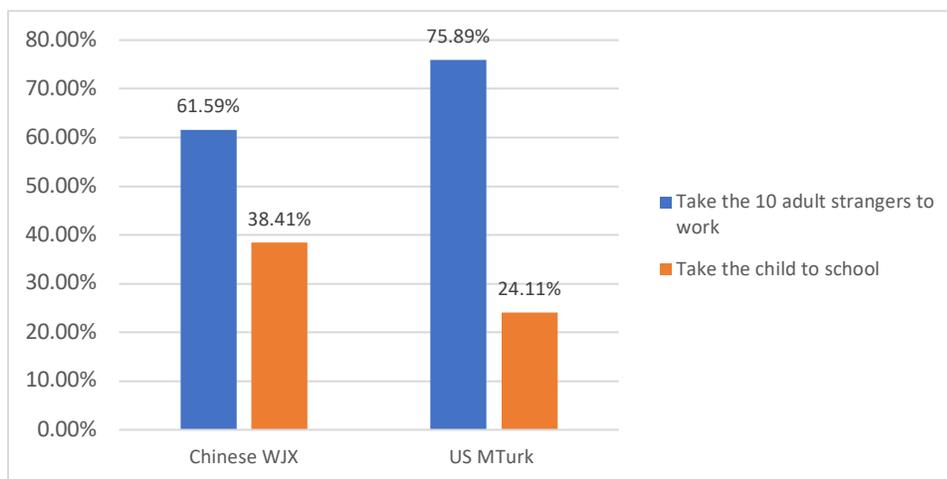


Figure 4.2. Proportion of US vs. Chinese Participants Who Would Choose to Take their Child to School over 10 Stranger(s) to Work, in an Emotional Loss Dilemma



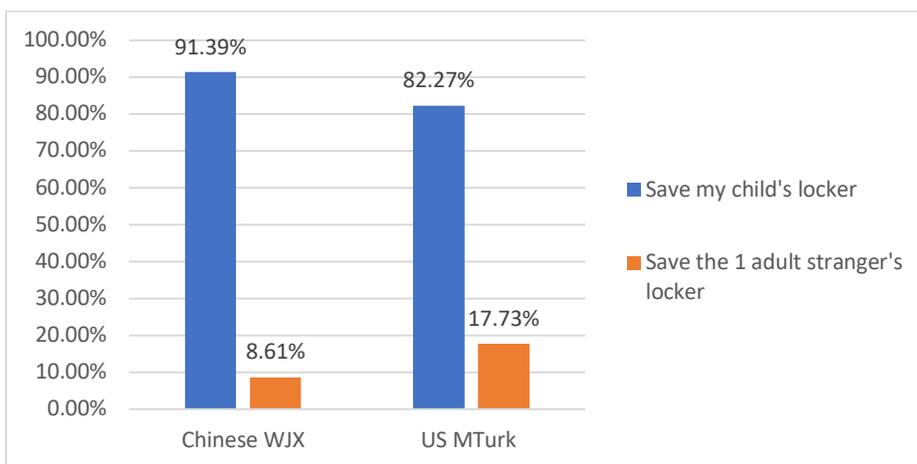
Cross-Cultural Responses to Economic Loss Dilemmas

Lastly, a set of binary logistic regressions found that cultural group was a significant predictor of participants' decision to save their child's locker (subsequently breaking strangers' valuables in their lockers) rather than saving the locker (and their valuables) of one stranger, as

depicted in an economic loss dilemma ($\beta = -0.893$, $SE = 0.37$, Wald (1) χ^2 , 5.71, $p = 0.02$, Odds Ratio = 0.41). However, cultural group was not a significant predictor of participants' decisions to save their child's locker compared to that of 10, 100 or 1000 strangers ($ps > 0.21$). Age and gender were also not significant predictors of participants' decision to save their child's locker / valuables over that of any number of strangers (1, 10, 100, or 1,000); all $ps > 0.05$.

A Chi square test of independence was run to disaggregate the significant effect of cultural group on participants' decision to save their child's locker / valuables, compared to that of one stranger. Results indicated that the proportion of US participants who indicated they would save their child's locker / valuables over that of a one stranger's was significantly *lower* than that of Chinese participants: χ^2 (1, N= 292) = 5.36, $p = 0.02$. A graphical representation of this set of results can also be found in Figure 4.3.

Figure 4.3. Proportion of US vs. Chinese Participants Who Would Choose to Save their Child's Locker / Valuables Over that of 1 Stranger, in an Economic Loss Dilemma



Timing Data for All Scenario Types – Trolley Car, Emotional, and Economic Loss

A series of independent sample t-tests were conducted to assess cultural differences in the average amount of time (in seconds) it took participants to make their moral decisions (i.e. to save their loved one over stranger(s), to take their child to school over stranger(s) to work, or to

save their child’s locker / valuables over that of stranger(s)).

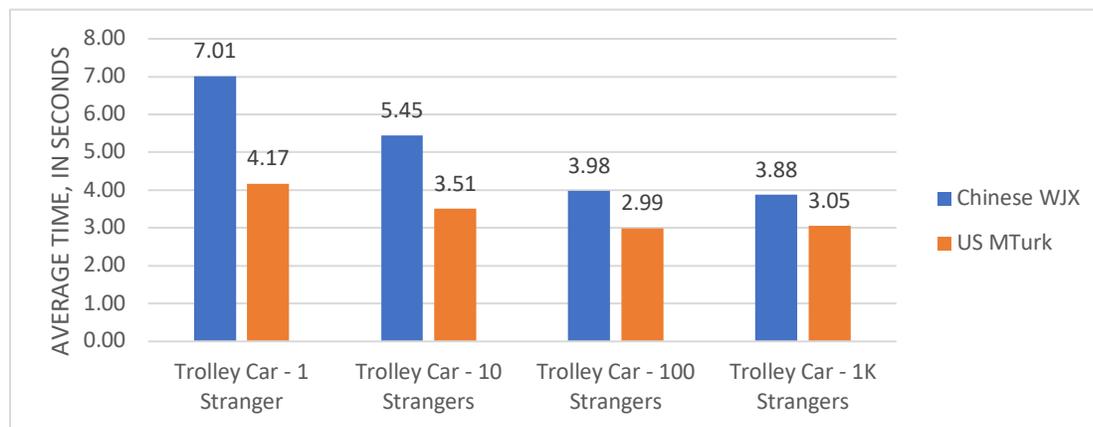
Regarding decision time for the revised trolley car scenario, results indicated that Chinese participants, on average, took longer to respond than US participants, across all stranger-iterations of the scenario. The statistics associated with these tests can be found in Table 4.6.

Additionally, a graphical representation of these cultural differences can be found in Figure 4.4.

Table 4.6 Independent Sample T-Tests Assessing Average Response Time to Trolley Car Dilemmas, by Cultural Group

Scenario:	t	df	p	M _{timing(sec)} (SD) Chinese Participants	M _{timing(sec)} (SD) US Participants
1 stranger	3.57	168.53	< 0.001	7.01 (9.2)	4.17 (2.2)
10 strangers	2.82	173.83	0.005	5.45 (8.13)	3.51 (2.2)
100 strangers	2.81	188.87	0.005	3.98 (4.05)	2.99 (1.42)
1000 strangers	2.66	244.70	0.008	3.88 (3.28)	3.05 (1.92)

Figure 4.4. Average Time to Make a Decision for Stranger-Iterations of the Trolley Car Dilemma, by Cultural Group



No significant differences by cultural group were found for average response times for any iteration of emotional loss scenarios (all ps > 0.14).

In scenarios involving economic loss, Chinese participants, on average, took longer to reply to

economic loss situations involving 100 and 1,000 strangers than US participants. The statistics associated with these comparisons can be found in Table 4.7.

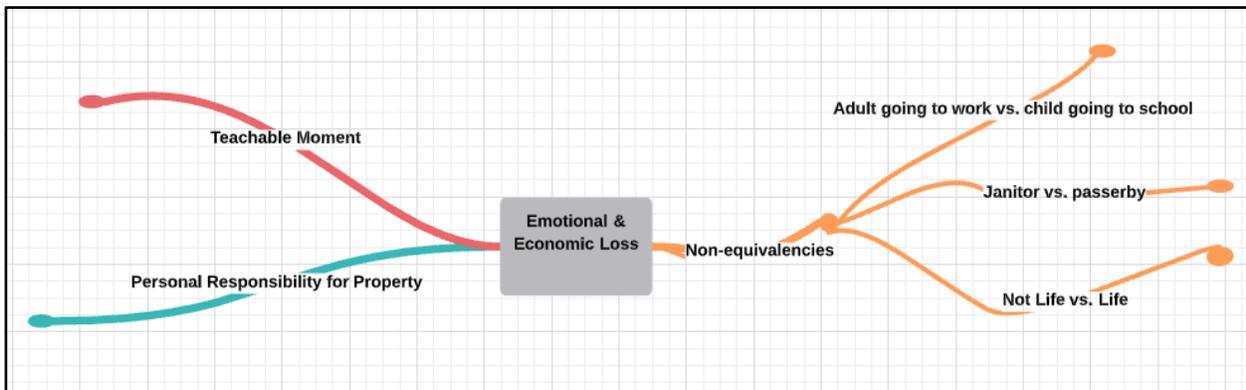
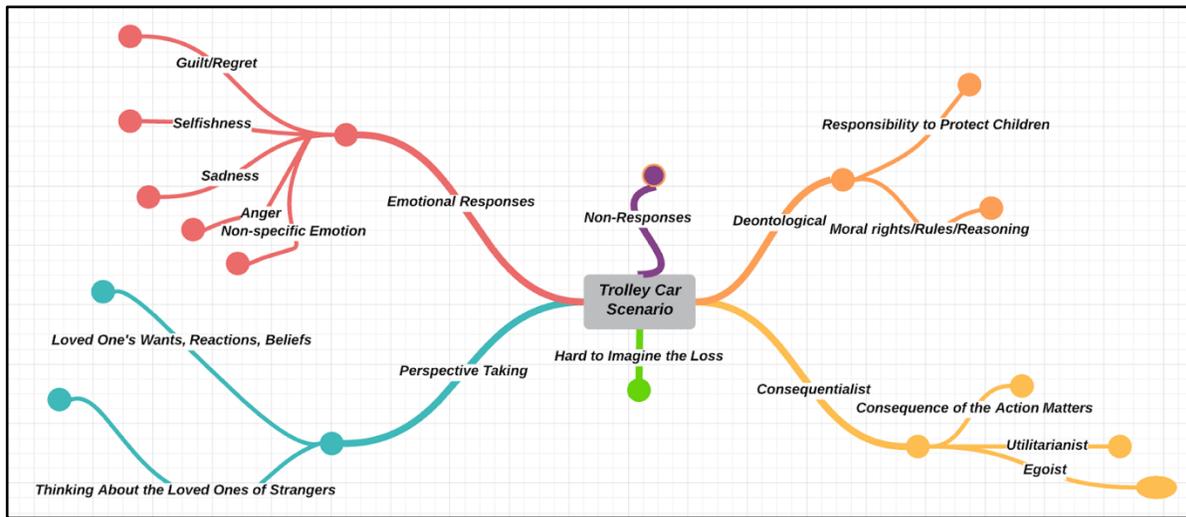
Table 4.7 Independent Sample T-Tests Assessing Average Response Time to Economic Loss Dilemmas, by Cultural Group

Scenario:	t	df	p	M _{timing(sec)} (SD) Chinese Participants	M _{timing(sec)} (SD) US Participants
100 strangers	2.07	169.67	0.04	3.51 (7.16)	2.26 (1.78)
1000 strangers	2.66	244.70	0.008	2.52 (1.83)	1.98 (1.10)

Thematic Analyses of Participants’ Open-Ended Responses

A set of research assistants thematically analyzed the open-ended responses given by American participants, in which they explained their rationale for their decision in both the trolley-car and economic and emotional loss scenarios. These themes were developed based on both a-priori predictions of certain concepts participants’ would bring up in these explanations (e.g. the distinction between consequentialist versus deontology-based explanations), as well as themes that emerged during data collection (e.g. the notion of a non-equivalency or teachable moment informing a participant’s ethical decision. An illustration of these themes and their sub-codes can be seen in Figure 4.5. A more detailed explanation of each code, accompanied with illustrative quotes for each, can be found in Appendix D.

Figure 4.6. Thematic Analysis of US Participants' Open-Ended Responses to Explain their Moral Decision-Making Process



Discussion

The current work sought to add personal, more realistic elements to previous research on cross-cultural moral decision making. This was accomplished via the inclusion of more realistic moral dilemmas, and the dichotomy between a close loved one and a stranger being added to participants' decision-calculus.

Cross-Cultural Moral Decisions

We hypothesized that across scenario-types (Trolley Car, emotional, or economic loss),

in a one-one situation (i.e. your loved one versus a stranger's life, emotional well-being to be on time, or economic well-being to protect valuables), participants across cultures would be more likely to select their loved one. However, contrary to hypotheses, Chinese participants were found to be more likely to save one, 10, 100, or 1,000 strangers within the Trolley Car paradigm, over their loved one. From the framework of collectivistic versus individualistic cultures, this finding seems to qualify definitions of collectivism, such that the selflessness of a collectivistic (in this case Chinese) individual is exemplified by choosing to save a stranger's life even when their own loved one's is at stake.

It was also hypothesized that the cultural differences (such that Chinese participants would be more likely to save 10, 100, or 1,000 strangers) would hold for both emotional and economic loss situations. However, results from the current study do not uphold this hypothesis: in situations of emotional loss, Chinese participants were less likely to take one or ten strangers to work (over taking their child to school) compared to their American counterparts. And in instances of economic loss, Chinese participants were less likely to save the valuables of a stranger (over those of their child) compared to their American counterparts.

These results could indicate that situations of emotional and economic loss invoke different moral principles that have differential value for Chinese versus American participants. Chinese participants in this study may care more about the emotional well-being (and in this case timeliness) of their own child, as well as protecting their belongings, compared to their American counterparts. Future work should investigate what specific, culturally-varied moral principles such scenarios may evoke, and how they could align, for example, with Graham and colleagues' five well-known Moral Foundations.

Cross-Cultural Differences in Reaction Time

Lastly, it was hypothesized that on average, Chinese participants would be faster than their American counterparts to answer questions (across scenarios) pitting their loved one against ten, 100, or 1,000 strangers. However, results from the current study indicated that Chinese participants actually took *longer* to answer all iterations of the Trolley Car problem, as well as iterations of the economic loss scenario involving 100 and 1,000 strangers. Chinese participants' overall slower reaction time to these situations could be indicative of the fact that these dilemmas (regardless of harm-type) pit together two choices that, if either were selected, both could in fact indicate a collectivistic response. That is, either choice (to save a loved one or stranger) is still an other-oriented response, tenets of which are key to collectivist thinking. As such, Chinese participants' slower reaction time could be indicative of the greater degree of contemplation needed to choose between options that both arguably align with their cultural group's dominant values.

Limitations and Future Directions:

Though the current work succeeded in adding a more human element to cross-cultural decision-making, it is not without its limitations, and there are many facets of the work that are ripe for future research.

One major limitation is methodological in nature: the presentation of the three moral dilemmas to participants was not counter-balanced, which could lead to detrimental, confounding, order effects. In other words, participants' responses to these dilemmas could be driven simply by the order in which they were presented, rather than based on an evaluation of their actual content.

An additional facet that blurs the ability to cleanly compare participants' responses to these three scenarios is that the economic loss scenarios depicted here does not clearly mirror the

Trolley Car paradigm. The Trolley Car dilemma depicted here involves actively bringing harm to a stranger by choosing to save your loved one (or vice versa), whereas in the economic loss scenario, both lockers in the locker room will fall down anyways, and it is simply a matter of which locker the participant would choose to save. As such, this situation does not involve actively harming another individual, and saving either locker could still be construed as creating some net good (from a utilitarian perspective).

Driven by prior work in moral decision making that has implemented such measures, future work should add questions regarding the perceived moral acceptability of an action, to get a fuller picture of participants' decision-making process. To assess cultural-specific values that may contribute to participants', questions should be added to assess participants' belief in fatalism, or letting things just run their course without intervening. Additionally, future studies should control for value differences (i.e. to what participants from each culture value punctuality, education, financial stability, etc.) that may be driving or overshadowing any actual differences in moral reasoning between the two cultures.

Lastly, the current work assumes a clear dichotomy between collectivism and individualism across Chinese and American cultures. However, much recent work has sought to dispel this clear cut divide, by discerning individual differences in collectivism, especially in Eastern cultures (CAO, 2009). As such, future work should assess moral decision making within more targeted sub-samples of individualistic individuals within otherwise collectivistic cultures.

REFERENCES

- Ahlenius, H., & Tännsjö, T. (2012). Chinese and westerners respond differently to the trolley dilemmas. *Journal of Cognition and Culture*, *12*(3–4), 195–201.
<https://doi.org/10.1163/15685373-12342073>
- Awad, E., Dsouza, S., Shariff, A., Rahwan, I., & Bonnefon, J. F. (2019). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(5), 2332–2337.
<https://doi.org/10.1073/pnas.1911517117>
- BLOOM, P. (2011). FAMILY, COMMUNITY, TROLLEY PROBLEMS, AND THE CRISIS IN MORAL PSYCHOLOGY. *The Yale Review*, *99*(2), 26–43.
<https://doi.org/10.1111/j.1467-9736.2011.00701.x>
- CAO, J. (2009). The Analysis of Tendency of Transition from Collectivism to Individualism in China. *Cross-Cultural Communication*, *5*(4), 42–50.
<http://www.cscanada.net/index.php/cc/article/view/j.ccc.1923670020090504.005>
- Conway, P., & Gawronski, B. (2013a). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, *104*(2), 216–235. <https://doi.org/10.1037/a0031021>
- Conway, P., & Gawronski, B. (2013b). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, *104*(2), 216–235. <https://doi.org/10.1037/a0031021>
- Gold, N., Colman, A. M., & Pulford, B. D. (2014). Cultural differences in responses to real-life and hypothetical trolley problems. *Judgment and Decision Making*, *9*(1), 65–76.
<https://philarchive.org/archive/GOLCDI>

- Gold, N., Pulford, B. D., & Colman, A. M. (2013). Your money or your life: Comparing judgements in trolley problems involving economic and emotional harms, injury and death. *Economics and Philosophy*, *29*(2), 213–233. <https://doi.org/10.1017/S0266267113000205>
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. <https://doi.org/10.1037/a0021847>
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*(3), 364–371. <https://doi.org/10.1016/j.cognition.2009.02.001>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. <https://doi.org/10.1037//0033-295x.108.4.814>
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog? *Journal of Personality and Social Psychology*, *65*(4), 613–628. <https://doi.org/10.1037/0022-3514.65.4.613>
- Khazan. (2014). *Is One of the Most Popular Psychology Experiments Worthless ?* 1–4. <https://www.theatlantic.com/health/archive/2014/07/what-if-one-of-the-most-popular-experiments-in-psychology-is-worthless/374931/>
- Lee, M., Sul, S., & Kim, H. (2018). Social observation increases deontological judgments in moral dilemmas. *Evolution and Human Behavior*. <https://doi.org/10.1016/J.EVOLHUMBEHAV.2018.06.004>
- Navarrete, C. D., McDonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality: emotion

- and action in a simulated three-dimensional “trolley problem”. *Emotion (Washington, D.C.)*, 12(2), 364–370. <https://doi.org/10.1037/a0025561>
- O’Neill, P., & Petrinovich, L. (1998). A Preliminary Cross-Cultural Study of Moral Intuitions. *Evolution and Human Behavior*, 19(6), 349–367. [https://doi.org/10.1016/S1090-5138\(98\)00030-0](https://doi.org/10.1016/S1090-5138(98)00030-0)
- Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology*, 74, 24–37. <https://doi.org/10.1016/j.jesp.2017.08.003>
- Sachdeva, S., Singh, P., & Medin, D. (2011). Culture and the quest for universal principles in moral reasoning. In *International Journal of Psychology* (Vol. 46, Issue 3, pp. 161–176). <https://doi.org/10.1080/00207594.2011.568486>
- Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal*, 94. <https://doi.org/10.5505/tjob.2016.33042>
- Triandis, H. (1988). Collectivism v. Individualism: A Reconceptualisation of a Basic Concept in Cross-cultural Social Psychology. *Cross-Cultural Studies of Personality, Attitudes and Cognition*, 60–95. https://doi.org/10.1007/978-1-349-08120-2_3
- Triandis, H. (1992). Collectivism and Individualism as Cultural Syndromes. *Cross-Cultural Research*, 27, 155–180.
- Triandis, H. (2018). *Individualism And Collectivism*. Routledge. <https://books.google.com/books?hl=en&lr=&id=gwDFDwAAQBAJ&oi=fnd&pg=PP1&dq=collectivism+and+individualism&ots=xCDIBjlhbt&sig=u6uVlxF80jtyQzdBKWIQoEyayHQ#v=onepage&q=collectivism and individualism&f=false>

APPENDIX A

Himpathy Study: Sampling of Workplace Harassment Scenarios Presented to Participants

Basic Scenario Version – Perpetrator (Male or Female) & Victim (Male or Female, who reports the harassment either after 2 weeks or 1 year) are Co-Workers:

Marcus and Lauren are co-workers on the same team at a consulting firm, having been working together on a project for about six months. They hold equivalent positions within their firm.

Recently, however, Lauren has been feeling uncomfortable with Marcus. Lauren claims that Marcus has been repeatedly asking her out on dates, although she has made it clear that she is not interested, and she has consistently rejected him. Marcus has also been making sexual commentary on Lauren's body and outfits. Lauren has told Marcus that his comments make her uncomfortable, and has asked him to stop, and to keep their conversations professional. Despite this, Lauren continues to be harassed by Marcus in the workplace. Marcus engages in this behavior when he and Lauren are alone, and as such, there are no other witnesses to his actions.

Lauren ultimately decides to report his actions to the company's Human Resources department. However, the company does not investigate the charges against Marcus and takes no remedial action (i.e. he is not given a warning, his position in the company stays the same, and he is not fired from the company).

Given her company's lack of action and Marcus' continued harassment, Lauren leaves the company. After leaving, two weeks after Marcus' inappropriate workplace behavior towards her began, she decides to pursue legal action against both Marcus and her company. She files a sexual harassment suit against them, suing specifically for pain and suffering damages (the emotional distress the harassment caused), and back wages (the salary and commission she lost after leaving the company).

Testimony Statements from the Victim & Perpetrator, Co-Workers

VICTIM STATEMENT

Victim (either a male – Marcus, or female - Lauren), Perpetrator is a Co-Worker

I had been working at the company for a few months when I first felt like Marcus was acting inappropriately towards me. We were working on a project together, so naturally we spent a lot of time together. He would always try to hang out with me after work hours, asking me to get dinner and drinks and go to the movies. He also began to make inappropriate sexual comments about my outfits and body.

I repeatedly denied his offers to spend time with me after work, but he continued to extend them. I also repeatedly told him that his comments about me were inappropriate and made me uncomfortable, and asked him to stop. Nonetheless, his commentary continued. He would always wait until we were alone to say these things, so no one witnessed them.

I eventually couldn't take it anymore, and reported these actions to Human Resources. However, they completely failed to investigate my report, and did not take any remedial actions against Marcus.

The situation eventually became so hostile and unpleasant for me that I realized I needed to leave the company, despite the income I knew I would lose from my departure. I simply couldn't continue working somewhere that had caused me such discomfort and emotional turmoil.

During this time, two weeks after Marcus' initial harassment began, I decided to pursue legal action against him and my company, filing a sexual harassment suit against them.

PERPETRATOR STATEMENT

Perpetrator (either a male – Marcus, or female - Lauren), Victim is a Co-Worker

I've known Lauren the duration of time she worked at this company and we had always gotten along very well. I thought she was very smart, capable, and honest as an employee and coworker. I was excited to have the opportunity to work with her more closely when we were assigned to the same account.

I was so surprised to hear, however, that she had reported me to Human Resources for sexual harassment. I really had not interpreted my behavior as inappropriate or disrespectful -- I did enjoy her company and had wanted to spend more time with her to discuss the project, as it was a cumbersome and time consuming client, but I never had any ulterior motives. I didn't think that my behavior was inappropriate in any way, I treated her as I would treat any other coworker (male or female).

I occasionally had asked her to grab dinner and hang out outside of work because I wanted to hang out with her and be friends, but no romantic feelings were involved. I respect her as a co-worker and I think she mistook my compliments to be offensive and demeaning.

There is literally no proof of her allegations except her word against mine. These allegations are completely false -- I'm just not that kind of person, ask anyone. On top of her report to Human Resources, she is now also suing me, further tainting my record. My life will be ruined - my family depends on me to support us financially - I don't know where I would go from here.

Basic Scenario Version – Perpetrator (Male or Female) & Victim (Male or Female, who reports the harassment either after 2 weeks or 1 year); Perpetrator is the Victim's Work Supervisor

Marcus is the lead project supervisor at the consulting firm Lauren works at, and has been supervising Lauren's work for the past six months.

Recently, however, Lauren has been feeling uncomfortable with Marcus. Lauren claims that Marcus has been repeatedly asking her out on dates, although she has made it clear that she is not interested, and she has consistently rejected him. Marcus has also been making sexual

commentary on Lauren's body and outfits. Lauren has told Marcus that his comments make her uncomfortable, and has asked him to stop, and to keep their conversations professional. Despite this, Lauren continues to be harassed by Marcus in the workplace. Marcus engages in this behavior when he and Lauren are alone, and as such, there are no other witnesses to his actions.

Lauren ultimately decides to report his actions to the company's Human Resources department. However, the company does not investigate the charges against Marcus and takes no remedial action (i.e. he is not given a warning, his position in the company stays the same, and he is not fired from the company).

Given her company's lack of action and Marcus' continued harassment, Lauren leaves the company. After leaving, two weeks after Marcus' inappropriate workplace behavior towards her began, she decides to pursue legal action against both Marcus and her company. She files a sexual harassment suit against them, suing specifically for pain and suffering damages (the emotional distress the harassment caused), and back wages (the salary and commission she lost after leaving the company).

Testimony Statements from the Victim & Perpetrator, Project Manager & Employee

VICTIM STATEMENT (Victim is an employee under the perpetrator, perpetrator is the victim's work superior):

Victim (either a male – Marcus, or female - Lauren), Perpetrator is a Work Superior:

I had been working at the company for a few months when I first felt like Marcus was acting inappropriately towards me. He was my project manager and was supervising my work on a very time consuming project, so naturally, we spent a lot of time together. He would always try to hang out with me after work hours, asking me to get dinner and drinks and go to the movies. He also began to make inappropriate sexual comments about my outfits and body.

I repeatedly denied his offers to spend time with me after work, but he continued to extend them. I also repeatedly told him that his comments about me were inappropriate and made me uncomfortable, and asked him to stop. Nonetheless, his commentary continued. He would always wait until we were alone to say these things, so no one witnessed them.

I eventually couldn't take it anymore, and reported these actions to Human Resources. However, they completely failed to investigate my report, and did not take any remedial actions against Marcus.

The situation eventually became so hostile and unpleasant for me that I realized I needed to leave the company, despite the income I knew I would lose from my departure. I simply couldn't continue working somewhere that had caused me such discomfort and emotional turmoil.

During this time, two weeks after Marcus' initial harassment began, I decided to pursue legal action against him and my company, filing a sexual harassment suit against them.

PERPETRATOR STATEMENT

Perpetrator (either a male - Marcus or female - Lauren), Perpetrator is a Work Superior:

I've known Lauren the duration of time she worked at this company and we had always gotten along very well. I thought she was very smart, capable, and honest as an employee, and as her project manager, I thought her work quality was excellent. I was excited to have the opportunity to work with her more closely when I was assigned to be her project manager.

I was so surprised to hear, however, that she had reported me to Human Resources for sexual harassment. I really had not interpreted my behavior as inappropriate or disrespectful -- I did enjoy her company and had wanted to spend more time with her to discuss the project, as it was a cumbersome and time consuming client, but I never had any ulterior motives. I didn't think that my behavior was inappropriate in any way, I treated her as I would treat any other employee I was supervising (male or female).

I occasionally had asked her to grab dinner and hang out outside of work because I wanted to hang out with her and be friends, but no romantic feelings were involved. I respect her as a company employee and as someone I supervise, and I think she mistook my compliments to be offensive and demeaning.

There is literally no proof of her allegations except her word against mine. These allegations are completely false -- I'm just not that kind of person, ask anyone. On top of her report to Human Resources, she is now also suing me, further tainting my record. My life will be ruined - my family depends on me to support us financially - I don't know where I would go from here.

APPENDIX B

Qualitative Coding Schema applied to Mock Jury Deliberations

Node	Coding	Affirmative Examples
<u>Plaintiff Injury Compare</u>	<p>0 = No comparison by any juror regarding the severity of the two plaintiffs' injuries/ how they were affected by their injuries</p> <p>1+ = Jurors compare the severity of the two plaintiffs' injuries/ how they were affected by their injuries</p>	<p>1: Alright. And then for, uhh, the pain and suffering for Susan Kessler. The one with the more severe impact.</p> <p>1: Yeah, Ms. Kessler's where it's chronic, and [Ms. Franklin's]</p> <p>Note: This code will likely overlap with SK/AF Bad or SK/AF Long, since the comparison will often be about the severity/length of their injuries (i.e. if you code something as Plaintiff Compare, it will likely also be coded as Bad/Long, or interfering; but not vice versa)</p>
<u>Plaintiff Award Compare</u>	<p>0 = No comparison by any juror regarding how the award for the two plaintiffs would or should compare (e.g. one should be proportionally smaller or larger)</p> <p>1+ = Jurors mention how the award for the two plaintiffs would or should compare (e.g. one should be proportionally smaller or larger)</p>	<p>2: [Int 5: Right, you could] see her getting a higher award, but award for her pain and [suffering, and] a lower reward for what she's lost, uh, in wages and such. Which is true, she has returned more or less to her previous career, even though the other woman cannot do so.</p> <p>4: [Int 1: Yes, yes.] If you think it's fair for the other one to get half, five million, then the other one we're not giving her enough.</p> <p style="text-align: center;"><i>**This example would also fit into one of the award-specific codes, since a specific dollar amount is mentioned**</i></p> <p>Note: Within their description of how the plaintiffs' awards should compare, jurors may mention a specific dollar amount. If so, this code, as well the award-specific codes (e.g. SK Pain & Suffering Award, SK Lost Award, etc.) should be applied to the text.</p>
<u>Plaintiff Compare – Other</u>	<p>0 = None of the jurors make a comparison statement between the two plaintiffs that is either general, or is not related to comparing the plaintiffs' injuries or award amounts (e.g. comparing their personal experiences).</p> <p>1+ = Jurors make a comparison statement between the two plaintiffs that is either general, or is not related to comparing the plaintiffs' injuries or award amounts (e.g. comparing their personal experiences).</p>	<p>3: She's obviously a lot better off than the first one.</p> <p>4: Yes but she's getting better and [better. She] doesn't have the life, I mean the two don't match anywhere</p> <p>5: Um, but then, yeah. And then, for, for Susan Kessler, it sounds like all her relationships are just worse, [and probably] will remain that way.</p>
<u>Malingering</u>	<p>Jurors mention possibility of malingering or attributes it to either plaintiff</p>	<p>I think that first lady was faking it for sure. The fact you can remember some things, but not others, like, that's definitely not how it works.</p> <p>I don't know if she was a bad actor or if she was malingering, like the doctor suggested.</p>

Node	Coding	Affirmative Examples
<u>SKBad</u>	<p>0 = No mention by any juror of how bad or severe the Kessler's injuries are</p> <p>1+ = Jurors mention how bad or severe Kessler's injuries are</p>	<p>She was the plaintiff who is more badly injured, right?</p> <p>Well when you lose your brain, that's a pretty significant bodily injury, I mean it- it affects everything</p> <p>She's not normal.</p> <p>Note: Characterization as "bad/not bad" should be in reference to her injuries, not the consequences of her injuries</p>
<u>SKLong</u>	<p>0 = No mention by any juror of how long the injuries will last/affect Kessler's life</p> <p>1+ = Jurors mention how long the injuries will last/affect the Kessler's life</p>	<p>Oh no it says for the next thirty years that she can't work ... she's not going back to work.</p> <p>She doesn't have the full use of her brain, basically, I mean for the rest of her life.</p> <p>She's never going to recover.</p> <p>She's never going to be the same.</p>
<u>SKInterfering</u>	<p>0 = No mention by any juror of how much the injury interferes with the Kessler's life</p> <p>1+ = Jurors mention how much the injury interferes with the Kessler's life (e.g., ability to attend work or school)</p>	<p>I can see why she has lower quality of life.</p> <p>Note: This code should be applied to general statements about the injury's interference (not related to social/family/spouse)</p>
<u>SKSocial Interfering</u>	<p>0 = No mention by any juror of how much the injuries interfere with the Kessler's <u>social life</u></p> <p>1+ = Jurors mention how much the injuries interfere with the Kessler's <u>social life</u> (e.g., ability socialize with friends)</p>	<p>She can't drive, so she can't visit friends. And she gets migraines, so she really can't be around people.</p>
<u>SKSpouse</u>	<p>0 = No mention by any juror of how Kessler's injuries/accident will affect her spouse's life</p> <p>1+ = Jurors explicitly mention how Kessler's injuries/accident will affect her spouse's life (he will need to work longer, make up for loss)</p>	<p>He's gonna have a hard time now that she can't do anything.</p> <p>Note: This code is looking at the general impact of her injuries on her <i>spouse</i> (only her husband, exclusive of her family)</p>
<u>SKFamily Interfering</u>	<p>0 = No mention by any juror of how much the injuries interfere with the Kessler's <u>family life</u></p> <p>1+ = Jurors mention how much the injuries interfere with the Kessler's <u>family life</u> (e.g., ability to attend kids events; go on dates)</p>	<p>That to me is the pain and suffering in this- is she's a parent of three young children. Imagine your parent not being able to parent you.</p> <p>Just, well it's affected her relationship with her husband, with her children, with her friends, you know, it's all encompassing.</p> <p>She could lose her marriage because of this.</p> <p>Note: This code is NOT looking at the monetary impact of her injuries on her family unit; this only refers to the non-financial effects</p>
<u>SKFamily Financial</u>	<p>0 = No mention by any juror of how Kessler's loss of income will financially affect her family unit</p> <p>1+ = Jurors explicitly mention how Kessler's loss of income will financially affect her family unit</p>	<p>Note: This code is looking at the monetary impact of her injuries on her <i>family unit</i></p>

Node	Coding	Affirmative Examples
<u>SKAge</u>	<p>0 = No mention by any juror of Kessler's age</p> <p>1+ = Mention of Kessler's age as either verbatim or gist (35 years; in her thirties; young; retirement)</p>	<p>She was 35, I think.</p> <p>But she was so young, like she could still be working if it weren't for the accident.</p> <p>Note: This code should capture any mention of her age and/or how her age is a factor in how much money they give (don't worry about gist/verbatim).</p>
<u>SKFirst P&SNum</u>	<p>1 – 6 = Juror number of first juror who proposes an award amount for Kessler's Pain and Suffering award</p>	
<u>SKLostWages Award</u>	<p>0 = No mention by any juror of a specific dollar award pertaining to Kessler's lost wages</p> <p>1+ = Mention of a specific dollar award pertaining to Kessler's lost wages (not the anchor)</p>	
<u>SKPain&Suffering Award</u>	<p>0 = No mention by any juror of a specific dollar award pertaining to Kessler's pain and suffering award</p> <p>1+ = Mention of a specific dollar award pertaining to Kessler's pain and suffering award (not the anchor)</p>	<p>2: And then what, 2 million for pain and suffering?</p>
<u>SKLostWages Anchor</u>	<p>0 = No mention by any juror of the lost wages salary anchor provided by Kessler's attorney</p> <p>1+ = Mention of the lost wages salary anchor provided by Kessler's attorney (\$50,582; \$53,000; about \$53,000)</p>	
<u>SKPain&Suffering Anchor</u>	<p>0 = No mention by any juror of the pain & suffering anchor provided by Kessler's attorney</p> <p>1+ = Mention of the pain & suffering anchor provided by Kessler's attorney (\$4,204,800; 4 million; about 4.2 mil)</p>	
<u>SKTotalAward</u>	<p>0 = No juror proposes a total damage award for Susan Kessler</p> <p>1+ = Any juror proposes a total damage award for Susan Kessler</p>	<p>3: So, you're thinking 2.5 total? And you're thinking 5?</p> <p>3: And then for the other person I put... some total that was around- it was like a hundred thousand? Yeah. I think.</p> <p>Note: This is a number that is not separated into its lost wages and/or pain & suffering components. It needs to be a specific statement about the numeric overall, total award (P&S + Lost Wages).</p>
<u>AFBad</u>	<p>0 = No mention by any juror of how bad or severe the Franklin's injuries are</p> <p>1+ = Jurors mention how bad or severe the Franklin's injuries are</p>	<p>Her injuries weren't as bad.</p> <p>She was the plaintiff who was less injured.</p> <p>Yeah, but she's only- she only suffered minimal, right?</p> <p>The pain and suffering were minimal, the duration was minimal, the interference was minimal.</p>
<u>AFLong</u>	<p>0 = No mention by any juror of how long the injury will last/affect the Franklin's life</p> <p>1+ = Jurors mention how long the injury will last/affect the Franklin's life</p>	<p>Two years is a long time for recovery.</p> <p>She was able to actually recover from it after a few years, unlike the other lady.</p> <p>I think she'll be back to normal soon.</p>

Node	Coding	Affirmative Examples
<u>AFInterfering</u>	<p>0 = No mention by any juror of how much the injury interferes with the Franklin's life</p> <p>1+ = Jurors mention how much the injury interferes with the Franklin's life (e.g., ability to attend work or school)</p>	<p>It clearly just didn't impact her as much, not nearly on the same level.</p> <p>*Note: This code should be applied to general statements about the injury's interference (not related to social/family/spouse)</p>
<u>AFSocial Interfering</u>	<p>0 = No mention by any juror of how much the injuries interfere with the Franklin's <u>social life</u></p> <p>1+ = Jurors mention how much the injuries interfere with the Franklin's <u>social life</u> (e.g., ability to attend work or socialize with friends)</p>	
<u>AFSpouse</u>	<p>0 = No mention by any juror of how Franklin's injuries/accident will affect her spouse's life</p> <p>1+ = Jurors explicitly mention how Franklin's injuries/accident will affect her spouse's life (he will need to work longer, make up for loss)</p>	<p>She's gonna need to be taken care of now.</p> <p>Note: This code is looking at the general impact of her injuries on her <i>spouse</i> (only her husband, exclusive of her family)</p>
<u>AFFamily Interfering</u>	<p>0 = No mention by any juror of how much the injuries interfere with the Franklin's <u>family life</u></p> <p>1+ = Jurors mention how much the injuries interfere with the Franklin's <u>family life</u> (e.g., ability to attend kids events; go on dates)</p>	<p>Note: This code is NOT looking at the monetary impact of her injuries on her family unit; this only refers to the non-financial effects</p>
<u>AFFamily Financial</u>	<p>0 = No mention by any juror of how Franklin's loss of income will financially affect her family unit</p> <p>1+ = Jurors explicitly mention how Franklin's loss of income will financially affect her family unit</p>	<p>Note: This code is looking at the monetary impact of her injuries on her <i>family unit</i></p>
<u>AFAge</u>	<p>0 = No mention by any juror of Franklin's age</p> <p>1+ = Mention of Franklin's age as either verbatim or gist (33 years; in her thirties; young; retirement)</p>	<p>She has so much more life to live.</p> <p>She's not gonna retire anytime soon, so I feel like we should give her at least 5 million.</p> <p>Note: This code should capture any mention of her age and/or how her age is a factor in how much money they give (don't worry about gist/verbatim)</p>
<u>AFFirst P&SNum</u>	<p>1 – 6 = Juror number of first juror who proposes an award amount for Franklin's Pain and Suffering award</p>	
<u>AFLostWages Award</u>	<p>0 = No mention by any juror of a specific dollar award pertaining to Franklin's lost wages</p> <p>1+ = Mention of a specific dollar award pertaining to Franklin's lost wages (not the anchor)</p>	
<u>AFPain&Suffering Award</u>	<p>0 = No mention by any juror of a specific dollar award pertaining to Franklin's pain and suffering award</p> <p>1+ = Mention of a specific dollar award pertaining to Franklin's pain and suffering award (not the anchor)</p>	<p>2: And then what, 2 million for pain and suffering?</p>
<u>AFLostWages Anchor</u>	<p>0 = No mention by any juror of the lost wages salary anchor provided by Franklin's attorney</p> <p>1+ = Mention of the lost wages salary anchor provided by Franklin's attorney (\$52,804; about \$53,000; \$52,000)</p>	

Node	Coding	Affirmative Examples
<u>AFPain&Suffering Anchor</u>	<p>0 = No mention by any juror of the pain & suffering anchor provided by Franklin's attorney</p> <p>1+ = Mention of the pain & suffering anchor provided by Franklin's attorney (\$175,200; \$175,000; about \$175K)</p>	
<u>AFTotalAward</u>	<p>0 = No juror proposes a total damage award for Ashley Franklin</p> <p>1+ = Any juror proposes a total damage award for Ashley Franklin</p>	<p>Note: This is a number that is not separated into its lost wages and/or pain & suffering components. It needs to be a specific statement about the numeric overall, total award (P&S + Lost Wages).</p>
<u>HourlyAnchor</u>	<p>0 = No explicit mention by any juror of \$10/hour anchor provided by the plaintiff attorney</p> <p>1+ = Explicit mention by any juror of \$10/hour anchor provided by the plaintiff attorney</p>	<p>I believe I wrote down in the closing remarks, the plaintiff asked for ten dollars an hour, is that right?</p> <p>Note: Reserved for just the mention of the anchor. Statement has no other characterization about the anchor.</p>
<u>HourlyAnchor Significance</u>	<p>0 = No explicit mention of \$10/hour anchor's <u>(in)significance</u></p> <p>1+ = Explicit mention of anchor's <u>(in)significance</u> (e.g., meaningful, important, significant, irrelevant, arbitrary)</p>	<p>4: [Ten] dollars an hour for the rest of her life would give her that amount. And I was kinda like... I dunno. That seemed like an arbitrary thing.</p> <p>Note: Reserved for explicit characterization of the anchor. (If coded as significant, it should likely also be coded as mentioned, but not vice versa)</p>
<u>LifeExpectancy</u>	<p>0 = No explicit mention by any juror of life expectancy anchor provided by the neurosurgeon witness</p> <p>1+ = Explicit mention by any juror of life expectancy anchor provided by the neurosurgeon witness, as either verbatim or gist (81 years; around 80 years)</p>	<p>We heard that the, um, average life of a woman-- which was very vague-- was eighty-one.</p>
<u>Fairness</u>	<p>0 = No mention by any juror of the award being fair, right, or accurate amount</p> <p>1+ = Jurors mention the award being a fair, right, or accurate amount</p>	<p>So, I feel like that's pretty fair because we've discussed, you know, the considerations of like social security and everything else.</p> <p>As for the amount it is what I would consider fair</p> <p>Note: This code refers to a juror's assessment of the moral/philosophical fairness of an award; not an agreement</p>
<u>LegalCosts</u>	<p>0 = No mention by any juror of legal costs/attorney fees/contingency fees associated with damage award</p> <p>1+ = Jurors mention legal costs associated with damage award</p>	<p>One third of the money will go to the attorney.</p> <p>I didn't take into consideration that-- um-- money would be going to pay legal attorney fees, uh other legal costs.</p>
<u>LegalCostsAdded</u>	<p>0 = Legal costs were mentioned but had no effect on the discussion of final award</p> <p>1 = Legal costs were mentioned AND had an effect on the discussion of final award (i.e. award was raised, lowered, split differently)</p>	<p>6: We could round it up to a [hundred] to make up her [lawyer fees].</p> <p>3: [Int 6: Alright].</p> <p>4: [Int 6: Yeah to] make up for her lawyer's fees, sounds good.</p> <p>Note: the code is binary (yes or no). It requires a holistic review of the entire transcript to see if they did end up changing their award to account for legal costs</p>

Node	Coding	Affirmative Examples
<u>Insurance</u>	<p>0 = No mention by any juror about the insurance of the defendants</p> <p>1+ = Jurors have some discussion about the insurance of the defendants (e.g. how it will/will not pay for the award or whether the defendants have any)</p>	<p>4: Well, one thing I think we should dedicate some discussion to is how these damage awards will be paid for, like, do- the defendants have insurance right?</p>
<u>Speculation</u>	<p>0 = No mention by any juror of speculating into the future</p> <p>1+ = Jurors mention speculation in the discussion of the plaintiffs' injuries/awards</p>	<p>I think we are only supposed to consider the evidence presented. We shouldn't speculate into the future about things we don't know.</p>
<u>Award Confusion</u>	<p>0 = No mention by any juror of being confused about coming up with any award</p> <p>1+ = Jurors mention being confused about arriving at any damage award</p>	<p>5: I don't know if lost wages only is retroactive</p> <p>I had some difficulty coming up with the award amount because it's so open ended and there's no specific range or amount stated.</p> <p>What is pain and suffering?</p> <p>Note: This code was amended to only include confusion about coming up with any award. We no longer care about general confusion about the deliberation. Use this code sparingly, not every sentence ending in a question mark should be considered Award Confusion</p>
<u>Judge</u>	<p>0 = No jurors mention the judge or judicial instructions</p> <p>1+ = Jurors explicitly mention the judge or judicial instructions (including their like or dislike of them)</p>	<p>I wish the judge gave us more instruction.</p> <p>I wrote it down. We have to use the "greater weight of evidence."</p> <p>I mean so by- by the definition of the judge, reasonable speed, he, he did not take reasonable – he, I mean he was going ten miles over the speed limit.</p>
<u>Personal Experience</u>	<p>0 = No jurors mention their personal connection/ experiences related to the case</p> <p>1+ = One or more jurors mentions their personal connection/experience related to the case (e.g. lawsuit, car crash, medical experience)</p>	<p>I've been through it and I got hit it by a truck driver who was high on Chrystal meth.</p> <p>Um, I don't really have any experience with accidents at all but I have a friend who's uh won a medical malpractice lawsuit because his grandfather, um- passed away in surgery due to like negligent- like serious negligence.</p>
<u>PResponsible</u>	<p>0 = No juror believes the plaintiffs were (not) liable or responsible (e.g., knowing the risks, should have known better)</p> <p>1 – 6 = Juror number of juror who believes the plaintiffs' responsibility or lack thereof</p>	<p>I agree with the fact that she was in the wrong place at the wrong time.</p> <p>I thought that the defense even knew that the plaintiff was in the right.</p>
<u>DriverResponsible</u>	<p>0 = No mention by any juror of the driver being liable or responsible</p> <p>1 – 6 = Juror number of juror who mentioned the driver's liability or responsibility (or lack thereof) for the injuries</p>	<p>There's no evidence whatsoever provided that he did anything.</p> <p>The other thing that I'm also strongly looking at is the, um, complete irresponsibility on the part of the defendant, both the driver and the company.</p> <p>[Note: this would be coded under both Driver Responsible and Company Responsible]</p>

Node	Coding	Affirmative Examples
<u>Company Liability</u>	<p>0 = No Jurors express (lack of) liability on the part of the Macklin Furniture Company, the defendant's employer.</p> <p>1 – 6 = Juror number of juror who expresses (lack of) liability on the part of the Macklin Furniture Company, the defendant's employer.</p>	<p>The company was totally negligent.</p> <p>The other thing that I'm also strongly looking at is the, um, complete irresponsibility on the part of the defendant, both the driver and the company.</p> <p>[Note: this would be coded under both Driver Responsible and Company Responsible]</p>
<u>Punitive</u>	<p>0 = No mention by any juror about the damage award being a punitive measure</p> <p>1+ = Jurors mention the damage award being a punitive measure (e.g. send a signal, punish recklessness)</p>	
<u>GoodQuote</u>	Code any quote that is particularly meaningful or interesting that can be used for the future publication of this work	
<u>JurorQuestions</u>	Code any questions related to the jurors' confusion about the procedure of the deliberation (i.e. things unrelated to the trial)	
<u>NumberPerson</u>	<p>0 = No juror was assigned to be the group "number person"</p> <p>1 – 6 = Juror number of juror assigned by group to be "number person" or "calculator" (i.e. in charge of calculating for the group)</p>	
<u>HandCalculate</u>	<p>0 = No juror calculated a damage award by hand/on paper</p> <p>1 – 6 = Juror number of juror who calculated a damage award by hand/on paper</p>	<p>6: I figured, [it's - it's free money anyway, right?]</p> <p>4: [Int 6: Three fifty, that's five hundred.] (calculating on paper)</p>
<u>PhoneCalculate</u>	<p>0 = No juror calculated a damage award on their phone/calculator</p> <p>1 – 6 = Juror number of juror who calculated a damage award on their phone/calculator</p>	<p>3: So, let's see... (Juror 3 uses phone calculator)</p>
<u>MathComfort</u>	<p>0 = No mention by any juror about their level of comfort with math</p> <p>1 – 6 = Juror number of juror who mentioned their level of comfort with math (bad, good, rusty, hate)</p>	<p>Umm, I'm not a math person, sorry, by far. I hate math.</p> <p>I wish I had my calculator. I took it out of my backpack.</p>

APPENDIX C

Qualitative Variables and Interrater Reliability for Jury B

Code	Cohen's Kappa	Agreement (%)
Ashley Franklin Award Numbers\AF Lost Wages Anchor	0.71	99.66
Plaintiff Compare\Plaintiff Injury Compare	0.75	99.71
Driver Responsible\J4 DResponsible	0.76	99.52
Personal Experience	0.77	98.25
Judge	0.83	99.25
Susan Kessler Commentary\SK Long	0.84	99.04
Speculation	0.85	99.8
Susan Kessler Commentary\SK Family Interfering	0.87	99.8
Driver Responsible	0.88	99.49
Susan Kessler Commentary\SK Age	0.88	99.75
Susan Kessler Award Numbers\SK Lost Wages Award	0.91	99.45
Ashley Franklin Award Numbers\AF Total Award	0.91	99.97
Susan Kessler Commentary	0.91	98.96
Ashley Franklin Commentary\AF Interfering	0.92	99.81
Susan Kessler Commentary\SK Family Financial	0.93	99.92
Ashley Franklin Commentary\AF Long	0.94	99.69
Plaintiff Compare\Plaintiff Compare - Other	0.94	99.66
Plaintiff Compare\Plaintiff Award Compare	0.94	99.48
Company Liability	0.95	99.82
Ashley Franklin Award Numbers\AF Pain & Suffering Award	0.95	99.37
Plaintiff Compare	0.96	99.47
Susan Kessler Award Numbers\SK Pain & Suffering Award	0.96	99.4
Susan Kessler Award Numbers\SK First P&S Number\J3 SK First P&S Number	0.96	100
Ashley Franklin Award Numbers	0.96	99.26
Susan Kessler Award Numbers	0.96	99.29
Ashley Franklin Commentary\AF Age	0.97	100
Ashley Franklin Commentary	0.97	99.65
Susan Kessler Award Numbers\SK Total Award	0.97	100
Ashley Franklin Award Numbers\AF First P&S Number\J3 AF First P&S Number	0.97	100
Malingering	0.97	100
Driver Responsible\J3 DResponsible	0.98	99.98

Code	Cohen's Kappa	Agreement (%)
LegalCostsAdded	0.98	99.99
Insurance	0.98	100
Susan Kessler Commentary\SK Bad	0.98	99.94
Ashley Franklin Award Numbers\AF Lost Wages Award	0.99	99.93
Ashley Franklin Commentary\AF Bad	0.99	99.95
Driver Responsible\J1 DResponsible	0.99	100
Punitive	0.99	99.99
Fairness	0.99	99.98
Math Comfort	0.99	100
Math Comfort\J1 Math Comfort	0.99	100
Susan Kessler Commentary\SK Interfering	0.99	99.99
Juror Questions	1.00	100
Legal Costs	1.00	99.99
Susan Kessler Award Numbers\SK Lost Wages Anchor	1.00	99.99
Driver Responsible\J2 DResponsible	1.00	100
Ashley Franklin Commentary\AF Social Interfering	1.00	100
Ashley Franklin Commentary\AF Family Interfering	1.00	100
Susan Kessler Commentary\SK Social Interfering	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number\J1 AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number\J2 AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number\J4 AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number\J5 AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF First P&S Number\J6 AF First P&S Number	1.00	100
Ashley Franklin Award Numbers\AF Pain & Suffering Anchor	1.00	100
Ashley Franklin Commentary\AF Family Financial	1.00	100
Ashley Franklin Commentary\AF Spouse	1.00	100
Award Confusion	1.00	100
Driver Responsible\J5 DResponsible	1.00	100
Driver Responsible\J6 DResponsible	1.00	100
Good Quote	1.00	100
Hand Calculate Juror	1.00	100
Hand Calculate Juror\Hand Calculate J1	1.00	100
Hand Calculate Juror\Hand Calculate J2	1.00	100

Code	Cohen's Kappa	Agreement (%)
Hand Calculate Juror\Hand Calculate J3	1.00	100
Hand Calculate Juror\Hand Calculate J4	1.00	100
Hand Calculate Juror\Number Person J5	1.00	100
Hand Calculate Juror\Number Person J6	1.00	100
Hourly Anchor	1.00	100
Hourly Anchor Significance	1.00	100
Length Defendant	1.00	100
Length Franklin	1.00	100
Length Kessler	1.00	100
Life Expectancy	1.00	100
Math Comfort\J2 Math Comfort	1.00	100
Math Comfort\J3 Math Comfort	1.00	100
Math Comfort\J4 Math Comfort	1.00	100
Math Comfort\J5 Math Comfort	1.00	100
Math Comfort\J6 Math Comfort	1.00	100
Number Person Juror	1.00	100
Number Person Juror\Number Person J1	1.00	100
Number Person Juror\Number Person J2	1.00	100
Number Person Juror\Number Person J3	1.00	100
Number Person Juror\Number Person J4	1.00	100
Number Person Juror\Number Person J5	1.00	100
Number Person Juror\Number Person J6	1.00	100
Phone Calculate Juror	1.00	100
Phone Calculate Juror\Phone Calculate J1	1.00	100
Phone Calculate Juror\Phone Calculate J2	1.00	100
Phone Calculate Juror\Phone Calculate J3	1.00	100
Phone Calculate Juror\Phone Calculate J4	1.00	100
Phone Calculate Juror\Phone Calculate J5	1.00	100
Phone Calculate Juror\Phone Calculate J6	1.00	100
Plaintiff Responsible	1.00	100
Plaintiff Responsible\J1 PResponsible	1.00	100
Plaintiff Responsible\J2 PResponsible	1.00	100
Plaintiff Responsible\J3 PResponsible	1.00	100
Plaintiff Responsible\J4 PResponsible	1.00	100

Code	Cohen's Kappa	Agreement (%)
Plaintiff Responsible\J5 PResponsible	1.00	100
Plaintiff Responsible\J6 PResponsible	1.00	100
Susan Kessler Award Numbers\SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK First P&S Number\J1 SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK First P&S Number\J2 SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK First P&S Number\J4 SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK First P&S Number\J5 SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK First P&S Number\J6 SK First P&S Number	1.00	100
Susan Kessler Award Numbers\SK Pain & Suffering Anchor	1.00	100
Susan Kessler Commentary\SK Spouse	1.00	100

APPENDIX D

Qualitative Coding Schema Applied to US Participants' Open-Ended Explanations for their Moral Decisions

Deontological:

- **Moral rights/rules/reasoning**
 - Ex: “killing is bad”; “family is family”
 - Ex: “Although I love and care deeply about the person the train is headed for, pulling the lever would kill another human being and that would be on me.”
 - Ex: “It's difficult to weigh the lives of individuals and especially if those you know and love to those who are strangers to you. But at the same time those strangers to you may not be strangers to others. In general it's hard to weigh people's lives and judge them to make the "right" decision. Because sometimes there isn't a "right" decision.”
 - Ex: “One life is as valuable as another, so who am I to choose who lives and who dies?”
- **Responsibility to protect children**
 - Ex: “Feeling responsible for bringing them into the world and the responsibility to go to any length to protect them, I would probably "pull the lever" on many more people to save my child.”

Consequentialist:

Consequence of the action matters

- Ex: “Adults going to work - I sympathize with the fact that they need to be on time. It could jeopardize their job, a pay raise they might otherwise get... and keeping a good status at work is really important to their whole livelihood. ”
- Ex: “However, I could not justify saving even the person I care most about when this would require the sacrifice of 100 or more strangers with lives of their own, contributions to society, and people who love them.”
- **Utilitarianist (i.e. creating the most good for the most people)**
 - **Weighing the consequences**
 - “Such as if it was 1 Bill Gates or 5 pedophiles. You can do a **cost-benefit** on society and how many lives you are actually helping or hurting by the decision”
 - **Numerical consideration of creating greatest good**

Description: These respondents often involve saying that they would save the most people, the greater amount of property getting damaged, or the greatest number of people getting hurt, in order to create the greatest good.

 - E.g. “For the rest of the scenarios I believe that saving the maximum amount of lives possible is the most important.”
 - E.g. “However, when there are more people involved in the situation, I would pick the choice where I can save more people.”
- **Egoist**
 - *Description: lack of remorse; creating the most good for one's self; I only care about myself, don't consider others*

- "...since I just made my decision based on the importance of each person to me."

Non-specific numerical considerations:

- *Description: these are statements where the number of people involved impacted respondents' decisions, but the reasoning (moral, emotional, etc.) behind it is unclear*
 - E.g. "Obviously, the higher the number of strangers, the more difficult it was to make the decision."
 - E.g. "The number of people killed impacted my decision. If it was only one person then I wanted to save my loved one but I would not save them if it meant killing more people."

Emotional Responses

Note: These are all emotions that respondents expressed that influenced or went into their decisions, or how they would feel after making a particular decision.

- Guilt / regret
 - Ex: "Even if I were to save the person I love in these scenarios, the remainder of both of our lives would likely be **ruined by guilt.**"
 - Ex: "I wouldn't kill my loved one because **I couldn't live with that** but a death of 1,000 strangers wouldn't affect me")
 - Ex. "**I regret** all four of my decisions now"
- Feeling selfish
 - Ex: "I feel like a selfish person. I feel like my choice is the wrong choice but it is what I would likely do"
 - Ex: "Eventually, I was forced to choose the option of saving strangers when I felt that I was being selfish by choosing the person I care about at the expense of SO many valuable lives."
- Sadness
- Anger
- Love and Cherish your loved one
 - Ex: " I believe in that short amount of time given to make a decision we are ruled not by our rational minds but by our emotions and **upon seeing somebody we love most in the world we will do anything to save them**, including killing any number of strangers."
 - Ex: "This was very difficult, as the person I love is extremely important to me and I do not want anything to ever happen to them.
 - Ex: "I taught about how that one person is my meaning of existence."
- Non-specific emotion
 - Ex. Feeling pain; blame; empathy

Perspective-Taking

- **Loved one's Wants, Reactions, Beliefs**
 - Ex: "The difficulty was very difficult because I wanted to save the person but I'm not sure if that person would rather me save the other people."
- **Thinking about the loved ones of strangers**

- Ex: “I know it would give the strangers' loved ones immense pain to know that they died”

Hard to Imagine the Loss

- Ex: “I felt that because this is unrealistic I just had to block out the decision of killing people. I also thought about how a train would probably not kill all 100 or 1000 people. ”
- Ex: “Then the picture of killing 1000 people for a second did not touch my heart because I did not see their faces and wanted to choose my loved one.”

Emotional and Economic Loss:

- **Teachable Moment**

- Ex: “A child experiencing challenges is a teaching moment. Particularly if it were my child, I would want to help them learn to work through the frustration and disappointment. ”

- **Personal Responsibility for Property**

- Ex: “Saving the cell-phone in the locker is a relatively easy decision. While it's certainly not pleasant to break your phone, I am not financially responsible for replacing the property of the 1, 10, 100, and 1000 people that placed their phone in the other locker. At this point in society, people are prepared to break their phone because it happens so often; if I can save myself a couple hundred bucks by saving my kid's phone, that'll be worth it because it's my direct responsibility. The other people's phones being damaged is just a risk they take when they store their items in a public-ish place.”

- **Non-equivalencies:**

Definition: the consequences for each action differ

- **Adult going to work vs. child going to school**

- Ex: “I know how important it is for my child to go to school, but I know it doesn't compare to the repercussions that an adult can face if they were late to school.”

- **Janitor vs. passerby**

- **Not Life vs. Life**

- Ex: “In the second scenario, we're talking about inanimate objects. Therefore these decisions were not that difficult to make. ”
“I felt this was significantly easier than the previous question as there was less at stake (e.g. lives were not lost).”

Non-answers