A Dissertation

entitled

Behavior in Situations Simulating the Tragedy of the Commons is Predicted by Moral

Judgment

by

Evan Clarkson

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the

Doctor of Philosophy Degree in Experimental Psychology

_____
John Jasper, Ph.D., Committee Chair

_____
Stephen Christman, Ph.D., Committee Member

_____
Jason Rose, Ph.D., Committee Member

_____
Andrew Geers, Ph.D., Committee Member

_____
John Sarnecki, Ph.D., Committee Member

_____
Amy Thompson, Ph.D., Dean
College of Graduate Studies

The University of Toledo

May 2022

An Abstract of

Behavior in Situations Simulating the Tragedy of the Commons is Predicted by Moral Judgment

by

Evan M Clarkson

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the
Doctor of Philosophy Degree in
Experimental Psychology

The University of Toledo
May 2022

Philosophers have long questioned the nature of morality and our moral sensibilities. Definitive answers have been difficult to come by. In response, psychologists have gathered mountains of empirical data to understand how these sensibilities are realized as moral beliefs or as moral judgments. Traditionally, researchers have required people to provide judgments about what is morally appropriate within morally charged hypothetical situations. Known as moral dilemmas, people's judgments are taken to be informative about their underlying ethical beliefs. While this common practice has produced powerful insights about our moral sensibilities, some have questioned whether the collected data is predictive of our real-world behaviors in morally significant situations. Thus, researchers have called for investigations that measure moral behavior, or at least associate behaviors, with moral judgments. While some investigations have used creative methods to target behavior, many have taken to studying people's behavior in games, associating it with their moral judgments or beliefs. In this quickly growing area that studies moral behavior in games, researchers have yet to investigate how people's moral judgments predict behavior in games that simulate the tragedy of the commons. The particular

importance of commons dilemmas comes to light when considering not only some influential arguments of past thinkers about the superiority of rational moral theories but also, the growing threat commons dilemmas pose to an ever-more globalized world, where viral pandemics and climate change have exerted increased pressure on human society. In recognition of these facts, a new package of research was initiated to investigate whether certain types of moral responders (and patterns of moral judgment) predict behavior in game situations that simulate commons dilemmas. Within this package, two pilot studies and three experiments were performed to test two key premises. First, do moral judgments predict behavior in games that simulate the tragedy of the commons? Second, do rational moral judgments predict behaviors that are less susceptible to bad outcomes, that is, a tragedy for the commons? The answer to the first question is unequivocal in the affirmative, as moral judgments predicted behavior in commons games. Answers to the second question are more difficult to come by. At the very least, rational moral judgments predict behavior that is better suited toward avoiding tragedies of the commons at least some of the time (in specific instances). Other types of moral judgments (which are often argued to be less rational) also predicted an increase in behaviors likely to avoid tragedies of the commons. The conclusions that can be drawn and the future directions that can be made will be discussed.

This document is dedicated to the many people that took it upon themselves to act for my benefit. To my parents and family, who brought me up. To my instructors, who taught me how to do things. To my co-workers, who handled my work responsibilities at the grocery store so I could study in the backroom for hours on end during my shifts. To my friends, who showed me to live a balanced life. And finally, to my wife, who has sacrificed much and provided me with the utmost purpose and happiness throughout our journeys together.

To you all, I can only say thank you in my most sincere voice. It cannot buy what you have given but know that your contributions will never be lost on me.

Acknowledgements

I would like to thank my advisor, Dr. John Jasper. Thank you for your insights, advice, and guidance throughout my graduate career at the University of Toledo.

Also, I would like to thank Dr. Raymond Voss for helping me to adjust during those first few tumultuous months in Toledo.

Finally, I would also like to thank my friend, Alyssa, who was across the hall on the fifth floor of University Hall during most of my time here. She helped me learn how to do the things that need doing in this line of work.

Table of Contents

List of Tables

List of Figures

List of Symbols

*..........Significance, P < .05
**........Significance, P < .01

Preface

Questions of morality speak deeply to what it means to be human. We are beings that are *thrown* into the world (as Heidegger would say) without answers and yet we are imbued with faculties that have the character of ascribing a type of certainty and correctness to the hard questions that we encounter in the world. What is morally right is one such question. In one sense, its answer seems to be helplessly subjective and inevitably beyond the realm of empirical inquiry. This may be. And yet, our sense of morality makes its voice known deeply from within, as a type of screaming conviction that calls us to action.

Philosophers and scientists have dedicated lifetimes of work that call our moral intuitions and beliefs into focus. Even the briefest analysis can offer us a clear answer as to why. That is, morality at some basic level addresses the how in which people should treat one another (and perhaps even themselves) in a world constrained by time and defined by change. We live and then we die. In the interim, we take up projects that require cooperation. And thus, morality is an all-important feature of our worlds. It is akin to atmosphere. It is difficult (if not impossible) to see directly, but it is everywhere. It is breathed in (when we encounter virtuous others) and exhaled (when we perform moral acts). We recognize its importance to our well-being but have trouble grasping it. Nevertheless, what precisely is moral has been subjected to a long history of folk and philosophical theorizing. Still vexed by the questions outlined above, like the subjectivity of morality, questions about what is moral continue to lack a decisive answer. Of course, these are big questions that do not lend themselves easily to solutions. Thus, psychologists have preferred a different tact – that ignores metaphysical pondering about

what morality is and where it comes from, in favor of empirical data about how morality is experienced and acted upon in specific contexts. That is, psychologists do not attempt to answer whether certain actions or beliefs are moral, but rather, focus on what our moral judgments are in situations that have the quality of seeming morally important.

In their pursuits, psychologists have collected data about peoples' moral intuitions and judgments across a wide variety of situations. Despite this, these situations often require participants to judge whether it is ethically appropriate to sacrifice a person to bring about a greater good. Thus, the situation of life and death is taken to have the quality of seeming morally significant and the participant's judgment is thought to provide data about their underlying moral sensibilities. While this technique has yielded many interesting findings, several key weaknesses in its approach have been pointed out. While these weaknesses of the common practice will be given considerable focus in this dissertation, they can be effectively outlined here as (1) failing to study moral behavior in an ecologically valid way, (2) failing to fully appreciate the nuances of certain key moral theories, and (3) failing to consider certain irreducible asymmetries between judgment and belief. In recognition of these weaknesses, there has been a push to extend the empirical analysis of moral judgment beyond sacrificial dilemmas and into other situations (such as games) that have behavior (instead of only judgment) as a measurable outcome variable in situations with greater ecological validity (like social traps and economic games in the context of collective action problems).

For human societies to function, individual rights and interests must be weighed against the rights and interests of the collective (for an in-depth analysis see Schneewind, 1983). This tension can be seen in many different areas of modern life, where societies

have considered policies that restrict individual freedoms to promote the collective good (such as emission policies to combat climate change). In 1968, Hardin powerfully exemplified such dilemmas in a hypothetical scenario called the tragedy of the commons. In commons dilemmas, individuals need to cooperate with the collective to prevent disaster, which requires a balancing of their own interests, with the community good. In the lab, behavior in these dilemmas has been investigated by observing the actions of participants in games that simulate commons dilemmas.

While morality has been theorized to play a pivotal role in determining the appropriateness of this balance between personal and community goals (Baatz et al., 2019; Darr, 2019; Greene, 2013), research has seldom investigated how particular types of moral agents behave in these situations. Thus, while several philosophers have argued that certain codes of ethics are better suited to solve commons dilemmas than others (see Greene, 2013), it is difficult to answer whether individuals, that adhere to different theories of ethics, actually behave differently in situations like the tragedy of the commons. That is, do a person's moral judgments (and the categorization of those judgments) predict their behavior in seemingly morally significant situations (especially when real consequences are tied to that behavior)? If so, several related questions should be asked, like whether certain types of moral responders (or even certain types of moral judgments) are predictive of behaviors better suited to avoid tragedies of the commons? These are key questions, as collective action problems that mirror the tragedy of the commons are represented in several significant social problems that we suffer from today (e.g., viral pandemics) and are expected to suffer from in the future (e.g., climate change, artificial intelligence).

Paper Overview

In recognition of the shortcomings of the conventional empirical approach and the importance of these questions, in this paper, I will review the research that has directly investigated behaviors of different moral agents in cooperative game situations before proposing a new line of research designed to extend knowledge in this area. This will demand context, and thus, a brief description of game theory and the major findings of moral judgment studies will be presented in the first and second chapters. Then, in the third chapter, a detailed review of studies that test moral behavior in games will be provided. In the fourth chapter, I will outline my program of research and offer my predictions that were established beforehand, after which, in the fifth chapter, I will discuss the results and analyses related to that research. Finally, in the sixth chapter, I will offer concluding remarks and speak to the potential future directions of this research.

Chapter One

Socrates, the Sicilian, and Social Dilemmas

If not unique to humans in kind, theory of mind is unique by degree, compared to other animal species. This seemingly innate ability allows humans to strategize in competitive situations against others in sophisticated ways. For instance, in many competitive situations or games, a player's best action (for them) depends on what they expect the other player(s) will do. Factoring in the motives and expected decisions of other competitors can make seemingly simple games (such as tic-tac-toe) exceedingly complex.

This complexity is humorously illustrated in the 1987 film, *The Princess Bride,* when the Sicilian and the Dread Pirate Roberts face off in a battle of minds. In this scene, these two characters play a game of life and death. The game begins when the Pirate pours deadly poison into one of two cups. While the Sicilian knows that the poison has been poured, he is not aware of which cup it has been poured into. After pouring the poison, the Pirate gives the Sicilian, and himself, each a cup. Before the game ends, both must drink, but the Sicilian has the final decision on who drinks from which cup (i.e., the Sicilian can ask to switch cups). Thus, the Sicilian must decide if the Pirate has put the poison into his own cup or the Sicilians' cup. As the Sicilian reasons out loud for the Pirate (and the audience) to hear, he attempts to predict which cup has been poisoned by taking into account the Pirate's own mental states, "I know that you know that I know" and so on.

While our ability to appreciate these complexities usually offers us a decisive advantage, it can sometimes lead to worse outcomes (Ross, 2019). For example, the

1

ancient philosopher Socrates reasoned that a soldier might be motivated to flee a battle if his side has an overwhelming force and is sure to win regardless of his actions. After all, why should the soldier continue to risk his life to achieve an already decided outcome? The key problem is that if this thought also occurs to the other soldiers in the ranks, then these soldiers will run also, and the battle will quickly turn into a deadly rout. Thus, the seemingly rational actions of individual decision-makers (i.e., fleeing the battle) can lead to a worse outcome intended by no one (i.e., a deadly rout).

Such scenarios are often referred to as social dilemmas in psychology, and while they have long been recognized (see Randall, 1964), their understanding was revolutionized in the 20th century by Jon Von Neumann, when he developed Game Theory. In short, game theory allows for an analysis of the expected success of strategies in game situations against other competitors (Camerer, 2003). It describes strategies that rational agents should use to secure rational ends (like maximizing utility). While some disagree about how to define and measure rationality, it is often thought to be connected to logic and result-based analysis (Julmi, 2019). Thus, in this paper, rationality will be used to refer to actions or rules that produce, or are expected to produce, the best result.

Adopting the definition just outlined, psychological research has found that humans routinely fail to behave rationally in many situations (for an in-depth analysis, see Ariely, 2008). Reasons for acting irrationally run the gamut; however, evidence indicates that morality plays a part. Indeed, in moral dilemmas, rational actions are often perceived negatively, and rational decision-makers are sometimes viewed as less warm and less preferable, compared to those who made principled or emotionally congruent

decisions (Rom et al., 2017). This negative perception of rational behavior also extends to game situations.

In a study by Grossman et al., (2020), people were asked to predict the behaviors of players who were described as "rational" versus "reasonable" in commons dilemmas and other economic games. Participants predicted that rational players would behave more selfishly. In addition, participants were more likely to describe rational players as emotionally suppressed, while reasonable players were more often described as honest, fair, and moral. Studies have even shown that in cooperation games, people who take longer to decide whether to cooperate are viewed as less trustworthy than individuals who cooperate quickly (Jordan et al., 2016). The interpretation is that participants viewed people who agreed to cooperate more quickly as adhering to something like a principle of trustworthiness, rather than a commitment to pragmatism. Thus, while it seems beneficial to calculate the potential costs and benefits of cooperating, this behavior may come with certain social costs that stem from the moral sensibilities of others. This is an important point, as certain moral inclinations require the explicit calculation of the consequences of one's actions, while others support acting in accordance with general principles (like fairness or reciprocity).

Playing Fair.

Humans have strong intuitions about what is fair (Hallsson et al., 2018) and often react negatively (or perhaps irrationally) when our expectations of fairness are violated. Psychologists have measured these intuitions, and the behaviors they instantiate, by having people play economic games (like the ultimatum game). In the ultimatum game, one player is given x number of resources that he must split between himself and the

other players. If the other players reject his proposed distribution, they all walk away with nothing. While a rational agent would accept any offer greater than zero (it is better to have some portion of x than none), players nonetheless routinely reject highly unfair offers (Camerer & Thaler, 1995). Of relevance to the current review, previous works have found that these rejections are motivated by emotional (e.g., Zheng et al., 2017) and moral (e.g., Hallsson et al., 2018) processes.

In the case of emotion, rejections of unfair offers are predicted by large emotional reactions (Reuben & Van Winden, 2005), high reactions of disgust (Chapman et al., 2009), and emotionally laden desires to punish negative behavior (Hallsson et al., 2018). While some researchers have argued that these rejections are defensible (e.g., Nowak et al., 2000), in that, decision-makers must balance their desire to profit with their social reputation (i.e., one does not want to have the reputation that he can be easily taken advantage of), studies have found that people reject unfair offers even in situations of anonymity (e.g., Ning Ma et al., 2012), when social concerns cannot be taking an active role in the rejection. Perhaps most alarmingly, people often reject unfair offers in conditions where the rejection does not penalize the offeror (Yamagishi & Horita, 2009). These findings have led some to argue that rejections of unfair offers are driven partially by moral concerns, as people often react to offers that violate one's sense of intuitive fairness with feelings of moral disgust (for a review, see Hallsson et al., 2018).

Social Dilemmas.

Beyond using ultimatum games, studies have also had participants play games that simulate social dilemmas, where the interests of the individual players are put into conflict with the interests of all players (for a review, see Van Lange et al., 2013). Many

scenarios have been created to research how people respond in such situations (e.g., stag hunt game, public goods game, the commons dilemma). Common's dilemmas, which are a type of social dilemma that leverages a player's short-term individual interests against long-term group interests, are particularly interesting due to their close approximation of collective action problems that face society today (e.g., climate change).

For a simple example of a commons dilemma, consider four villagers sharing fish as a resource in a pond that has a capacity of twelve fish and that every night, each pair of fish conceive a second fish. Given these basic requirements, the best-case scenario (in respect to expected utility) is that each villager only fishes one fish per day. This way, every day, the number of fish in the pond is reduced to eight and then increases back up to twelve the next day. In the long run, this will lead to the best result for the farmers collectively. However, if one of the villagers takes more than their fair share of fish, while this selfish action allows an immediate short-term benefit for him, the cost is spread out amongst all the villagers in the community. Much like in Socrates' battle scenario, where one soldier fleeing is expected to cause a chain reaction of the army fleeing in mass, the action of this one villager would likely cause a chain reaction, where the other villagers would also overfish the pond, leading to a long-term tragedy for the commons. Thus, not only is it irrational to overfish the pond, but it would also be morally significant, as such actions, expected to lead to collective tragedy, seem prima facie immoral. As a quick caveat, whether this is actually immoral is something that philosophers may contest. While a deeper consideration of morality will be provided in a later chapter, the seeming immorality of such actions is assumed for now.

Acknowledging the connections between rationality and morals in commons dilemmas, some have argued that following a characteristically rational moral theory (i.e., utilitarianism) should better equip individuals to prevent tragedies of the commons than does our common-sense moral intuitions (Greene, 2013). However, to date, few studies have investigated how different types of moral agents behave in tragedy of the common's situations. Instead, most of the research has investigated human morality in hypothetical thought experiments like the trolley problem (e.g., Greene et al., 2001) or with questionnaires assessing moral beliefs (e.g., Dickinson et al., 2016), instead of using more real-life scenarios (Ellemers et al., 2019). Thus, it is difficult to answer whether people that show more characteristically rational moral judgment will behave differently in commons dilemmas (see Clarkson, 2022 for a review). The limited research that has attempted to answer this question will be reviewed in detail in chapter three. However, chapter two will first provide a brief overview of the work in moral philosophy and moral psychology to lay context.

Chapter Two

Psychological Theories and Moral Judgment

In the philosophy of normative ethics there exist many different theories about what constitutes moral rightness and wrongness. A lengthy discussion of these many theories is beyond the scope of this project. It is, however, important to outline two conflicting ethical theories (i.e., deontology and utilitarianism) since they have exerted a large impact on psychological investigations of morality. In Kantian ethics, often referred to as deontology, morality has a theorized basis in the laws of reason (Korsgaard, 1996). Therefore, to Kant, obligatory moral principles are discernable by reason alone. For certain reasons that I will not fully explore here, deontology is typically associated with commonsense morality. That is, deontology has been argued to be based on common-sense morality in a post hoc fashion because it instantiates the emotionally based gut-level, and thus, common-sense moral intuitions of the common man (for an in-depth treatment of this argument see Greene, 2008). Still, this is not to say that deontological precepts cannot be rational; rather, it is argued that the adherence to deontological moral rules sometimes leads to irrational decisions (Sunstein, 2005).

By contrast, for utilitarian ethics (which is a category of consequentialism), morality is assessed by the consequences of one's actions (Troyer, 2003), and it often diverges from common sense moral intuition. In the words of famous utilitarian philosopher John Stuart Mill, "the creed which accepts as the foundations of morals "utility" or the "greatest happiness principle" holds that actions are right in proportion as they tend to promote happiness; wrong as they tend to produce the reverse of happiness"

(Troyer, 2003, p. 99). Simply put, for utilitarianism, actions that promote the greatest good for the greatest number are morally appropriate.

Building upon these philosophical theories and other interdisciplinary work, psychologists have extensively investigated morality across thousands of research studies (Ellemers et al., 2019). Based on the empirical findings of this research, several popular theories have been proposed (e.g., morality as cooperation, moral foundations theory, dual-process theory). In contrast to those in philosophy, theories in moral psychology do not attempt to identify what is moral or to make prescriptions about what humans ought to do. Instead, they are concerned with describing the many psychological processes that instantiate morality and are important insofar as they can explain our moral judgments and ultimately, predict our behaviors in situations where morality matters. While these theoretical approaches are different in kind, interdisciplinary investigations have used the degree to which people follow the prescriptions of utilitarianism and deontology as evidence of their underlying moral motivations (Conway & Gawronski, 2013). For instance, utilitarian decisions are often argued to be the product of utilitarian motivations (e.g., Conway et al., 2018). While this could be so, some moral theories in psychology offer a set of additional or alternative explanations for our moral judgments. Thus, it is important to review theories in philosophy and moral psychology. For the sake of clarity, I will begin the next section by defining some relevant terminology before discussing specific theories in moral psychology.

Moral Theories in Psychology.

Psychologists have used many terms to describe different and overlapping psychological aspects of morality. Indeed, these terms are not always clearly defined in

the literature and to a varying degree, are used imprecisely. Intuitions and beliefs represent a good example of this, as they have often been used interchangeably (e.g., Sinnott-Armstrong et al., 2010). Here, intuitions and beliefs generally refer to attitudes, feelings, or dispositions about what is morally appropriate. Moral inclinations are similar, in that, they also assume a disposition about what is appropriate. Indeed, it is up for debate how much daylight exists between these terms (e.g., Ramsey, 2019). It is clear, however, that intuitions, beliefs, and inclinations are not the same as judgments or behaviors.

A moral judgment is the assessment of the morality of a specific action (Saltzstein, 1994). In this way, moral judgments are less abstract than intuitions and beliefs. For instance, a person may hold the belief that it is morally wrong to commit murder, but this is not in itself a moral judgment. Instead, a moral judgment would require that a person has evaluated the moral appropriateness of committing murder in a situation (real or hypothetical). Moral behaviors are elevated beyond moral judgments in that they typically require a person to commit an action that has consequences for themselves and others (Saltzstein, 1994). That is, for a person to engage in a moral behavior, they must make a choice or perform an action that carries with it some level of moral responsibility. While psychologists have used several interesting methods to study moral behavior, these efforts often involve having participants play games that have morally relevant consequences. For instance, if a person chooses to act unfairly (or cheat) in a game, this action has morally relevant consequences for the cheater and the other players. These consequences can be restricted to the game or be affected upon the real world, typically with monetary incentives.

Having defined the necessary terms, space will be given below to outline the dominant theory in moral psychology, that is, Dual-Process Theory. However, it should be noted here, that there are other theories within moral psychology, which were given passing reference above (i.e., moral foundations theory and morality as cooperation theory), that pay less attention to the moral judgments people make and instead focus primarily on diagraming the moral beliefs that people hold. These theories will be relevant when considering potential future directions resulting from this project and will thus be taken up in some detail in chapter six.

While there are many models and accounts of morality that fall within the purview of dual-process theory, it commonly holds that our moral intuitions are the product of often conflicting emotional and rational processes (Greene et al., 2001) that roughly instantiate deontological and utilitarian ethics respectively (Greene, 2013). Indeed, while the theory is not universally accepted, it is supported by many empirical studies that have investigated moral judgment. This research has often investigated how individuals make decisions in morally charged hypothetical dilemmas modeled after philosophical thought experiments (for a review, see Ellemers et al., 2019). These thought experiments, sometimes referred to as "trolley problems", typically require that decision-makers select an option they consider to be ethically appropriate from a larger set. In their common form, they present decision-makers with two solutions to a dilemma. These solutions differ, in that, one often adheres to deontological ethics (i.e., a moral principle is not violated), while the other adheres to utilitarian ethics (i.e., the best consequences are produced). In this way, researchers can record participants' responses

to determine whether they are making deontological judgments (consistent with the emotional process) or utilitarian judgments (consistent with the rational process).

For instance, in a famous example from Greene et al., (2001), individuals are told to consider a scenario where a runaway trolley is about to run over and kill five people further down a track. The only way to avoid this outcome is to pull a switch that will divert the trolley onto a separate track. However, on this alternative track, there is another person who will be run over and killed if the switch is pulled. Participants are then asked to decide what they ought to do. Should they pull the switch (i.e., the option advised by utilitarian ethics as it leads to the best consequences)? Or should they refuse (i.e., the option advised by deontological ethics as it is considered an objective moral principle not to kill)? While many modifications have been made to moral dilemmas like the switch case, they usually hinge on forcing people to endorse either a utilitarian or deontological solution. For instance, in another similar dilemma (known as the footbridge dilemma) from Greene et al., (2001), individuals are again told that a runaway trolley is about to run over and kill five people further down a track. The only way to avoid this outcome is to fatally push a large man off a bridge that runs over the track, so that he falls in front of the trolley, stopping it before it runs over and kills the five people further down the track.

Notice, that in each of these two scenarios, the consequences are kept logically consistent. That is, in each case, decision-makers must judge whether it is appropriate to kill one person to save five. Interestingly, among the most common findings in moral psychology is that these two scenarios are typically responded to in different ways. That is, people usually judge it appropriate to pull the switch but not to push the man off the bridge (e.g., Bloom, 2011; Gleichgerrcht & Young, 2013). This lack of consistent

judgment between logically equivalent scenarios has often been explained, within the context of dual-process models, by appealing to emotion (e.g., Greene, 2013; Haidt, 2001). That is, viewing these dilemmas appears to differentially activate brain areas associated with emotion (Greene et al., 2001) and scenarios like the bridge case, where the deontological response is more common, illicit increased levels of emotion (e.g., Horne & Powell, 2016). Conversely, when participants endorse the utilitarian solution to a dilemma, this is assumed to indicate that the participant is motivated by a cognitive concern to bring about a greater good (e.g., Greene et al., 2001). Taken together, these and other findings have been interpreted as evidence in favor of a dual-process model of moral judgment (e.g., Bretz & Sun, 2018; Greene, 2008; Haidt, 2001).

In accordance with the model, many studies analyze participants' utilitarian vs deontological decisions to make inferences about the nature and strength of each process (e.g., Conway & Gawronski, 2013; Friesdorf et al., 2015). However, this practice has been challenged by research showing that people who make utilitarian decisions in sacrificial dilemmas are more likely to display negative personality traits, like psychopathy (Bartels & Pizarro, 2011), are less moral in a variety of other situations (Kahane et al., 2015), and are more likely to view cheating as appropriate (Crittenden et al., 2009). The upshot of these findings is that these sacrificial judgments may not reflect a stronger rational process, but rather, a lack of morality. This has led to an unresolved debate in the area. Do utilitarian judgments in sacrificial dilemmas reflect a true concern for the greater good or a lack of moral principle and an antisocial disposition?

In a direct response to Kahane et al., (2015), Conway et al., (2018) argue that sacrificial utilitarian judgments do represent a concern for the greater good and that, their

association with negative personality traits is an artifact of a second variable (i.e., low deontological inclinations). Thus, they argue that individuals can be sensitive to both utilitarian and deontological motivations simultaneously. Using an approach known as the process dissociation technique, which allows for an independent analysis of participants' utilitarian and deontological inclinations, their results show that low deontological inclinations (not high utilitarian inclinations) are associated with the negative personality traits and immoral behaviors identified by Kahane et al., (2015). This debate remains unresolved and will be discussed further in the following section; however, it is important to recognize here that using methods that only record participants' moral judgments, rather than their behavior in morally relevant situations, may not provide sufficient data to settle it.

Weaknesses of only Measuring Moral Judgment.

A recognition of the inherent limitations of studying morality by collecting data on participants' moral judgments is not new. Indeed, it has been acknowledged for some time that a fundamental shortcoming of this approach is that it asks for, and collects data on, participants' moral judgment, rather than their actual behavior (Bloom, 2011). To be clear, while collecting data on participants' moral judgments is valuable – it is interesting to detail what people think one ought to do in a specific moral situation; it may not predict behavior in real-life situations. After all, moral judgment data has typically been gathered from extreme moral dilemmas (like the trolley problem), which rarely happen in day-to-day life. While an extensive body of research has yet to be gathered on the differences between moral judgment and behavior (Patil & Silani, 2014), several studies have found key discrepancies between the two (Bostyn et al., 2018; Gold et al., 2014;

Gold et al., 2015; Francis et al., 2016; Tassy et al., 2013). For instance, research by Gold et al., (2015) found that when people are asked to perform a moral action with real-world consequences, the previously noted differences in judgment between dilemmas like the switch and footbridge case fall away.

Of the studies that have investigated potential differences between moral judgment and behavior, several did not actually measure behavior but instead measured participants' predictions of how they would act if the dilemma was real (Tassy et al., 2013). Interestingly, results show that peoples' predictions of how they would act are more utilitarian than their judgments about what is morally right (e.g., Tassy et al., 2013). This is consistent with other research observing increased utilitarian judgment when dilemmas are presented in virtual reality rather than conventional text formats (Francis et al., 2016). While this research has interesting implications (i.e., perhaps people would show more utilitarian behavior than their judgments in conventional sacrificial dilemmas indicate), it nonetheless falls short of measuring actual moral behavior.

The small body of research that has risen to this standard, by associating participants' moral judgments with morally relevant behaviors, has yielded mixed results. In a well-known study conducted by Bostyn et al., (2018), participants' judgments in hypothetical trolley problems were compared to a similar real-life decision. That is, participants were told that five mice were about to receive an electric shock unless they decided to push a button, which would redirect the electrical current to a different mouse instead. When compared to their judgments in hypothetical versions of the trolley problem, participants were twice as likely to make the utilitarian decision in the real-life mouse dilemma. In addition, the authors did not find evidence that participants'

hypothetical judgments (to trolley problems) predicted their actual behavior in the mouse

dilemma. However, some have argued that this conclusion goes too far, citing that

alternative analysis on the same data reveals that participants' hypothetical moral

judgments marginally predicted their behavior (Plunkett & Greene, 2019). Indeed, other

studies have found that studying peoples' responses to hypothetical moral dilemmas can

be useful for predicting certain patterns of moral behavior in games, like antisocial

money burning (Dickinson & Masclet, 2019). Thus, while the precise relationship

between moral judgment and behavior continues to be debated, there does seem to be a

growing body of evidence indicating that people may act in ways that are more utilitarian

than their judgments in sacrificial dilemmas are letting on.

      Another weakness comes from outside of psychology. In philosophy, some have

questioned the legitimacy of using moral judgment data from thought experiments like

the trolley problem (for an in-depth analysis, see Ramsey, 2019). In this vein, it is argued

that peoples' intuitions about how they would act in moral dilemmas are not equivalent to

judgments or beliefs about how they would, or even should, act. Instead, as argued by

Ramsey (2019), participants' intuitions about the morally appropriate action are less of a

judgment and more of an intellectual seeming. To Ramsey, these intellectual seemings

are more like "defeasible pro-attitude" mental states toward a position that does not

necessitate a belief (Ramsey, 2019, p.80). Specifically applied to the trolley problem, a

person who intuits that they would pull the switch (i.e., saving five at the cost of killing

one) believes that pulling the switch will save lives (as this is a feature of the scenario

itself) but is merely reporting a defeasible attitude in favor of pulling the switch (2019).

Ramsey argues that these intuitions may be akin to the intuition that two equally long

lines are different lengths (as produced by the Muller-Lyre illusion) while still holding

the belief that they are not.

Chapter Three

Moral Psychology and Games

While behavior in economic games, and to a larger extent, social dilemmas, have been commonly studied in psychology, associations between behavior in these situations and morality have only recently been studied (Alfano et al., 2018) and remain underexamined (Brodbeck et al., 2013; Ellemers et al., 2019). To this point, only a few recent studies have investigated the relationship between participants' moral beliefs, attitudes, and judgments, with their behavior in game situations. Of these, they often differ across several factors, including how they measure morality.

While some studies have used questionnaires that assess support for specific moral precepts, like those outlined by moral foundations theory (Clark et al., 2017; Dickinson et al., 2016; Kim et al., 2013; Weaver & Lewis, 2012), the more common approach is to use participants' utilitarian versus deontological judgments in sacrificial dilemmas as an individual difference factor (Bostyn & Roets, 2017b; Capraro et al., 2018; Chen, 2011; Everett et al., 2016; Everett et al., 2018; Noclain, 2018; Sacco et al., 2017). Within this set, some studies have used additional or alternative methods, such as white lie dilemmas (Noclain, 2018) or beneficence dilemmas (Everett et al., 2018), or have asked participants to evaluate justifications for utilitarian versus deontological actions (Kreps & Monin, 2014).

In addition to the noted differences in measuring morality, the games themselves, which measure economic and moral behavior, often differ from study to study (e.g., the trust game, dictator game, etc). Even particular games, like the trust game, have been manipulated to explore different interacting factors (such as the number of players in the

17

game). Another major differentiating factor is whether studies investigate how a particular moral agent is perceived by others (e.g., if participants trust a utilitarian player) versus how different moral agents behave in game situations (e.g., how a utilitarian player actually behaves in the game). Taken together, these differences have allowed for a small number of often exploratory studies to make tentative conclusions about how moral judgment predicts behavior in some game situations. However, this research remains largely incomplete and as such, it cannot settle certain debates about conflicting results and leaves other important questions entirely uninvestigated.

Findings Across Different Games and Dilemmas.

Studies that have sought to associate moral judgments with behavior in economic games have mostly tested participants in cooperation games (e.g., trust games, public goods games). In these games, players need to cooperate to achieve the optimal utilitarian result (i.e., the greatest good for the greatest number). Among the different types of cooperation games, the trust game is most popularly used (Bostyn & Roets, 2017; Capraro et al., 2018; Everett et al., 2016; Everett et al., 2018; Noclain, 2018). In their common form, trust games involve two players (Player A and Player B).

Player A starts the game with some amount of money and can transfer any amount of that money to Player B. Then, the entire amount that was transferred is increased by some factor, and Player B gets to decide how much (if any) of the money to transfer back to Player A. In this game, the best-case scenario (in terms of utility) is that Player A transfers the entirety of their starting funds to Player B, as this allows for the greatest monetary sum to be awarded in the game. However, to do this, Player A needs to trust that Player B will return a portion of the money that is at least equal to the amount

that they originally transferred to Player B. Thus, in trust games, Player A's actions are treated as an individual measure of trust, and Player B's actions are a measure of trustworthiness (i.e., how much does Player B transfer back to Player A).

In these studies, participants often make moral judgments in hypothetical sacrificial dilemmas before playing trust games with other players. Participants' responses can be used to categorize them as utilitarian responders or deontological responders. Once they start the game, participants are told the moral judgments of the other players in the game. Based on this information, participants often rate co-players on several metrics, such as likeability (Sacco et al., 2017), or perceived trustworthiness (Bostyn & Roets, 2017b; Capraro et al., 2018; Everett et al., 2016; Everett et al., 2018; Noclain, 2018). Finally, behavioral indices of trusting and trustworthiness are calculated by examining the cooperative behavior of players in the game.

Trust and Cooperation. The studies that have directly associated moral judgment with behavior in trust games have yielded mixed results, indicating that a perception gap exists between perceived trustworthiness and actual trustworthiness. That is, while it has regularly been found that participants perceive (and act as though) players who made deontological judgments in sacrificial dilemmas are more trustworthy, the behavioral data indicates that these players are not any more trustworthy than those who made utilitarian judgments. In addition, some interesting but limited results have been found on the group level, when looking at a different cooperation game, i.e., the public goods game.

Across five studies, Everett et al., (2016) found that participants viewed people who made the deontological choice in the footbridge dilemma as more trustworthy,

transferred more money to them in the trust game, expected that they would transfer more money back to them in later rounds, and preferred them over utilitarian partners as much as 80% of the time. Interestingly, this preference was maintained even when controlling for the participant's own moral judgments. That is, utilitarian and deontological participants showed the same preference for deontological partners. These findings were replicated (study one) and extended (study two) in Bostyn and Roets (2017b).

In study one, participants rated deontological partners as more trustworthy than utilitarian partners and transferred more credits to deontological partners. However, despite these preferences, deontological responders did not transfer more money than their utilitarian counterparts, indicating the existence of the aforementioned perception gap between perceived and actual trustworthiness. In study two, participants played a hypothetical public goods game with three other sham participants, in which, they began with 100 credits. Participants could donate any amount of their credits to the public good, where this total (from all players) would be doubled and redistributed evenly, regardless of the player's original donation. Finally, participants were told how the other players had responded to sacrificial dilemmas. Like in the trust game, participants trusted deontologist responders more; however, in the public goods game, deontological responders were found to be more cooperative than utilitarian responders.

Conversely, recent investigations by Capraro et al., (2018) found that deontological responders were not more trustworthy in trust games than utilitarian responders, after analyzing the amount of money returned in later rounds of the game. There are several potential explanations for this conflicting result. First, Capraro et al., (2018) classified participants as deontological or utilitarian by assessing their responses

to a modified version of the footbridge dilemma, known as the trapdoor dilemma. While the footbridge and trapdoor dilemmas are logically consistent, the trapdoor dilemma teases apart two arguably morally significant variables, i.e., personal force and the doctrine of double effect, which have been shown to heavily influence moral judgments (e.g., Greene et al., 2009). In addition, while both studies measured the behavior of different types of moral responders in cooperation games (i.e., the public goods game and the trust game), research has shown that response patterns between these games sometimes only display small to medium correlations (Haesevoets et al., 2015). Still, the overall finding that deontological judgment increases one's perceived trustworthiness has strong support in the literature.

      Altruism and Competition. A few studies have associated moral judgment with behavior in non-cooperation games, like the dictator game. In dictator games, Player A starts with a sum of money, which they can share with Player B or not. In these games, Player B has no active role in the game. Since Player B cannot reciprocate by returning some of this money to Player A (as in the trust game), Player A's giving is taken as a measure of altruism. Research has previously shown that cooperation is common in social dilemmas, like the dictator game, even when cooperation is not expected to generate better outcomes (Capraro et al., 2014). This research adds to other accounts (e.g., evolutionary psychology) that argue that cooperation is a type of heuristic that is generally adaptive. However, this study did not investigate potential associations between morality and cooperation. A later study by Capraro et al., (2018), where participants played dictator games in pairs, after first being given information about their partner's moral judgment in the trapdoor dilemma, came closer to the mark. While participants

expected deontological partners to be more altruistic, deontological and utilitarian responders did not differ in their actual donations.

      Moral Conformity. The research outlined above clearly demonstrates that a reputational advantage is to be had if one is observed making deontological judgments in sacrificial moral dilemmas. Other research indicates that individuals may be sensitive to this reputational advantage, as people have been shown to offer different moral judgments when responding in groups (Bostyn & Roets, 2017a; Kundu & Cummins, 2013). In Bostyn and Roets, (2017a), participants were asked to make moral judgments after first being presented with data about how other participants responded to the same dilemmas. Between conditions, participants were informed that their peer participants had made a majority of utilitarian (65-85%) or a majority of deontological (65-85%) decisions. Results showed an interesting asymmetric conformity between conditions. While participants heavily conformed to the deontological majority (i.e., they were more likely to make the deontological judgment), participants did not show increased utilitarian judgments in the utilitarian majority condition.

      This asymmetric conformity has been argued to stem from peoples' desires to represent themselves as attractive and cooperative social partners (Noclain, 2018). One piece of evidence consistent with this view comes from research by Gold et al., (2015), who compared rates of utilitarian judgment between observers and deciders in typical sacrificial dilemmas. In this study, participant-observers were more likely to endorse the utilitarian option as ethically correct than those making the moral decision. This result potentially indicates that decision-makers are sensitive to potential social costs (like a

loss of credibility) that may be suffered from "getting ones' hands dirty" by deciding to sacrifice someone for a greater good.

While extensive research has not yet tested the validity of such interpretations, recent (e.g., Baumard et al., 2013; Curry et al., 2019; Haidt, 2012) and distant (e.g., Schneewind, 1983) works have argued that our moral intuitions are finely tuned to facilitate social functioning. Given this, it makes sense that people motivated by reputational concerns would conform to the deontological majority (as deontological judgments are associated with trustworthiness) but not the utilitarian majority (which should theoretically be motivated by concerns to maximize outcomes). However, to date, investigations of group-level morality have been rare, even though researchers have called for additional studies investigating moral behavior in groups (Ellemers et al., 2013). Indeed, the lack of research that has focused on addressing this question is representative of the larger state of the literature that has investigated moral behavior. While many other variables (such as gender) may influence the relationship between moral judgment and behavior in game situations, a better understanding of this basic relationship is first needed.

Chapter Four

Moral Judgment and the Tragedy of the Commons

Thus far, I have previewed arguments suggesting that moral agents (who adhere to utilitarianism over deontology) may be more likely to avoid tragedies of the commons. After outlining the basic features of commons dilemmas, I broadly covered game theory and cooperation, as concepts that are related to how science understands human behavior in game situations. Then, I sketched several relevant moral theories before delving into a deeper analysis of the dominant findings of past psychological investigations on moral judgment. In the third chapter, I reviewed the current state of literature that has directly associated moral judgments with behavior in mostly cooperative game situations. In the current chapter, I return to the topic previewed at the outset. That is, I will address the works of Greene and others who have argued that utilitarianism is better suited to solve the tragedy of the commons than deontology, before cataloging a new line of research that aims to test the validity of these arguments in chapter five. This research and its predictions were informed by several pilot studies, which will also be covered in the current chapter.

Moral Judgment and Hardin's Tragedy of the Commons.

In an essay titled, *The Secret Joke of Kant's Soul*, Greene argues that deontological ethics is largely based on gut-level emotional reactions, and not, rationally derived moral principles (Greene, 2008). Indeed, as has already been mentioned, participants' emotional reactions have been predictive of deontological judgment in many hypothetical moral dilemmas (e.g., Greene et al., 2001). While Greene concedes that many of these emotional reactions stem from the evolutionary benefits they allow (e.g.,

people commonly react to incestual acts with disgust), he argues that a reliance on them for one's morality has a major flaw. That is, these emotional reactions, which some have argued to be akin to moral heuristics (Sinnott-Armstrong et al., 2010; Sunstein, 2005), are likely to lead us to moral error, much like heuristics in other decision-making domains that lead to cognitive errors. In his 2013 book, *Moral Tribes: Emotion, Reason and the Gap Between Us and Them,* Greene argues that commons dilemmas are one such area, where a morality based on emotion, will likely lead individuals to make poor decisions.

First, let me say that this project is not attempting to litigate basic moral truths. Insofar as this can be done at all, this is the job of scholars in ethical philosophy. While game theory can identify rational behaviors based on economic payoffs, no universally accepted theory that outlines rational moral ends exists (Gold & Colman, 2020). While some have argued that a utilitarian consequentialism does succeed in establishing rational moral ends (see Harris, 2010), Greene's argument, which favors utilitarianism over deontology, is based only on the consequences that these theories instantiate. On this point, many researchers and philosophers have argued that moral decision-making, based on deontological ethics, leads to predictably sub-optimal results (e.g., Baron & Ritov, 2009; Giubilini et al., 2018; Greene et al., 2009; Sunstein, 2005).

It is an analysis of these expected results that precipitate Greene's conclusion, that, utilitarian consequentialism is uniquely poised to prevent tragedies of the commons. Indeed, utilitarianism is deeply pragmatic and prioritizes results over all else (i.e., morality is defined as the greatest good for the greatest number). Thus, in commons dilemmas, if individuals act in a way that brings about poor results, they are, ipso facto, behaving immorally according to utilitarian ethics. However, while factors that predict

25

utilitarian over deontological judgments have been identified, it is unclear whether

utilitarian judgments are actually indicative of a larger commitment to utilitarian

principles or rather, a reflection of immoral and antisocial traits (for an analysis of this

debate, see Conway et al., 2018; Everett & Kahane, 2020; Kahane et al., 2015).

  While many psychological studies have investigated human behavior in social

dilemmas, like the tragedy of the commons (for a review see Van Lange et al., 2013), the

current, and other recent reviews (i.e., Clarkson, 2022), have failed to identify a single

study that has explicitly collected data on participants' moral judgments and then had

them play actual commons games. While this is not surprising, as the area of research

that has investigated associations between moral judgment and behavior in games is

minimal, it remains an important and undetermined research question. From reviewing

this area, which has mostly used trust games, it is difficult to predict if utilitarians would

behave in ways better or worse suited to divert tragedies of the commons. In fact, recall

that much of the research shows that people are less trusting of utilitarians in situations

where cooperation is important to secure a mutually good result. Thus, despite Greene's

argument, people seem to expect that utilitarians are motivated by selfishness more than

the greater good. Indeed, past research has shown that people who make sacrificial

judgments in hypothetical moral dilemmas are more narcissistic (Bartels & Pizarro,

2011), more likely to believe that cheating is appropriate (Crittenden, 2009), and show

less empathic concern (Partil & Silani, 2014).

  While much of this research indicates that utilitarian responders will fail to be

better at successfully navigating through commons dilemmas, results are mixed as to

whether utilitarian responders are less cooperative in game situations. For instance,

research that has connected personality traits to behavior in economic games has found that extroverts exhibited more selfish individual behaviors, which led to worse community outcomes (Prentice & Sheldon, 2014; Sheldon & McGregor, 2000). Relatedly, recent evidence has connected extroversion to fewer utilitarian judgments in sacrificial dilemmas (Tao, 2020). Still, other research has found that emotions, such as guilt, are important predictors of reduced resource extraction (Tarditi et al., 2018).

Since it is theoretically deontologists, not utilitarians, that are thought to base their moral decisions on emotion, it may be the case that utilitarians will extract more of the common resource pool. This potential result is theoretically supported by separate findings that have associated narcissism with increased utilitarian judgment (Bartels & Pizarro, 2011) and damaging patterns of resource extraction in commons dilemmas (Campbell et al., 2010). Thus, as a whole, past studies support conflicting expectations about how utilitarian and deontological responders will behave. In fact, many of the collected data points tentatively support the existence of a perception gap, where utilitarians are incorrectly expected to be worse egalitarian operators than their non-utilitarian peers. Thus, when considering the theoretical arguments of Greene and the limited empirical research that has investigated how moral judgment predicts behavior in game situations (and the conflicting results of that research), it is difficult to make specific predictions pertaining to the topic at hand.

While specific predictions may be hard to come by, any research that investigates how moral judgment predicts behavior in commons games must be sensitive to several key outcomes. First, do moral judgments predict behavior in games that simulate the tragedy of the commons? Specific predictions will be offered in the next chapter on a

27

study-by-study basis; however, a main effect of moral responding is generally predicted. A second question gets at the direction of this effect, specifically, do rational moral judgments predict behaviors that are less susceptible to bad outcomes, i.e., a tragedy for the commons? Here, no general predictions are on offer. While specific predictions will be tentatively addressed in the next chapter, the directional effects of morality's influence on behavior in these situations are expected to vary depending on several other factors. Briefly, these factors speak to the circumstances of the game itself and the behavior of the other players within them. Thus, no specific predictions will be offered here except to say that moral responding is expected to interact with other factors that have been programmed into the commons games.

In commons games, participants typically play multiple rounds where each player can extract some amount of a pooled resource in every round. For instance, in Sheldon and Mcgregor (2000), participants played a multi-round commons dilemma game in groups of four. Each participant acted as a "company" that could harvest 0-10 hectares of forest every round. Before deciding how much to harvest, participants were informed that the forest totaled 200 hectares, and that, after each round, the forest would replenish at a rate of ten percent. Thus, if participants prioritized short-term goals (i.e., harvesting as much as possible in early rounds) the forest would fail to replenish, and everyone would be worse off in the long run (resulting in a tragedy for the commons). The scenario presents participants with an easily understood commons dilemma. Due to this reason, as well as its frequent use in prior research, this dilemma and slightly modified versions of it were used in several pilot studies.

Pilot Research.

Pilot Study One. Data on 202 participants were collected via MTurk to establish an initial relationship between moral judgment and behavior in tragedy of the commons-like dilemmas. As part of a larger and separate study, participants responded to a variety of scales and scenarios that are not relevant to the current project. However, all participants first responded to several sacrificial moral scenarios (from Conway & Gawronski, 2013) and then were introduced to a common's dilemma (from Prentice & Sheldon, 2014; Sheldon & McGregor, 2000).

Method. The main purpose of the original study, from which this pilot data comes, was to associate differences in decision-making and social judgment with different levels of autistic traits in people. For this reason, over half the data set (60.4%) was self-identified as autistic. To measure individual differences in autistic-like traits, the Iowa Screener (Gaeth et al., 2016) was administered to all participants. For the current project, all analyses were conducted while controlling for autistic-like traits. In keeping with the designs of Conway and Gawronski (2013), all participants responded to 20 sacrificial moral scenarios that allowed for a separate evaluation of a participant's utilitarian and deontological inclinations (see Appendix A). Then, in keeping with the designs of Sheldon and McGregor (2000), participants read the foresting dilemma and then rated their apprehension, desire to profit, and first-year bid amount in reference to the scenario (see Appendix B). While the scenario can go on for multiple rounds (allowing for multiple participant bids), participants only played one initial round in this study. Considering the prior literature reviewed in the previous chapters, it was hypothesized that there would be significant associations between participants' moral judgments and their bid amounts (from 0-10 hectares) on the foresting dilemma.

29

Results. To evaluate the relationship between specific patterns of moral

responding and bid amounts in the foresting dilemma, partial correlations were run

between the continuous variables that tracked moral responding (i.e., participants'

frequency of utilitarian decisions, deontological, and utilitarian inclinations), and their

selected bid amounts in the foresting dilemma, while controlling for autistic

characteristics as measured by the Iowa Screener. Results showed that participants' bid

amounts were significantly correlated with their frequency of utilitarian judgments $r =$

.141, $p = .049$, and their utilitarian inclinations $r = .165$, $p = .021$. On the other hand,

deontological inclinations were not related $r = -.037$, $p=.608$.

Discussion. This pilot study was conducted to test a fundamental research

question, that is, whether patterns of moral judgment are related to behavioral differences

in tragedy of the commons dilemmas. As hypothesized, these data provide the first

evidence that this is indeed the case, as participants' frequency of utilitarian judgments

and calculated utilitarian inclinations, were associated with higher bid amounts in the

foresting dilemma. While no such relationship was found between deontological

inclinations and bid amounts, the direction of the descriptive relationship may also be

informative. That is, higher deontological inclinations descriptively correlated with lower

bid amounts. Thus, in addition to showing that an individual's patterns of moral judgment

matter when it comes to behavior in commons dilemmas, these results also tentatively

indicate that a utilitarian-based morality may not lead to better outcomes, despite

countervailing theoretical claims (see Greene, 2013). Finally, while the resultant

correlations are relatively weak, it should be noted that this study only collected

participants' first-year bid amounts. It is possible, and even likely, that playing additional

rounds may have increased the power of this analysis, and potentially, the robustness of these effects.

Pilot Study Two. While pilot study one showed that a participants' pattern of moral judgment (i.e., increased utilitarian judgment and utilitarian inclinations) is associated with behavior in the foresting dilemma (i.e., increased bid amounts), it only measured participants' bids over a single round. Although not abnormal, (see, Sheldon and McGregor (2000) who associated personality traits with bid amounts after only collecting bid data in a single round) designs with multiple rounds, where the community resource pool can actually expire, better simulate tragedy of the commons dilemmas. To investigate whether the observed effects in pilot study one generalized to commons scenarios with multiple rounds, a second pilot study was conducted with 95 participants at the University of Toledo.

As part of a larger study, participants also responded to a variety of scales and materials that are not relevant to the current project. However, one of these measures, the Rasch-based numeracy scale, is relevant and was used to track participants' ability to understand and manipulate numerical information (Weller et al., 2013). Since participants are asked to make numerical decisions, which balance desires to profit against threats of depleting the forest, numerical ability was thought to perhaps correlate with certain patterns of behavior. Finally, all participants first responded to sacrificial moral scenarios (from Gleichgerrcht & Young, 2013) and then were introduced to a slightly altered version of the common's dilemma used in pilot study one.

Method. The key aim of the original study, from which the pilot data originates, was to investigate whether individual differences (e.g., numeracy, degree of handedness)

interact with emotion to predict certain patterns of decision-making in a ratio bias task. After participants performed these tasks, they then responded to two moral scenarios (see Appendix D) and participated in the foresting dilemma. This study differs from pilot study one in two important ways. First, participants responded to a different set of moral dilemmas. While these dilemmas do not allow for a separate analysis of utilitarian and deontological inclinations, they are advantaged in that participant responses can be used to classify individual responders along a moral quadrant (see Appendix E). This classification technique has been used previously (see Gleichgerrcht & Young, 2013), and identifies participants as utilitarian, typical, weird, and deontological.

The second way that this study differs from pilot study one is in the changes made to the foresting dilemma itself. Specifically, participants played two rounds (i.e., made two bids) in the foresting scenario to get preliminary data regarding the effect that playing multiple rounds has on the relationship between moral judgment and bidding behavior. As this was a pilot study, within a larger study that used many different measures, only two rounds could be included. However, to make the risks of a potential tragedy of the commons (i.e., a total depletion of the forest) more salient, the other three "companies" always bid the maximum amount (i.e., 10 hectares each). Thus, in pilot study two, data were collected on participants' moral judgments, first-round bid amount, participants' rated levels of apprehension and desire to profit, and additionally, their second-round bid amounts (after finding out what the other "companies" collectively bid in Round 1, i.e., after feedback). Finally, to capture how participants responded to the other companies over-harvesting the forest, participants' bid differences, between round one and round two, were calculated.

For pilot study two it was predicted that a main effect of moral judgment on behavior would be observed in the commons dilemma. Informed by the results of pilot study one, it was also predicted that participants classified as utilitarians would make larger bids than deontological participants overall. Third, it was hypothesized that utilitarians would show greater differences in bid amount between rounds one and two. This is supported by a straightforward rationale, that utilitarians are theorized to primarily be focused on consequences, whereas, deontologists are focused on principle. Thus, it was anticipated that utilitarians would show a greater response to the aggressive (i.e., maximum) bids of the other three companies, as the bids of those companies would change the expected consequences of later bids on the forest. No other predictions regarding the other measures or individual difference factors were made.

Results. Based on their responses to the moral dilemmas, participants were classified as follows: 52 typical, 25 utilitarian, 14 deontological, and 4 weird. In keeping with the design of Gleichgerrcht and Young (2013), the four participants classified as weird were excluded from future analyses. To test the relationship of moral judgment on participants' total bid amounts, a one-way ANOVA was performed. There was not a significant effect of moral classification on bid amount, $F(2,88) = .910$, $p = .406$, $\eta^2 = .020$. An additional one-way ANOVA was performed to test the effect of participants' moral judgments on their bid differences between round one and round two. This time a marginally significant effect was found, $F(2,88) = 2.79$, $p = .067$, $\eta^2 = .060$. On average, participants increased their bids from round one to round two (M= .78 hectares). However, while typical and utilitarian responders increased their bids (M =.88 and M= 1.17 respectively), deontological responders decreased their bids (M= -.36). Post-hoc

analysis revealed that while utilitarians differed from deontologists ($p = .063$), they did not differ from typical responders ($p = .817$). In addition, typical responders did not differ from deontologists ($p = .107$).

An exploratory MANOVA was run to check for differences between participants' moral judgments and the other measures in the foresting scenario. Participants' bid amounts for rounds one and two and their rated levels of apprehension and desire to profit were entered into the model as dependent measures. No differences were found between moral responders, $F(8, 170) = 1.25$, Wilk's $\Lambda = .275$, $\eta^2 = .055$.

Finally, participants' scores on the Rasch-based numeracy scale did not correlate with participants' bids in their first ($r = .06$, $p = .55$) or second year ($r = .02$, $p = .84$) decisions. No correlation was found between numerical ability and any other dependent variable related to the foresting scenario or moral responding.

Discussion. As a preliminary follow-up to pilot study one, pilot study two was designed to investigate whether the previously observed differences between participants' moral judgments and their bidding behavior in the foresting scenario were generalizable to a different set of moral dilemmas (and classification technique) and a slightly modified commons dilemma. The results partially confirmed hypotheses. The first (i.e., that a main effect of moral responding would be observed) and second hypothesis (i.e., that utilitarians would exhibit larger bids across both rounds) were not supported. While utilitarian participants did descriptively make larger bids than deontological participants in total, this difference was not significant. However, this lack of a significant effect may be a product of the study being underpowered ($\beta = .20$) due to in-person data collection being halted with the emergence of the Covid-19 pandemic.

The third hypothesis, that utilitarians would exhibit a larger change than deontologists, between their first and second-round bids, was confirmed by the detection of a marginally significant difference. Again, this analysis was also found to be underpowered ($\beta = .54$), and as such, it is possible that collecting additional data would have strengthened the observed effect, especially considering the reported medium effect size with an observed partial eta squared of $\eta^2 = .060$. While no hypothesis was made regarding the direction of the effect, it may be surprising that the observed difference between utilitarian and deontological responders' changes in bid amount (between round one and round two) was driven by a tendency for utilitarians to increase their bids, while deontologists decreased their bids, in the second round.

On its face, this result tentatively indicates that utilitarian responders may not actually be concerned with avoiding tragedies for the commons. However, these results can be explained by multiple and conflicting interpretations. First, it could be that utilitarians increased their bids in round two because they were motivated to selfishly gather as many resources as possible for themselves. This interpretation is supported by some research reviewed in chapter three (e.g., Bostyn & Roets, 2017) that found utilitarians to be less cooperative in trust games. However, there are conflicting results in this literature, and it is not clear whether utilitarians are less cooperative or if they are merely perceived to be less cooperative by others. In addition, utilitarians did not actually make larger bids in round one than deontologists. Finally, utilitarian responders did not show an increased desire to profit compared to deontological responders. Thus, it seems that utilitarians' and deontologists' differences in bidding behavior in the second round

could be motivated by something other than an uncooperative or selfish desire to collect more resources.

One additional explanation may be that utilitarians are increasing their bids due to gamesmanship or other strategical motivations. Recall, that in the modified version of the foresting dilemma, used in pilot study two, the other three companies always took the maximum amount in round one. After learning this, participants then had to make their bids for round two. It could be that utilitarians were engaging in a tit-for-tat strategy to get the other companies (who had extracted the maximum amount in the prior round) to back down and reduce their bids in later rounds. Thus, these results cannot determine conclusively whether utilitarians increased their bids due to purely selfish motivations or matters of strategy. What is more, these results also cannot fully address whether utilitarians, deontologists, or typical responders, are better suited to avoid tragedies for the commons. However, these results do have critical importance. Just like in pilot study one, they show that different types of moral responders behave differently in commons dilemmas. For this reason, these pilot studies are valuable, as they, along with the reviewed literature, inform predictions about how utilitarians and deontologists will behave in commons dilemmas in future research. An overview and proposal of this planned future research is the subject of the next chapter.

Chapter Five

The Research on Moral Judgment and Behavior in Commons Dilemmas

A specific but far-reaching question has been raised about whether moral judgment predicts behavior in situations that simulate the tragedy of the commons. The importance of commons dilemmas was made clear in chapter one, and it was recognized that morality might have a role in our behavior in these situations. After reviewing normative theories of ethics in philosophy and certain well-established findings in moral psychology in chapter two, it was argued that additional research should be conducted on moral behavior in situations with greater ecological validity. In chapter three, the research that has investigated the association between moral judgment and behavior was reviewed in detail. Finally, in chapter four, research that has associated participants' moral judgments with behavior in games was closely reviewed before addressing arguments asserting that increased rational (in this case, utilitarian) moral judgments should be better suited to avoid tragedies of the commons. After conducting this review and performing two pilot studies, some significant questions remain. First, do moral judgments predict behavior in games that simulate the tragedy of the commons? Second, do rational moral judgments predict behaviors that are less susceptible to bad outcomes, that is, a tragedy for the commons? Thus, three additional experiments were conducted to offer answers to these and other related questions.

Programmatic Overview

To summarize, the pilot research found moral judgments to be associated with behavior (i.e., resource harvesting) in limited versions of commons dilemmas that have been used in prior research. These results, however, only presented participants with an

opportunity to harvest resources over a one- or two-round design. While interesting, these

results cannot speak fully to how participants might behave across many rounds – when

resources can be *fully* harvested by other players in the game. What's more, the pilot

research cannot fully appreciate how different types of moral responders are influenced

by different opposing player strategies. While pilot study two did require participants to

make a second bid after seeing the bids of the other players in round one, this only offers

limited information.

First, participants in pilot study two only had a single opportunity to adjust to the

bids of the other players. In a situation such as this, where the risk of eliminating the

forest should be salient, it is clear that additional rounds offer more complete information

about how participants' behaviors might adapt to different opposing strategies over time.

Second, pilot study two presented participants with only one type of opposing player

strategy. Recall, that participants were told that the other players took the maximum

allowable amount in the previous round (let us call this an Aggressive opposing player

strategy). Of course, this is not the only strategy that a decision-maker is likely to face in

such situations. Instead, participants would likely be exposed to an array of different

strategies, and certain patterns of moral judgment might be associated with behaviors

when faced with some opposing strategies and not others. In recognition of these points,

all three experiments that follow required participants to play a commons dilemma game

for up to 12 rounds and with different opposing player bid strategies (i.e., Aggressive,

Equilibrium, Soft). While the design of the experiments was kept constant in these

respects, other variables were changed between them.

Since it has already been established that as a field, moral psychology has been disadvantaged by studying participants' moral judgments in situations that have low external validity, experiment two presented participants with a new, modified version of the commons dilemma, that is, the Covid Reefs dilemma (see Appendix C). Additional details will be presented in later sections; however, in the Reefs dilemma, participants' were asked to make decisions about harvesting an important resource that could save people from dying from Covid-19. All data were collected during the heart of the Covid-19 pandemic and before the vaccine was readily available in the United States of America. Thus, experiment two stands in stark contrast to experiments one and three. It presents participants with another scenario that may feel closer to the real world. This effort to increase the relatedness of the game to the real world was also a primary aim of experiment three, which is differentiated by its attaching of real-world consequences (i.e., actual monetary incentives) to participants' behaviors in the standard commons dilemma. Finally, other related exploratory variables were also included across these three experiments.

The analysis for Experiments 1-3 will have a uniform structure. That is, for each experiment, the analyses will first test participants opening bids (in round one only) with a one-way ANOVA. Then, several two-way ANOVAs will be conducted to test for main effects of and interactions between the moral judgment and the opposing player strategies on participants' behaviors in the game. Finally, a hierarchical linear regression will be conducted across three steps. In step one of the model, the opposing player condition will be dummy coded and used as the lone predictor. In the second step of the model, participants' continuous utilitarian and deontological process strengths will be entered. In

the third step, participants' relatedness to nature and empathic concern scores will be added to the model. Once these analyses have been reported, the results of other exploratory analyses, such as partial correlations between participant bidding behavior and their moral processes or self-report scales will be assessed, where appropriate. With this now sketched out, the next section will provide detailed descriptions about the method and results of Experiments 1-3 before offering general conclusions and an assessment of the types of future research needed in the subsequent chapter.

Experiment 1. As an extension of pilot studies one and two, experiment one was designed to detect whether an effect exists between moral judgment and behavior in a commons dilemma. In addition, the direction of effect, and whether moral judgment interacts with other variables (such as the behavior of other players) were also key questions.

   Method. Experiment one commenced across three phases. In phase one, participants completed a short demographics questionnaire and several surveys to assess potentially relevant individual differences, i.e., the Edinburgh Handedness Inventory (which assesses handedness consistency and is a proxy for degree of hemispheric lateralization; Oldfield, 1971), the Nature Relatedness Scale (which assesses one's concern about the environment, which might be important in decisions to harvest a forest or an ocean reef; Nisbet & Zelenski, 2013), the Social and Conservatism Scale (which assesses political orientation; Everett, 2013), and the Empathic Concern subscale from the Interpersonal Reactivity Index (Davis, 1980). In phase two, participants provided judgments to two sets of moral dilemmas (i.e., Conway & Gawronski, 2013 and Gleichgerrcht & Young, 2013). By assessing a participant's moral judgments to the

Conway and Gawronski (2013) set, a continuous value of the participant's Utilitarian and Deontological process strengths was calculated. A participant's score on this measure assesses the strength of their Utilitarian or Deontological Process. At the same time, a participant's responses to the Gleichgerrcht and Young (2013) set allow for a categorical assessment of their moral disposition (i.e., Utilitarian, Deontological, Typical, Weird). Additional details on these variables can be seen in chapter two. Finally, in phase three, participants played in one of three conditions of the foresting commons game, in which the behavior of the other players in the game (all other players were computer players) was programmed to exhibit one of three fixed strategies (i.e., Aggressive, Equilibrium, or Soft) throughout all rounds of the game. Using Qualtrics software, this new game was modeled after the old paper and pencil game created by Sheldon and McGregor (2000) and was programmed to appreciate these three different conditions, which are described below.

Recall, that in the foresting game, in every round, players decide how many hectares (i.e., 0 – 10) they would like to harvest from the forest, which initially is 200 hectares in size, and replenishes at a rate of 10 percent between each round. The game can be played for 12 rounds unless the forest is completely harvested in an earlier round. In the aggressive condition, computer players collectively bid between 27-30 hectares in all rounds (irrespective of the participant's behavior). Bids of this size cannot be sustained over 12 rounds, and thus, the forest is guaranteed to expire before the 12th round. Thus, participants' behavior in this condition indicates how they respond to extremely aggressive opponents that are operating irrationally and perhaps, even in bad faith. In the equilibrium condition, computer players collectively bid between 14-16

hectares. By equilibrium, it is meant that the computer players' bids are rational, in that, if the bids of all four players were just so – between 4-6 hectares for each player per round, the forest would never be exhausted and the game could perpetuate indefinitely. Thus, in this condition participants can cooperate with the commons (the other players) by also biding at equilibrium, and therefore, behavior here indicates the participant's willingness to cooperate when it is rational. Finally, in the Soft condition, the computer players' bids collectively range from 6-8 hectares per round. If the participant follows suit and makes bids in each round of between 2-3 hectares, the forest will not be harvested below its replenishment rate. These bids are too small to reduce the forest's size and thus, a participant's behavior in this condition indicates how willing the participant is to cooperate when it is irrational.

In keeping with prior designs (e.g., Prentice & Sheldon, 2014) participants' average bid amount per round and total bid amount across all rounds were recorded as dependent variables. These data were used to test several hypotheses. First, a main effect of moral judgment on participants' behavior in the game was predicted (Hypothesis 1). A series of secondary hypotheses were also offered. A significant interaction between participants' moral classification and opposing player condition was predicted, with utilitarian participants predicted to make larger bids in the aggressive and soft conditions (Hypothesis 2).

The logic for this is straightforward. In this context, utilitarians are generally thought to be motivated by consequences. Since the other player's (i.e., the bots) behaviors will alter the expected consequences of the human players bid (e.g., in the soft bid condition participants can take a larger number of resources without putting the forest

in jeopardy) utilitarian participants should make larger bids. Such actions will only improve outcomes, as this would increase the players' profit at no long-term cost to the forest. However, deontological participants, who might be making decisions based on a principle of fairness, are expected to be less likely to increase their respective bids. Similar logic holds in the aggressive bid condition, where utilitarian participants were expected to reduce their bids for the sake of long-term consequences and deontological participants were expected to remain steady. However, such predictions must be tempered by an acknowledgement that utilitarians have not always been shown to be motivated by securing the greatest-good for the greatest number (but instead sometimes appear to be motivated by selfishness or – the greatest good for them). Even still, since utilitarians are theoretically more sensitive to consequences, their behaviors should update to reflect the changing situation (between bid conditions). Deontologists, on the other hand, who seem to be more motivated by principle, may be less inclined to adjust their behavior between conditions.

Across all conditions, it was hypothesized that the strength of participants' Utilitarian and Deontological processes would show opposite correlations with their behavior in the commons game. Specifically, the Deontological process was predicted to negatively correlate with that same variable (Hypothesis 3), while the Utilitarian process was predicted to positively correlate with resource extraction (Hypothesis 4).

Results. A total of 258 undergraduate students participated in Experiment One through the Universities Sona System account on Qualtrics. Following the procedure of other methods, three participants were removed from analysis for failing quality checks

(i.e., they were classified as Weird in the moral categorization set). This left 255 participants for analysis.

Descriptive statistics for all subject variables can be seen in Table 1. Two 3-level one-way ANOVAs were performed on participants' Moral Classification: (i.e., Utilitarian, Deontological, Typical) to detect differences in desire to profit and apprehension between these groups. No differences were found. A one-way 3-level (Moral Classification) ANOVA was then conducted to detect differences in participants' bids in the opening round of the game before their bids were influenced by the opposing player behavior condition (starting from round two on). As predicted, a significant effect of moral classification was found, $F(2,254) = 3.22$, $p = .04$, $\eta2 = .03$. Tukey post hoc comparisons did not reveal any significant differences between the three groups, but the Utilitarian group extracted marginally more resources than the Typical group ($p = .08$) and the Deontological group ($p = .10$) in round one. Statistics for the means and standard deviations associated with this analysis can be seen in Table 2.

Two separate 3 (Moral Classification) X 3 (Opposing Player Condition: Aggressive, Equilibrium, Soft) ANOVAs were conducted on participants' total resource extraction and their average bid amounts across all played rounds in the foresting game. On resource extraction, a main effect of Opposing Player Condition was found, $F(2,246) = 36.43$, $p < .01$, $\eta2 = .05$. Post-hoc Tukey analyses found that all bid conditions significantly differed from each other (see Figure 1), with the most resources being harvested in the soft condition and the least in the aggressive condition. No main effect of Moral Classification, however, was found, $F(2,246) = 1.10$, $p = .34$, $\eta2 = .01$. However, a significant interaction was found between these variables, $F(2,246) = 2.88$, $p = .02$, $\eta2 =$

.05. Post hoc analysis of simple effects showed that the utilitarian group harvested more

resources in the equilibrium condition (p <.01) than the deontological group.

The second ANOVA, on participants' average bid amounts, showed a significant

main effect of Opposing Player Condition, $F(2,246) = 10.94$, p <.01, $\eta2 = .05$, on

participants' behaviors. Post-hoc Tukey analyses found no differences between these

conditions, although bids were descriptively the largest in the soft condition. No effect of

Moral Classification was found, $F(2,246) = 1.32$, $p = .27$, $\eta2 = .01$. However, there was a

marginally significant interaction between these variables, $F(2,246) = 2.15$, $p = .08$, $\eta2 =$

.03. Post hoc analysis of simple effects showed that the utilitarian group harvested more

resources in the equilibrium condition (p <.01) than the deontological group.

A hierarchical linear regression was conducted on participants bidding behavior

across all rounds of the game. The first step in the model was significant, $F(2,252) =$

43.05, p <.01 $R^2 = .26$, indicating that opposing player condition significantly predicted

participant's bid decisions. The second step in the model was also significant, $F(2,250) =$

4.82, $p < .01$, $\Delta R^2 = .03$, indicating that participant's moral processes predicted their bid

decisions. Finally, the third step in the model was also significant, $F(2,248) = 7.86$, p

<.01, $\Delta R^2 = .06$, indicating that participants' relatedness to nature and empathic concern

predicted their bid decisions. Full results can be seen in Table 3.

Finally, partial correlation analyses were conducted between participants' biding

behavior and several other continuous variables (i.e., Utilitarian and Deontological moral

process strength, empathic concern, relatedness to nature, political orientation,

handedness). Participants assigned Opposing Player Condition were dummy coded and

controlled for in these analyses. Participants' Deontological Process strength negatively

correlated with total resource extraction, $r = -.14$, $p = .02$, and average bid per round, $r = -.16$, $p = .01$. In contrast, participants Utilitarian Process showed a marginal negative correlation with total resource extraction, $r = -.11$, $p = .09$ but did not correlate with the average bid per round. Full results can be seen in Table 4.

      Discussion. As the first step in this larger project, experiment one was designed to provide an in-depth analysis of how participants' moral judgments might predict moral behavior in game situations that simulate the tragedy of the commons. Several predictions were offered. The first hypothesis was supported, as participants' moral judgments did predict differences in behavior in the foresting game. That is, participants moral judgments (specifically their moral classification) predicted differences in their first-year bids (or their opening, first round strategies) and significantly interacted with the bid condition variable throughout the entire game, where utilitarians made larger bids in the equilibrium condition than deontologists.

      Interestingly, neither two-way ANOVA (on total resource extraction or average bid type) showed a significant main effect of moral type. While there was a significant interaction between participants' moral classification and opposing player strategy on average bid amount on total resource extraction, this seems to have been driven by utilitarians making larger bids in the equilibrium condition. Thus, hypothesis two, that a significant interaction would be found between these variables, driven by larger utilitarian bids in aggressive and soft conditions was only partially supported. That is, the hypothesized interaction was found between these variables, but it was driven by differences in the equilibrium condition. It is possible that the moral classification procedure, which labels participants as utilitarian, deontological, or typical, lacks the

46

necessary sensitivity to detect subtle differences in moral judgment that could predict differences in behavior in commons games. While participants' moral type did predict differences in the opening round, these differences appear to have been overshadowed by the large effect of the opposing player conditions when considering behaviors throughout the game.

The measurement of participants' utilitarian and deontological processes represents a more sensitive assessment of participants' moral judgments. Participants' utilitarian and deontological processes were used as key predictors in a three-step hierarchical regression. All three steps in the model were significant. Importantly, in step two, participant's behavior in the foresting game was significantly predicted by their utilitarian (t= -2.09, p = .04) and deontological processes (t = -2.58 p= .01). Thus, while hypothesis three was supported (i.e., a stronger deontological process did predict reduced resource extraction), hypothesis four (i.e., a stronger utilitarian process did not predict increased resource extraction) was not only not supported, but an inverse relationship was found. That is, just as stronger deontological processes predicted decreased resource extraction, so too did stronger utilitarian processes. These results indicate something interesting, that is, stronger moral processing of either type is predictive of less resource extraction in the foresting game.

While this may not be surprising, it does offer some tentative information about how different types of morality (as measured by participant's moral judgments) relate to behavior in commons dilemmas. Recall, that there exists a camp of researchers and philosophers that think deontological processing to be insufficient (or at least less than optimal) to avoiding tragedies of the commons. With these doubts held in one hand, they

have commonly offered utilitarian processing as a long-term solution, in the other hand. Of course, it would be ridiculous to argue that anything conclusive about the long-term quality of their arguments has been shown here. It is, however, notable that increased deontological and utilitarian processing both predicted reduced resource extraction in the game. While something interesting has no doubt been shown here (i.e., it is the first time research has found that specific types of morality predict behavior in a commons dilemma), several caveats must be made.

Importantly, the stakes in the foresting game were low. Participants' behaviors here are connected to the extraction of fake resources from a fake resource pool. To be clear, the lack of real-world consequences being connected to participants' behaviors is not atypical in the context of a larger research area that associates participants' judgments or decisions with behavior in monetary games (Hassovests, 2014). Instead, this seems to be normal practice. Granting this, it is still the case that the hypothetical consequences are low. What is the imagined tragedy for the commons in the foresting game? The forest is annihilated but what does this really mean? The tragic nature of causing a tragedy for the commons in the foresting game may not have been salient. This issue is seriously considered and addressed in experiment two.

In addition to the moral variables, other individual variables (e.g., relatedness to nature) also significantly predicted participants' behavior in the foresting game. Of these, participants' relatedness to nature ($t = -2.49$, $p = .02$) and political conservatism ($t = 2.61$, $p = .01$) were significant predictors of behavior. These results make sense, in that, it seems reasonable that participants' harvesting of the forest would decline as their relatedness to nature increases. It is also unsurprising that as participants' conservatism

increased that their profit-seeking behavior would increase in the game as well. While these findings speak well to the external validity and generalizability of this research, it is notable that empathic concern did not predict behavior. This might indicate that our concerns addressed in the prior paragraph (about whether the tragic nature of a tragedy of the commons was salient) are legitimate, especially since empathic concern should be a motivating factor to avoiding community-wide suffering (i.e., a tragedy for the commons).

Experiment 2. While experiment one showed for the first time that specific patterns of moral judgment predict differences in behavior in a common's dilemma, a key shortcoming was identified. That is, the external validity of the results from experiment one was called into question for two related reasons. First, participants' behaviors were not connected to real-world consequences. Second, the nature of the commons game itself may not have succeeded in making the tragic nature of a tragedy for the commons salient. This possibility is consistent with the fact that empathic concern did not significantly predict participants' behaviors in the foresting game. Thus, experiment two was designed to present participants with a scenario with greater external validity, where the tragedy at risk for the commons was more salient and severe.

Method. Experiment two commenced across three phases. In phase one, participants completed a short demographics questionnaire and several surveys to assess potentially relevant individual differences: the Edinburgh Handedness Inventory (Oldfield, 1971), the Nature Relatedness Scale (Nisbet & Zelenski, 2013), the Social and Conservatism Scale (Everett, 2013), and the Empathic Concern subscale from the Interpersonal Reactivity Index (Davis, 1980). In addition, participants also answered a

single-item question about their perceived seriousness of the Covid-19 pandemic. In phase two, participants responded to the same sets of moral dilemmas as in experiment one. Finally, in phase three, participants played in one of three conditions of a new commons game (i.e., the Covid Reefs dilemma), in which the behavior of the other players in the game (all other players were computer players) was programmed to exhibit one of three fixed strategies (i.e., Aggressive, Equilibrium, Soft) throughout all rounds of the game. Using Qualtrics software, the Covid Reefs dilemma was modeled after the foresting dilemma used in experiment one but with several key alterations.

Recall, that in the foresting game, in every round, players decided how many resources (i.e., 0 – 10) to harvest from the forest, which initially was 200 units of resources in size, and replenished 10 percent between each round for up to 12 rounds. While these basic features were held constant in the Covid Reefs dilemma, there are several notable changes. The first of these changes is in the role that the participants play in the game itself. In the Covid Reefs game participants no longer act as CEOs attempting to harvest a resource for profit but instead act as heads of state attempting to harvest a resource to save the lives of citizens in the middle of the Covid-19 pandemic. Specifically, in the Covid Reefs scenario participants are told that a cure for Covid-19 has been found in a rare species of coral reefs located in international waters, of which, they can harvest to extract the cure, to subsequently save the lives of their infected citizens.

Finally, one other key addition differentiates these two dilemmas. In the Covid Reefs dilemma, participants are told that 10,000 of their country's (and 10,000 of each of the other country's) citizens are dying monthly from Covid-19, and that, they can expect to save about 1,000 of their citizens for every acre of the reef that they harvest per round.

Thus, if they harvest the full number of acres allowed per month (i.e., 10 acres) they can save the lives of all their citizens, who would have died of Covid-19 during that particular month. Of course, if participants (acting as heads of state) decide to continually harvest the maximum number of acres per month, they increase the risk of depleting the reef in the early rounds, quickly leading to a tragedy for the commons. Thus, in certain experimental conditions, participants must allow some people to die in earlier rounds to save more citizens in the long run. For this reason, the reef scenario is an example of a disaster triage situation, where participants must balance a willingness to accept mass casualties to promote a greater good (see Petrini, 2010 for an analysis of these situations). Thus, the Covid Reefs dilemma has been designed to make the tragic nature of a tragedy of the commons more salient, and the desire to avoid a tragedy for the commons more motivating.

In keeping with experiment one participants' average bid amount per round and total bid amount across all rounds were recorded as dependent variables. Like experiment one, there were four hypotheses. First, a main effect of moral judgment on participants' behavior in the game was predicted, with deontological participants making the largest bids (Hypothesis 1). Hypothesis one was driven by the recognition that deontologists will likely find allowing citizens to die in the short term (to save more lives in the long term) to be less acceptable. This is something utilitarians should be more willing to do, considering how they respond to typical sacrificial dilemmas. Second, the strength of participants' moral processes was hypothesized to predict their behavior in the game (Hypothesis 2). Specifically, the strength of the utilitarian process was predicted to show a negative relationship with resource extraction in the game (Hypothesis 3). The reasons

for this are straightforward. Utilitarian ethics, which allows for the sacrificing of lives for a greater good, should subsequently predict less resource extraction (since this behavior should lead to more lives being saved in the long run). On the other hand, the strength of participants' deontological processes was predicted to show a positive relationship with resource extraction in the game (Hypothesis 4). Like hypothesis three, this prediction was informed by how participants respond to sacrificial moral dilemmas – i.e., deontology is less sanguine on the prospect of sacrificing a few to save many.

Results. A total of 259 undergraduate students participated in Experiment Two through the Universities Sona System account on Qualtrics. Following the procedure of other methods, eight participants were removed from analysis for failing quality checks (i.e., they were classified as Weird in the moral categorization set). In addition, three participants were removed from analysis for failing other basic attention checks. This left 248 participants for analysis. All data were collected between January and March 2021 – before vaccines against Covid-19 had become readily available in the United States.

Descriptive statistics for all subject variables can be seen in Table 1. Two 3-level one-way ANOVAs were performed on participants' Moral Classification: (i.e., Utilitarian, Deontological, Typical) to detect differences in the desire to save lives and apprehension between these groups. No differences were found. A one-way 3-level (Moral Classification) ANOVA was then conducted to detect differences in participants' bids in the opening round of the game before their bids were influenced by the opposing player behavior condition (starting from round two on). No differences between moral classification groups were found, $F(2,247) = .088$, $p = .92$, $\eta2 < .01$. Statistics for the means and standard deviations associated with this analysis can be seen in Table 2.

Two separate 3 (Moral Classification) X 3 (Opposing Player Condition:

Aggressive, Equilibrium, Soft) ANOVAs were conducted on participants' total resource

extraction and their average bid amounts across all played rounds in the reefs dilemma

game. On resource extraction, a main effect of Opposing Player Condition was found,

$F(2,239) = 67.80$, $p < .01$, η2 = .36. Post-hoc Tukey analyses found that all bid conditions

significantly differed from each other (see Figure 2), with the most resources being

harvested in the soft condition and the least in the aggressive condition. In addition, a

main effect of Moral classification was found, $F(2,239) = 3.09$, $p <.05$, η2 = .03. Post hoc

Tukey analyses revealed that Deontological participants harvested more resources than

typical participants ($p = .049$). No interaction, however, was found between these

variables, $F(2,239) = .24$, $p = .92$, η2 < .01.

The second ANOVA, on participants' average bid amounts, showed a significant

main effect of Opposing Player Condition on participants' behaviors, $F(2,239) = 6.72$, p

<.01 η2 = .05. Post-hoc Tukey analyses again found that all bid conditions significantly

differed from each other. A main effect of Moral Classification was also found on

participants average bid amounts, $F(2,239) = 3.81$, $p = .02$, η2 = .03. Post hoc Tukey

analyses revealed that Deontological participants harvested more resources than typical

participants ($p = .02$). Once again, though, there was no interaction between these

variables, $F(2,239) = .43$, $p = .79$, η2 = .01.

A hierarchical linear regression was conducted on participants bidding behaviors

across all rounds of the game. The first step in the model was significant, $F(2,243) =$

86.91, $p <.001$, $R^2 = .42$, indicating that opposing player condition significantly predicted

participant's bid decisions. The second step in the model was also significant, $F(2,241) =$

3.1, $p < .05$, $\Delta R^2 = .02$, indicating that participants' moral processes predicted their bid decisions. The third step in the model was not significant, $F_{(4,238)} = .99$, $p = .41$, $\Delta R^2 = .01$. Full results can be seen in Table 3.

Finally, partial correlation analyses were conducted between participants' bidding behavior and several other continuous variables (i.e., Utilitarian and Deontological moral process strength, empathic concern, relatedness to nature, politics, handedness, perceived seriousness of Covid-19, opinion about a vaccine mandate for Covid-19). Participants assigned Opposing Player Condition were dummy coded and controlled for these analyses. Participants' Utilitarian Process strength negatively correlated with total resource extraction, $r = -.15$, $p = .02$, and average bid per round, $r = -.18$, $p < .01$. In contrast, participants' Deontological Process did not correlate with total resource extraction or with average bid per round. Full results can be seen in Table 4.

Discussion. To extend the findings of experiment one, experiment two was designed to provide an in-depth analysis of how participants' moral judgments might predict behavior in a hypothetical commons situation where the tragic nature of the tragedy of the commons is increased. Several predictions were offered. The first hypothesis was supported, as participants' moral judgments did predict differences in behavior in the Covid Reefs game. That is, deontological participants (according to their moral classification) harvested more resources in the game than Utilitarian participants and made larger bids on average per round. Thus, these results indicate that deontological participants were less willing to make the rational choice (in respect to long-term consequences) when lives, rather than profits, were the object of sacrifice. Interestingly, this is likely exactly what would be predicted by philosophers and researchers, who decry

the potential for decision-making based on deontological ethics to avoid tragedies for the commons. Thus, unlike in experiment one, this result supports these arguments and casts doubt on the utility of deontological ethics in these situations.

It is interesting, however, that post hoc tests revealed that behavior (total resource extraction and average bid per round) of deontological participants only differed from typical participants and not utilitarian participants. While it was not specifically predicted, one might have suspected that utilitarians would harvest the least number of resources, not typical participants, if they were following a truly utilitarian edict. Indeed, while participants classified as utilitarian made descriptively lower bids than deontological participants, their bids (M = 6.10 acres) were still well above equilibrium (5 acres). In a similar fashion to the findings of experiment one, this result indicates that the moral classification procedure may not be precise enough to capture subtle differences in participants' utilitarian motivations.

For a more sensitive measure of participants' moral sensibilities, their utilitarian and deontological processes were calculated and used as key predictors in a three-step hierarchical regression. The first step of the model, with opposing player condition as the lone predictor, was unsurprisingly significant. More importantly, the second step in the model was also significant, accounting for 1.5% of the variance above and beyond step one (supporting hypothesis 2). Here, participants' behavior in the Covid Reefs game showed a negative relationship with their bidding behavior and was significantly predicted by their utilitarian ($t= -2.44$, $p = .02$) process strength (supporting hypothesis 3). However, the strength of participants' deontological process did not predict their behavior in the game (failing to support hypothesis 4).

Interestingly, the hypothesized direction of the relationship (positive) was not shown in these data. Instead, a non-significant but negative relationship was shown between total resource extraction ($r = .03$, $p = .64$) and average bid amount ($r = .04$, $p =.49$). While non-significant, it appears that greater deontological processing may not necessarily inhibit specific types of sacrificial behavior. Indeed, prior evidence is consistent with this assessment (e.g., Baron & Goodwin, 2020). That is, stronger deontological processing has been linked to a greater preference for inaction in sacrificial dilemmas (Baron & Goodwin, 2020). In the Covid Reefs game, participants must actively bid above equilibrium to avoid sacrificing their citizens in the short term for a greater long-term benefit. Thus, it could the case that deontological processing showed a descriptively negative relationship with resource extraction because this required an active decision to not sacrifice citizens in the short-term by making larger bids in early rounds.

The third step in the hierarchical regression model was not significant. In addition, partial correlation analyses of these variables found a marginal relationship between political conservatism and total resource extraction ($r = .11$, $p = .10$) but did not find any relationship between relatedness to nature, empathic concern, and perceived seriousness of Covid-19 with participants behavior in the game. It is not surprising that the relationship between relatedness to nature and bidding behavior (that was found in experiment 1) fell away, given that resource extraction was motivated by saving lives rather than by profit. It is again interesting, though, that participants' empathic concern failed to show any relationship with behavior in the game. These results are somewhat surprising, as it seems reasonable to expect that at least empathic concern would have

correlated (i.e., positively) with behavior in the Reefs game. However, it is notable that empathic concern did not predict behavior in experiment one (albeit that such a relationship seems less likely considering that desire to profit, not lives, should have been the motivating variable). These results (i.e., the lack of associations between these variables and behavior in the Reefs game) offer increased (but indirect) support for the importance of considering morality in designs that put participants in disaster triage situations, where decisions about life and death must be made on a large scale.

Taken together, the results of experiment two provide additional evidence that morality is a key predictor of behavior in commons games. In the Covid Reefs game, the consequences of what a tragedy would mean for the commons were severe (greater loss of life of one's citizens). As predicted, deontologically classified participants extracted more resources and made larger bids in the game (which would eventually lead to a tragedy for the commons). Also, in line with predictions, stronger utilitarian processing significantly predicted resource extraction, showing a negative relationship. Conversely, stronger deontological processing did not predict resource extraction. Thus, the results of experiment two add initial support to certain philosophers and researchers who have argued that an increased subscription to utilitarian ethics might better allow societies to avoid tragedies of the commons in multiple, pressing real-world domains (e.g., climate change). However, experiments one and two were limited in that, each failed to connect real-world consequences with participants' behavior in the games. This was the main task of experiment three.

Experiment 3. To further test the relationship between participants' moral judgments and behavior, experiment three replicated the methods of experiment one with one key

alteration. That is, in experiment three participants' behaviors in the commons game were connected to real-world monetary incentives. Conducted on Amazon's Mechanical Turk (MTURK), participants were awarded cash bonuses for each hectare of forest that they harvested in the foresting game. Funding for payment was secured by the Department of Psychology Meritorious Research Grant at the University of Toledo.

While real-world incentives are not typically thought to be required for a study to reach the threshold of behavior (e.g., Saltzstein, 1994), studies that connect participants' decisions to real-world consequences are often thought to be of a higher standard. Even so, some may debate whether such studies are in fact measuring behavior. In these cases, one wonders what their standard for a behavioral measure would be. Would behavior only be possible to investigate in real-world situations beyond the psychological laboratory? If so, entire psychological disciplines reporting to study behavior would need to be amended. And do people not behave in games? It seems unnatural and incorrect to call the repeated decisions that people make in games just "judgments". Still, for those of this opinion, let us at least say that the experiments one and two come closer to the mark of studying behavior, and that experiment three comes closer still. Thus, experiment three was designed with several key questions in mind. Namely, do the effects of experiments one and two hold when behavior is connected to real-world consequences, in this case money.

Method. Like experiment one, experiment three commenced across three phases. In phase one, participants completed a short demographics questionnaire and several surveys to assess potentially relevant individual differences: the Edinburgh Handedness Inventory (Oldfield, 1971), the Nature Relatedness Scale (Nisbet & Zelenski, 2013), the

Social and Conservatism Scale (Everett, 2013), and the Empathic Concern subscale from the Interpersonal Reactivity Index (Davis, 1980). In phase two, participants provided judgments to two sets of moral dilemmas (i.e., Conway & Gawronski, 2013 and Gleichgerrcht & Young, 2013). By assessing a participant's moral judgments to the Conway and Gawronski (2013) set, a continuous value of the participant's Utilitarian and Deontological process strengths was calculated. Participants' responses to the Gleichgerrcht and Young (2013) set allowed for a categorical assessment of their moral disposition (i.e., Utilitarian, Deontological, Typical, Weird). Additional details on these variables can be seen in Chapter three. Finally, in phase three, participants played in one of three conditions of the foresting commons game, in which the behavior of the other players in the game (all other players were computer players) was programmed to exhibit one of three fixed strategies (i.e., Aggressive, Equilibrium, Soft) throughout all rounds of the game.

In keeping with the designs of experiments one and two, participants' average bid amount per round and total bid amount across all rounds were recorded as dependent variables. Several hypotheses were made, which broadly predict that experiment three will replicate the findings of experiment one. Thus, a main effect of moral judgment on participants' behavior in the first round of the game was predicted for the initial one-way ANOVA (Hypothesis 1). A replication of the previously observed interaction between opposing player condition and participant's moral classification was predicted (Hypothesis 2) for the planned two-way ANOVA on total resource extraction. Finally, Step 2 (with Utilitarian and Deontological processes entered as predictors) in the hierarchical linear regression model was predicted to be significant (Hypothesis 3),

replicating experiment one. In addition, participants' deontological process strength was predicted to negatively correlate with resource extraction (Hypothesis 4). No correlation between the utilitarian process and resource extraction was predicted, seeing as the previously observed correlation between these variables in experiment one was only marginally significant. However, unlike in experiment one, since real consequences were on the line, Step 3 in the regression model (that includes participants' relatedness to nature) was predicted to not be significant (Hypothesis 5). The notion here is that it seems unlikely that participants' relatedness to nature would predict how they would behave when making decisions about harvesting timber from a hypothetical forest when real money is to be had.

Results. A total of 420 participants were recruited via MTURK in exchange for payment. Participants who successfully completed the study and passed basic quality checks were paid $1.00. In addition, participants were awarded a bonus of $0.02 per hectare they harvested in the foresting game. Thus, participants could earn up to $2.40 (.02 X 10 hectares per round X 12 rounds) as a bonus due to their harvesting behavior in the game. A total of 360 participants successfully completed the survey on MTURK. Since this study was conducted on MTURK, other quality checks were used to ensure the quality of the data. For instance, participants were asked if they completed the survey on an "electronic device", "crystal ball", or a "book". Here, 62 participants failed quality checks of this kind, indicating that they weren't reading the materials carefully. In addition, 35 participants were removed for completing the study in under five minutes. This left 263 participants for analysis.

Descriptive statistics for all subject variables can be seen in Table 1. Two 3-level one-way ANOVAs were performed on participants' Moral Classification: (i.e., Utilitarian, Deontological, Typical) to detect differences in desire to profit and apprehension between these groups. No differences were found.

A one-way 3-level (Moral Classification: Utilitarian, Deontological, Typical) ANOVA was conducted to detect differences in participants' bids in the opening round of the game before their bids were influenced by opposing players (starting from round two on). As predicted, a significant effect of moral classification was found, $F(2,260) = 3.13$, $p < .05$, η2 = .02. Tukey post hoc comparisons showed that deontological participants made larger bids in year one than utilitarian participants ($p = .04$); however, no other differences were found. Statistics for the means and standard deviations associated with this analysis can be seen in Table 2.

Two separate 3 (Moral Classification) X 3 (Opposing Player Condition: Aggressive, Equilibrium, Soft) ANOVAs were conducted on participants' total resource extraction and their average bid amounts across all played rounds in the foresting game. On resource extraction, a main effect of Opposing Player Condition was found, $F(2,254)$ = 66.20, $p < .01$, η2 = .34. Post-hoc Tukey analyses found that all bid conditions significantly differed from each other (see Figure 3), with the most resources being harvested in the soft condition and the least in the aggressive condition. There was no main effect of Moral Classification, $F(2,254) = .60$, $p = .34$, η2 = .01, nor any interaction between these variables.

The second ANOVA, on participants' average bid amounts, showed a significant main effect of Opposing Player Condition $F(2,254) = 6.23$, $p < .01$, η2 = .05 on

participants' behaviors. Post-hoc Tukey analyses again found that all bid conditions significantly differed from each other except for the equilibrium and aggressive conditions ($p =.99$). However, there was no effect of Moral Classification, $F(2,254) = .59$, $p = .56$, η2 $< .01$, and no interaction between these variables was found.

A hierarchical linear regression was conducted on participants bidding behaviors across all rounds of the game. The first step in the model was significant, $F(2,258) = 66.06$, $p <.01$, $R^2 = .34$, indicating that the opposing player condition significantly predicted participants' bid decisions. The second step in the model was marginally significant, $F(2,256) = 2.59$, $p = .08$, $\Delta R^2 = .02$. Finally, the third step in the model was not significant, $F(2,254) = 1.44$, $p =.24$, $\Delta R^2 < .01$. Full results can be seen in Table 3.

Finally, partial correlation analyses were conducted between participants' bidding behavior and several other continuous variables (i.e., Utilitarian and Deontological moral process strength, empathic concern, relatedness to nature, politics, handedness). Participants' assigned Opposing Player Condition were dummy coded and controlled for in these analyses. Participants' Deontological Process strength negatively correlated with total resource extraction, $r = -.14$, $p = .02$, and marginally correlated with average bid per round, $r = -.12$, $p =.06$. Participants' Utilitarian Process, however, did not correlate with participants' behavior in the game. Full results can be seen in Table 4.

Discussion. Experiment three was designed to test whether the effects found in experiments one and two (but specifically experiment one) would generalize to a situation with real-world consequences. Generally, it was predicted that experiment three would replicate the findings of experiment one. The first hypothesis was supported, as participants' moral judgments did predict differences in behavior in the foresting game

(just as with experiment one). That is, participants moral judgments (specifically their moral classification) predicted differences in their first-year bids (or their opening strategies). However, this effect seems to have flipped when comparing it to the results of experiment one. That is, in experiment one, participants classified as utilitarian descriptively made the largest opening bids (although no post hocs were significant). In contrast, in experiment three, utilitarians made the lowest opening bids (and it was significant, i.e., post hoc analyses showed that utilitarians made lower bids than deontologists, $p = .04$). Reasons for these differences will be expounded upon in the following section on inter-experiment comparisons, but for now, it seems that utilitarian participants' opening strategies were more in line with their strategies in experiment two (where the consequences, while hypothetical, were designed to be more motivating) than in experiment one.

Replicating the results of experiment one, there was not a significant effect of moral classification on behavior in the foresting game (supporting Hypothesis 1). Interestingly, the observed interaction between opposing player condition and moral classification that was found in experiment one was not replicated (therefore, failing to support Hypothesis 2). Thus, like experiment one (but to an even larger degree), while the participant's moral type did predict differences in the opening round, these differences appear to have been overshadowed by the larger effect of the opposing player conditions when considering behaviors throughout the game.

In the hierarchical linear regression, steps one (opposing player condition) and two (moral processes) were hypothesized to be significant in the model. While step one was significant, step two was only marginally significant (thus only partially supporting

Hypothesis 3). However, unlike the utilitarian process, participants' deontological

process was a significant predictor ($t = -2.55$, $p = .025$) of resource extraction. This effect

partially replicates the results of experiment one, which found a negative relationship

between participants' strength of the deontological process and their resource extraction

in the foresting game (supporting Hypothesis 4). In addition, participants' deontological

process strength showed a significant negative correlation with their resource extraction.

The marginal negative correlation observed in experiment one, between participants'

utilitarian process and bid behavior was not found. Finally, step three in the hierarchical

linear regression was not significant (supporting Hypothesis 5).

Taken together, the results of experiment three partially replicated experiment

one. Importantly, for the first time, these results show that participants' moral judgments

(even in hypothetical scenarios) predict their behavior in commons games, where real-

world monetary consequences are at stake. While they do not seem to offer a ringing

endorsement of the adherence to either deontology or utilitarian ethics as a superior

motivator to avoiding tragedies of the commons, they do inform how these judgments in

line with these theories predict behavior in a variety of situations and contexts. The larger

discussion below will address this line of research as a whole, and inter-experiment

analyses (which are the subject of the next section) will better inform that discussion.

However, for now, it is important to stress the novelty of this finding. Surely, the results

here can be extended into many different future directions.

Inter-Experiment Comparisons and Exploratory Analyses.

The prior sections in this chapter outlined the results of experiments 1-3. These

experiments investigated whether moral judgments predict behavior in game situations

that simulate collective action problems for two key reasons. First, they were designed to test the accuracy of prior philosophical arguments which have celebrated utilitarian morality (in their application) as superior to deontological ethics in specific situations, like commons dilemmas. Second, these studies sought to bring greater insights into the relationship between moral judgments and behavior in game situations that simulated the tragedy of the commons. The context and consequences of these games were manipulated between experiments 1-3. However, the basic dependent variables in these experiments were kept constant and thus, allow for several inter-experiment comparisons. These comparisons, and a few other exploratory analyses, are the subject of the present section. Results here speak to how differences in context and consequences influence the connection between judgment and behavior and may identify other potential future directions of research.

Comparison 1. The first comparison of interest is whether money had a significant effect on behavior, along with moral processing, in the foresting game. Here, it's relevant to statistically compare the results of experiments one and three.

A hierarchical linear regression was conducted on participants bidding behavior across all rounds of the game with the computer bot conditions, participants' moral processes (U- and D- processes), and whether monetary incentives were provided (entered as a dummy coded variable) were entered as predictors. The first step in the model was significant, $F(1,516) = 170.88$, $p < .01$ $R^2 = .25$, indicating that the opposing player condition significantly predicted participants' bid decisions. The second step in the model was also significant, $F(2,514) = \Delta 4.78$, $p < .01$, $\Delta R^2 = .01$, indicating that participants moral processes accounted for an additional one percent of the variance

above and beyond bid condition. Finally, the third step in the model was significant, $F(1,513) = \Delta89.22$, $p <.01$, $\Delta R^2 = .11$, indicating that the presence of monetary incentives accounted for an additional 11 percent of the variance.

Comparison 2. The second comparison of interest is whether the game scenario itself (i.e., the foresting game versus the Covid Reefs game) had a significant effect on behavior, along with moral processing. Here the relevant comparison are the data of experiments one and two.

A hierarchical linear regression was conducted on participants bidding behavior across all rounds of the game. The opposing player condition, participants' moral processes, and whether the game situation (entered as a dummy variable indicating whether they played in the foresting or reef game) were entered as predictors.

The first step in the model was significant, $F(1,504) = 202.47$, $p <.01$ $R^2 = .29$, indicating that the opposing player condition significantly predicted participants' bid decisions. The second step in the model was also significant, $F(2,502) = \Delta 4.55$, $p = .01$, $\Delta R^2 = .01$, indicating that participants moral processes accounted for an additional one percent of the variance above and beyond bid condition. Finally, the third step in the model was significant, $F(1,501) = \Delta35.82$, $p <.01$, $\Delta R^2 = .05$, indicating that differences in the game scenario accounted for an additional five percent of the variance.

In previous sections, the results were reported (and plotted) from ANOVAs that evaluated the average and total amount of resources that participants harvested from the common resource pool. As such, the total number of harvested resources were compared between computer-bid strategy and moral type factors (see Figures 1-3). While these analyses were critical to the current work, a more detailed picture of participants'

behaviors (between these variables at every round of the game) is also important to any comprehensive viewing of the results of this line of research. Thus, participants' average resource harvesting, in each round of the games, between computer bid strategy type and their moral classification type, are plotted for each experiment in Figures 4-6.

Finally, to better assess how participants' moral processing influenced their decisions across all three experiments, participants' behaviors were analyzed in a Repeated Measures ANOVA. Rounds in the game was entered as a within subjects variable, while game scenarios (hypothetical profit, real profit, lives) and computer bid strategy condition (aggressive, equilibrium, soft) were entered as between subjects variables. Continuous variables of participants moral processing (U - and D - process) were entered as covariates. A main effect of behavior between rounds of the game was found, $F(11,8338) = 19.30$, $p < .01$, $\eta^2 = .03$. A number of two-way interactions were also found between rounds of the game and other variables. These include game rounds interacting with the game scenario, $F(22,8338) = 4.03$, $p <.01$, $\eta^2 = .01$, the computer player bid condition, $F(22,8338) = 174.36$, $p <.01$, $\eta^2 = .32$, and the U-process $F(11,8338) = 1.98$, $p = .03$, $\eta^2 = .01$. Finally, a three-way interaction was found between game round, game scenario, and computer player bid condition, $F(44, 8338) = 3.78$, $p < .01$, $\eta^2 = .02$. Importantly, between-subjects tests showed that the U-process, $F(1,758) = 4.69$, $p = .03$, $\eta^2 = .01$ and D-process, $F(1,758) = 9.75$, $p = .01$, $\eta^2 = .01$, were significant covariates in the model.

Exploratory Analysis. All data for experiments 1-3 were collected between October 2020 and May 2021 (during the Covid-19 pandemic). Thus, in all experiments, participants responded to two questions to assess their perception of the seriousness of

Covid-19 and their opinion about a governmental vaccine mandate. The first question, "Please rate on the scale below how serious you believe the current Covid-19 global pandemic is?", was assessed on a 100-point scale, ranging from 0 (*Not very serious*), 50 (*Somewhat serious*), to 100 (*Very serious*). The second question, "If a vaccine to Covid-19 were to be created and become publicly available, do you think the government should require that all adult citizens take the vaccine?", was assessed on a 7-point scale, ranging from 1 (*Not at all*) to 7 (*Very Much).*

For purposes of the current project, it is an interesting and relevant question as to whether moral judgments are associated with different opinions regarding Covid-19 (both in terms of the seriousness of the pandemic and favorability with a government-imposed vaccine mandate). Thus, a hierarchical linear regression was performed on these data to investigate this question. Participants' seriousness ratings of Covid-19 were entered into the first step of the model. In step 2, participants' moral processes (U and D) were entered, along with their rates of political conservatism. The first step of the model was unsurprisingly significant, $F(1,764) = 161.37$, $p < .01$, $\Delta R^2 = .17$. The second step of the model also was significant, $F(3,761) = \Delta 23.24$, $p < .01$, $\Delta R^2 = .07$, indicating that participants moral processes and conservatism accounted for 7 percent of the variance above and beyond their perceived seriousness of the pandemic. While participants' conservatism ($p < .01$) and deontological process strength ($p < .01$) were both negatively associated with support for a government-imposed vaccine mandate, their U process strength showed no relationship. Correlation analyses showed no relationship between moral processes and the perceived seriousness of the pandemic.

These results importantly show that differences in moral processing are predictive of how people view policies (like mandatory vaccination) related to real-life collective action problems (like vaccine adherence) and have been recently published (Clarkson & Jasper, 2022). What's more, these results also are only partially in line with theoretical predictions. Interestingly, participants' U-process strength did not significantly correlate with or predict support for the mandate. This is somewhat surprising, as it seems likely that a mandated vaccine would lead to better consequences. Research by Kahane et al., (2018), which indicates that utilitarianism is not unidimensional, but rather, motivated by two distinct factors (i.e., impartial beneficence and instrumental harm), might offer an explanation. That is, it could be that while the utilitarian process (as a whole) is not associated with opinions about a vaccine mandate, specific dimensions (i.e.., impartial beneficence) may be. Since this 2-dimensional account has relevance to the results of this line of research as a whole, it will be discussed at some length in the next chapter.

Chapter Six

Conclusions, Future Directions, and Implications on Moral Judgment's Association

with Behavior in Situations Simulating the Tragedy of the Commons

Morality is a difficult thing to define, much less study. Philosophers have long questioned the basis of morality and have constructed theories that offer us recommendations about how to behave morally and lead an ethical life. For several decades now, scientists (and some philosophers) have been doing yeoman's work in outlining our moral judgments, beliefs, and ethical intuitions. As was the topic of previous chapters, this work has yielded significant knowledge on these fronts. It is clear that morality, in practice, is not governed just by a rational or emotional process, but instead, a combination of the two. Thus, it seems that rational and emotional beliefs instantiate our ethical intuitions and produce judgments about what it means to behave morally; in a given situation. Indeed, behavior is key, not only for our immediate purposes here but for the field at large. The behavior is there in the world. It is observed daily by all who live in it. However, when it comes to a laboratory-based empirical study, behavior is difficult to target, as our real-world moral sensibilities (that would decry any attempt at a real-life trolley problem being used in controlled research settings) keep us from measuring moral behavior of real consequence in these settings.

Taking the most famous example from the field, we do not actually require a person to decide to pull the switch or not. That is, we do not require a person in actuality to kill another to realize a greater good. Instead, we ask for judgments or beliefs related to such scenarios, with the notion that these judgments can stand in for what people might actually do in such situations. Yet, this practice has several flaws, which were described

70

earlier, and some have doubted whether the collected data on moral judgment can speak to moral behavior at all. Indeed, such critiques have motivated some researchers to investigate morality in other ways that get closer to studying behavior rather than just judgment. Thus, the primary aim of the current project was to redress this common shortcoming and to answer a key question – do moral judgments predict moral behavior?

One popular method in this quest has been to study the relationship between moral judgments and behavior in game situations. In these studies, people have provided judgments in moral dilemmas before playing games (like the trust game). Interesting but sometimes conflicting conclusions have been drawn from this research. However, it alone cannot answer certain outstanding debates within the field – like whether utilitarians are better suited in practice at avoiding social traps and collective action problems. On this point, philosophers (like Greene) have famously argued that utilitarian ethics is better equipped to solve these problems than other theories, like deontology. In addition, these philosophers have argued that an increased following of utilitarian ethics should reduce our susceptibility to falling into social traps that result in tragedies for the commons.

Of course, in one sense, this is an explicitly philosophical question, as it asks about the expected consequences of following one moral theory over others. In situations like those outlined by Hardin's tragedy of the commons, it is prima facie obvious that if individuals follow the prescription of utilitarian ethics, they should better navigate these dilemmas (this rationale is outlined well in Sunstein, 2005). Yet, another related question, deeply entrenched within the realm of empirical study, remains. That is, do individuals that demonstrate utilitarian judgments actually behave in accordance with utilitarian prescriptions in these situations? If not, do moral judgments predict behavior in these

scenarios (those that simulate tragedies of the commons) at all? To date, these questions have no (or very little) record of study in the literature.

Across three experiments, the current project collected data to answer these questions. One, do moral judgments predict behavior? Two, do utilitarian judgments predict utilitarian behavior in situations that simulate the tragedy of the commons? Three, do judgments aligned with other moral theories (like deontology) predict behavior in these situations? In respect to these questions, many hypotheses were suggested and the status of each was diagnosed in the last chapter. Instead of further investigating these points in detail, it is here that we will speak to the overarching conclusions that can be drawn from this work. Then, after taking away from it what we can, several important future directions will be discussed.

General Conclusions.

Whether moral judgments predict behavior, indeed, is a foundational question in moral psychology. While some studies have been conducted on this front, clear answers about the association between moral judgments and behavior are difficult to decipher. The current research does not conclusively answer the question, but it does offer a number of insights. Across all three experiments, evidence shows that moral judgment does predict behavior. While this is not new, it does add to a limited and muddled body of research and stands to improve the field's knowledge about such associations.

A more precise question was asked about utilitarian judgment and whether it predicts utilitarian behavior in situations like the tragedy of the commons. To this question, the data offer a less decisive answer. Only in experiments one and two was the strength of a participant's utilitarian process negatively associated with (and predictive

of) their resource extraction in a commons game. At the same time, however, the strength of participants' deontological process was associated with (and predictive of) less resource extraction in experiments one and three. Thus, it seems that increased moral processing (along either utilitarian or deontological lines) predicts less resource extraction (and ipsofacto, less risk of causing a tragedy of the commons) in general.

It is interesting that only in experiment one, did an increased strength of both moral processes predict reduced bidding behavior. Recall, that experiment one was the lowest stakes condition (i.e., participants played as CEOs harvesting fake timber for fake profits). This result indicates that perhaps increased moral processing, not just increased utilitarian processing, might lead to better outcomes in real-world situations where collective tragedy is a risk. However, it should be stated that the participants who were classified as utilitarian or deontological (via the moral classification technique) did not show differences (in terms of main effects) in their behavior in experiment one, except for their opening bids in round one, in which descriptively, utilitarian participants made the largest bids.

A significant interaction, however, was found between moral classification and the opposing player strategy condition. The interaction seems to have been driven by deontologically classified participants making lower bids in the equilibrium condition (a condition that is meant to simulate fair rational cooperators). Thus, this result indicates that deontologists might be particularly successful at avoiding tragedies of the commons when dealing with other fair or rational actors in these situations. Even still, it should be pointed out that, in the equilibrium condition, deontologically classified participants made bids below equilibrium. That is, these participants left money (or hypothetical

profit) on the table. In other words, deontologically classified participants (from a rational perspective) made poor bids and harvested less than the competition, even when the competition behaved fairly and rationally. This result does not appear to have been driven by differences in their desire to profit (or to save lives) or even apprehension about the game; across all three experiments, these variables did not differ between moral responders.

It is interesting that in experiment two, where participants were asked to act as heads of state (in a hypothetical scenario) where they could extract resources from a reef to save their own citizens from dying of Covid-19 in the short-term (against the risk of long-term concerns about causing a tragedy of the commons) only the utilitarian process was negatively associated with (and predicted) resource extraction. This result is consistent with analyses related to the moral classification approach, where deontologically classified participants harvested more resources than typically classified participants and (descriptively) more than utilitarian classified participants. These results indicate that in particularly grievous situations, where lives are on the line, utilitarians may be better suited to avoid tragedies of the commons than those who adhere to other ethical theories (as Greene and others have argued). This was expected.

Notice, that experiment two has some key similarities with typical moral dilemmas that experiments one and three do not. That is, in experiment two, participants must decide whether to sacrifice some people (in the short-term) to save a greater number of people (in the long-term). The only differences between the scenario in experiment two and other typical sacrificial dilemmas are the scale (i.e, it is very big), their leadership role (they are heads of state), and the time horizon (deaths from under-

harvesting are immediate while deaths from over-harvesting are delayed). Still, this is no doubt informative. It is perhaps the case that utilitarian ethics is especially important for navigating commons dilemmas when desire to profit is not motivating the conflict (or incentivizing the escalation towards collective tragedy). Profit does seem to be a key variable, especially when considering the results of experiment three, where MTurk participants played in the foresting game for real-world money.

In experiment three, where participants received a cash bonus for each hectare of the forest they harvested, deontological processing was negatively associated (and predictive of) less resource extraction. Despite this result, deontologically classified participants began the game with the largest opening bids. Thus, it seems that deontologists, while they may have started with large bids, were willing to cooperate with the players and reduced their bids over time, resulting in behavior that was less likely to lead to collective tragedy. Since there was no relationship between utilitarian processing and resource extraction in experiment three, it is difficult to say conclusively that increased adherence (or demonstration of utilitarian judgment) is predictive of behaviors better suited to avoid tragedies of the commons. Instead, it seems that it depends. What it depends on might include what is at risk if the tragedy of the commons is realized (i.e., lives or profit). Still, it is notable that the only experiment that did not show a negative relationship between utilitarian processing and resource extracting behavior was the one that connected behaviors in the game to real-world consequences.

So, what about the third question? That is, do judgments aligned with other moral theories (like deontology) predict behavior in these situations? Results from experiments one and three indicate that stronger deontological processing is predictive of reduced

resource extraction in the foresting game regardless of incentives. The fact that only in experiment two (where participants played the reefs game to save hypothetical lives) was deontological processing not related to reduced resource extraction is informative. It indicates that deontological processing may play a selectively important role in mitigating behavior that leads to tragedies of the commons when that behavior is motivated by a desire to profit. Thus, in these situations, where the desire to profit and monetary incentives are at their greatest, deontological processing may be particularly important to divert tragedies of the commons.

Taken together, the results from experiments one through three strongly indicate that increased moral processing (of either a deontological or utilitarian nature) generally predicts behavior that is less likely to lead to tragedies of the commons (i.e., reduced resource extraction from the community pool). However, this general conclusion is complicated by the observation that while utilitarian processing seems to be selectively important for diverting tragedies of the commons when lives must be sacrificed in the short-term to save a greater number in the long term (like disaster triage situations), deontological processing appears to be selectively important for avoiding such tragedies when large profits are on the line. Notice that many commons situations can take both shapes (in terms of disaster triage versus profit situations).

Limitations and Future Directions. While this research stands to make important contributions to the field of moral psychology, specific limitations and ideas for future research will be discussed in the current section.

The measurement of Utilitarianism. To measure utilitarian judgment, the current study recorded participants' responses to conventional sacrificial dilemmas, where one

person must be sacrificed for a greater good. The strength of participants' utilitarian processes also hinged on a similar factor, that is, whether participants thought it appropriate to sacrifice a person for a given result. While this second technique is advantaged by the fact that not all sacrificial harms were utilitarian (i.e., not all sacrificial actions led to a greater good) it is the case that all utilitarian judgments required participants to endorse sacrificing one person for a greater good. Recently, researchers (i.e., Kahane et al., 2018) have argued that these measurements of utilitarian processing may (at least partially) miss the mark, in that, they ignore half of the core features of utilitarianism (i.e., altruism). Recall that utilitarianism exalts the greatest good for the greatest number as its metric for moral goodness. Obviously, in most situations, this does not require that a person commits a harmful action (e.g., murder) to bring about a greater good. Instead, utilitarianism also, and perhaps more commonly, requires positive acts of altruism (Singer, 1972). Thus, according to Kahane et al., (2018), studies exclusively relying on judgments in sacrificial dilemmas to measure utilitarianism offer an incomplete picture of a participants' actual utilitarian motivations and beliefs.

In recognition of this, some researchers (e.g., Capraro et al., 2019; Kahane et al., 2018) have taken to measuring one's support for utilitarian ethics across two dimensions, that is, instrumental harm (e.g., murder) and impartial beneficence (e.g., donating to charity). Known as a two-dimensional approach, research has found that much of the negative personality traits correlating with sacrificial utilitarian judgment are associated only with the instrumental harm dimension of utilitarianism (Kahane et al., 2018). Conversely, positive personality traits (like empathic concern) are associated with the impartial beneficence dimension (Kahane et al., 2018).

Considering these findings, recent investigators have called for additional studies that compare one's adherence to these two dimensions of utilitarianism with their real-life moral behaviors (e.g., Capraro et al., 2019). Thus, to address the questions in the present work, future research should measure whether adherence to these different dimensions of utilitarianism predicts different patterns of behavior in commons dilemmas. Such investigations would add to the findings of the current research, which show that utilitarian processing (when it is predictive of behavior) predicts lower levels of resource extraction in commons dilemmas. Indeed, new research using the 2-dimensional model may find that particular dimensions (like impartial beneficence) specifically predict behavior that is likely to lead to avoid tragedies of the commons.

The measurement of just Utilitarianism and Deontology. In its focus on morality, the current project only measured participants' moral judgments and processes concerning philosophically based theories like utilitarianism and deontology. However, building upon these philosophical theories and other interdisciplinary work, psychologists have extensively investigated morality across thousands of research studies (Ellemers et al., 2019). In addition to the dual-process model of moral psychology (which was discussed in chapter two), psychologists have also posed alternative theories (i.e., morality-as-cooperation-theory and moral foundations theory) that rely less on theories from moral philosophy in favor of data that has been collected on peoples' moral intuitions, beliefs, and judgments.

For instance, according to a recently proposed theory, known as morality as cooperation theory, morality exists as a kind of finely tuned operating system, rooted in biology and culture, that promotes beliefs and behaviors suited to solving cooperation

problems humans commonly face in social life (Curry et al., 2019). Since these problems are socially ubiquitous, the theory identifies seven categories of cooperative behavior that should be culturally universal: helping kin, helping your group, reciprocity, bravery, deferring to superiors, dividing resources under dispute, respecting prior possession (Curry et al., 2019). Indeed, in a recent study of 60 cultures around the world, all identified these seven behaviors as morally good (Curry et al., 2019). To morality as cooperation theory, our need to cooperate drives our moral sensibilities. This contrasts with another popular theory, moral foundations theory.

While moral foundations theory also emphasizes the importance of evolution and culture in engineering our morality (Haidt, 2012), our need to cooperate is not its sole driver. It holds moral judgment to be primarily intuitive, non-rational, and like morality as cooperation theory, it divides morality into different categories of moral beliefs and behaviors: harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, purity/sanctity (Haidt, 2012). While it has been widely studied (for a review, see Simpson, 2017), the theory has been criticized by supporters of morality as cooperation theory for including certain categories (i.e., care and purity) that do not directly facilitate cooperation as moral bedrock (Curry et al., 2019). Thus, the objection goes that those moral precepts, principles, or categories that do not directly offer solutions to cooperation problems are simply culturally specific sentiments, and insofar as they do not generalize to other cultures, are not truly universal moral foundations (Curry et al., 2019). Still, researchers have collected extensive evidence that ones' belief in these foundations (e.g., fairness) is a useful individual difference variable for predicting moral judgment and behavior.

Taken together, these two moral theories in psychology represent an interesting future avenue of research. Due to necessity, and the direct connections made between utilitarianism and diverting tragedies of the commons (see Greene, 2013) the current study prioritized testing the connection between participants' adherence to (or judgment aligned with) particular theories in moral philosophy and their behavior in commons dilemmas. This research should be seen as an initial, albeit important first step toward determining the relationship between morality and behavior in these situations.

Real-world examples. Future research would do well to use methods that measure specific dimensions of utilitarianism and other ethical theories (like Moral Foundations Theory or Morality as Cooperation Theory) when further investigating the observed link between morality and behavior in commons dilemmas. Indeed, such issues are important, as commons dilemmas are increasingly frequent, with ever-growing stakes in modern globalized societies (e.g., autonomous vehicles, climate change, viral pandemics). For instance, do people want cars to make the utilitarian choice or the deontological choice in similar real-life scenarios on the road? Like the trolley problem, results from several studies indicate that individuals are not consistent in their response (Shariff et al., 2017) and that cultural differences also exist (Awad et al., 2018).

This situation's relevance to collective action problems quickly becomes clear when considering evidence showing that participants generally prefer utilitarian programmed self-driving cars, unless it is their car, in which case these preferences are reversed (Bonnefon et al., 2016). That is, while people favor utilitarian cars, that act to bring about the best consequences, they would rather their car be programed with a deontological principle to save the life of the driver above all else, despite the

consequences to others. Thus, debates surrounding the ethical programming of self-driving cars represent a pressing social dilemma, where despite peoples' general preferences for utilitarian vehicles, people may purchase deontological cars, which would likely lead to worse long-term outcomes for everyone (Bonnefon et al., 2016).

Moral views surrounding another contemporary issue (i.e., climate change) have also been the subject of recent investigations. In this vein, several research studies have sought to associate moral judgment with opinions about behaviors thought to mitigate climate change (e.g., Dickinson et al., 2016; Kahn & Lourenco, 2002; Knez, 2016; Markowitz & Shariff, 2012; Thogersen, 1996). While these studies have measured the association between moral judgment and views about climate change (not behavior), they did yield several interesting results. First, several studies found support for the existence of a "deontic proposition" or moral obligation in favor of behaving pro-environmentally (Kahn & Lourenco, 2002; Knez, 2016).

Recall that Greene's argument (about utilitarianism and the tragedy of the commons) largely only compared utilitarian and deontological ethics. Thus, considerably less focus was given to other moral theories like moral foundations theory or morality as cooperation theory. While (in its application to how moral judgment predicts behavior) Greene's argument may not be valid (based on the findings of the current project), it is an unanswered question as to how support for specific moral precepts (like equaling dividing resources under dispute) outlined by other moral theories in psychology associate with behavior.

In light of this question, a study by Dickinson et al., (2016) offers some initial data. While it still did not measure behavior, the study assessed the association between

support for precepts in moral foundations theory and one's willingness to take personal action to combat climate change in the United States. Most notably, it was found that strong valuations of fairness supported willingness to act, while purity was a marginal predictor (Dickinson et al., 2016). However, several studies have found that moral concerns may not be sufficient to motivate pro-environmental behaviors (e.g., Markowitz & Shariff, 2012; Thogersen, 1996).

Considering these findings, researchers have pointed out a need for increased experimental studies to investigate the relationship between moral judgment and pro-environmental behavior (e.g., Dickinson et al., 2016). Despite the relevance of this research and its promise as a more ecologically valid test of morality, it still relies mostly on participants' judgments rather than their moral behaviors. Thus, considering the findings of the current research project, these domains represent a promising avenue for future study that measures participants' behavior in game situations that simulate another commons dilemmas like climate change.

Strategies and tactics. Across all three experiments in the current research project, participants were presented with three types of opponent player strategies that were broadly characterized as soft, equilibrium, or aggressive based on the preprogrammed behavior of bots in the commons games. Of course, future research could build upon these methods in several respects. First, research could program more adaptive bots that adjust in real-time to the behavior of participants (such as tit-for-tat programs). Second, research should also use designs that require participants to play these games against other human players.

Due to this project's commencement during the peak of the Covid-19 pandemic, in-person data collection was not possible, and as such, was not performed in this research. It is no doubt critical to use methods in the future that require participants to play commons games with others. Along these lines, studies could even manipulate how different types of moral responders interact with each other. For instance, such research could investigate if certain compositions of moral responders (such as a group of people high in utilitarian processing) are more likely to have certain outcomes (like a tragedy for the commons) than other groups. Thus, it could be that morality has effects, not just at the individual, but also at the group level on behavior.

Deeper philosophical truths remain obscure. This paper does not attempt to speak to deeper philosophical truths about the nature of morality. For instance, just because utilitarian processing did not predict reduced resource extraction in experiment three, this cannot speak to the nature of utilitarian theory itself. Instead, this result only shows that individuals who make utilitarian judgments in sacrificial dilemmas demonstrated behavior less likely to live up to the ethical requirements of the most basic models of utilitarianism in this specific situation. While philosophers have outlined many different versions of utilitarian (and deontological) ethics, psychological studies typically only investigate the degree to which people follow the general tenets of these basic or sometimes, as it has been suggested, cartoon versions of these theories.

The current work cannot remedy this issue but instead can only note that some philosophers may correctly point out that future research could be benefited by investigating whether support for specific types of utilitarianism (like act versus rule utilitarianism) predict different patterns of behavior. One issue here, however, is that the

typical folk morality of the common man may not be able to appreciate the sometimes-subtle distinctions between precise theories such as these. Thus, such studies may be difficult to come by beyond those studies that recruit those with expertise in moral philosophy. While this would no doubt be interesting, it would be of less practical value to the research aims that motivated this project.

Final Remarks. Past research has called for increased study of human behavior in morally charged situations. A key reason for such declarations is twofold. First, most research in moral psychology has only collected data on moral judgment, which carries with it a host of weaknesses, that studies of moral behavior can avoid. Second, the studies that directly investigate moral behavior may be better equipped to answer a major debate in the field, i.e., whether utilitarian responders are motivated to instantiate a greater good. This research serves as an important first step in determining the relationship between morality and behavior in collective action problems and commons dilemmas. Across all three experiments, participants' moral judgments or processes were associated with (and often predicted) behavior in situations simulating the tragedy of the commons.

Specifically, increased strength of either moral process (i.e., utilitarian or deontological) was negatively associated with participants resource extraction from the community pool (behavior which makes tragedies of the commons less likely). Yet, increased deontological processing also predicted less resource extraction behavior in experiments when profit was the explicit motive. Thus, while this research cannot speak to Greene's philosophical argument (about whether a true commitment to utilitarian ethics will offer societies the best chance to avoid tragedies of the commons), it does have implications about a related empirical idea. That is, this research does not show that

those who make utilitarian decisions (or demonstrate increased utilitarian processing) are more likely to demonstrate behavior better suited to avoid tragedies of the commons in anything approaching a universal sense. Thus, future research should further investigate how belief in different moral precepts (outlined by particular theories in moral psychology) predict behavior in association with adherence to consequentialist or non-consequentialist ethics (outlined by moral philosophy), across different commons dilemmas (like climate change games), and against more sophisticated opponents (like other real human participants) in game situations.

References

Alfano, M., Rusch, H., & Uhl, M. (2018). Ethics, morality, and game theory. *Games,*
9(2).

Ariely, D. (2009). *Predictably irrational: The hidden forces that shape our decisions.*
Harper Collins publishers.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., … Rahwan, I.
(2018). The moral machine experiment. *Nature. 563,* 59-64.

Baatz, Christian., & Voget-Kleschin, Lieske. (2019). Individuals' contributions to
harmful climate change: The fair share argument restated. *Journal of Agricultural*
and Environmental Ethics, *32*(4), 569-590.

Baron, J., & Ritov, I. (2009). Protected values and omission bias as deontological
judgments. *The psychology of learning and motivation: Vol. 50. Moral judgment*
and decision making, 133–167

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial
personality traits predict utilitarian responses to moral
dilemmas. *Cognition*, *121*(1), 154–161.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality:
The evolution of fairness by partner choice. *Behavioral & Brain Sciences*, *36*(1),
59–78.

Bloom, P. (2011). Family, community, trolley problems, and the crisis in moral
psychology. *Yale University, 99*(2), 26-43.

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous
vehicles. *Science, 352*(6293), 1573–1576.

Bostyn, D., & Roets, A. (2017a). An asymmetric moral conformity effect: Subjects

conform to deontological but not consequentialist majorities. *Social

Psychological and Personality Science*, *8*(3), 323-330.

Bostyn, D., & Roets A. (2017b). Trust, trolleys and social dilemmas: A replication

study. *Journal of Experimental Psychology General*. *146*(5):e1-e7.

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys:

Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas.

Psychological Science, 29(7), 1084–1093.

Bretz, S., & Sun, R. (2018). Two Models of Moral Judgment. *Cognitive Science*, *42*, 4–

37.

Brodbeck, F., Kugler, K., Reif, J., & Maier, M. (2013). Morals matter in economic

games. *Plos One 8*(12).

Camerer, C. (2003). *Behavioral game theory*: *Experiments in strategic interaction.*

Princeton University Press.

Camerer, C., & Thaler, R. (1995). Anomalies: Ultimatums, dictators and manners. *The

Journal of Economic Perspectives, 9*(2), 209-219.

Capraro, V., Everett, J. A. C., & Earp, B. D. (2019). Priming intuition disfavors

instrumental harm but not impartial beneficence. *Journal of Experimental Social

Psychology, 83, 142–149.*

Capraro, V., Jordan, J., & Rand, D. (2014). Heuristics guide the implementation of social

preferences in one-shot prisoner's dilemma experiments. *Scientific Reports, 4*

(6790).

Capraro, V., Perc., M. (2021). Mathematical foundations of moral preferences. *Journal of the Royal Society Interface, 18*(175).

Capraro, V., Sippel, J., Zhao, B., Hornischer, L., Savary, M., Terzopoulou, Z., Faucher, P., & Griffioen, S. F. (2018). People making deontological judgments in the trapdoor dilemma are perceived to be more prosocial in economic games than they actually are. *PLoS One*, *13*(10), 1–16.

Clark, C. B., Swails, J. A., Pontinen, H. M., Bowerman, S. E., Kriz, K. A., & Hendricks, P. S. (2017). A behavioral economic assessment of individualizing versus binding moral foundations. *Personality and Individual Differences, 112, 49–54.*

Clarkson, E. (2022). Moral behavior in games: A review and call for additional research. *New ideas in psychology,* 64.

Clarkson, E., & Jasper, J. (2022). Individual differences in moral judgment predict attitudes towards mandatory vaccinations. *Journal of personality and individual differences*, 186.

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1225.

Chen, D. (2011). Markets and morality: How does competition affect moral judgment. *SSRN Electronic Journal.*

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality & Social Psychology*, *104*(2), 216–234.

Conway, P., Goldstein-Greenwood, J., Polacek, D., & Greene, J. D. (2018). Sacrificial

utilitarian judgments do reflect concern for the greater good: Clarification via

process dissociation and the judgments of philosophers. *Cognition*, *179*, 241–265.

Crittenden, V., Hanna, R., & Peterson, R. (2009). The cheating culture: A global societal

phenomenon. *Business Horizons,* 52, 337-346.

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing

the theory of morality-as-cooperation in 60 societies. *Current Anthropology,*

60(1).

Darr, R. (2019). Climate change, individual obligations and the virtue of justice. *Studies*

in Christian Ethics, *32*(3), 326-340.

Dickinson, D. L., & Masclet, D. (2019). Using ethical dilemmas to predict antisocial

choices with real payoff consequences: An experimental study. *Journal of*

Economic Behavior & Organization, 166, 195–215.

Dickinson, L., McLeod, P., Bloomfield, R., & Allred, S. (2016) Which moral foundations

predict willingness to make lifestyle changes to avert climate change in the USA?

PLoS ONE 11(10):e0163852.

Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The psychology

of morality: A review and analysis of empirical studies published from 1940

through 2017. *Personality and Social Psychology Review*, *23*(4), 332–366.

Ellemers, N., Pagliaro, S., & Barreto, M. (2013). Morality and behavioral regulation in

groups: A social identity approach. *European Review of Social Psychology*, *24*(1),

160–193.

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being

consequentialist: Social inference from instrumental harm and impartial

beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216.

Everett, J., & Kahane, G. (2020). Switching tracks? Towards a multidimensional model

of utilitarian psychology. *Trends in Cognitive Science*.

Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness

from intuitive moral judgments. *Journal of Experimental Psychology.*

General, *146*(6), 772–787.

Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., &

Terbeck, S. (2016). Virtual Morality: Transitioning from Moral Judgment to

Moral Action? *PLoS ONE*, *11*(10), 1–22.

Friesdorf, R., Conway, P., & Gawronski, B. (2015). Gender differences in responses to

moral dilemmas: A process dissociation analysis. *Personality and Social*

Psychology Bulletin, *41*(5), 696-713.

Gold, N., & Colman, A. (2020). Team reasoning and the rational choice of payoff-

dominant outcomes in games. *Topoi,* 39, 305-316.

Gold, N., & Colman, A., Pulford, B. (2014). Cultural differences in responses to real-life

and hypothetical trolley problems. *Judgment and Decision Making. 9*(1).

Gold, N., Pulford, B. D., & Colman, A. M. (2015). Do as I say, don't do as I do:

Differences in moral judgments do not translate into differences in decisions in

real-life trolley problems. *Journal of Economic Psychology, 47, 50–61.*

Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong

(Ed.), *Moral psychology, Vol. 3. The neuroscience of morality: Emotion, brain*

disorders, and development (p. 35–80). MIT Press.

Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them.*

Penguin Books.

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.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001).

An fMRI investigation of emotional engagement in moral

judgment. *Science*, *293*(5537), 2105–2108.

Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian

moral judgment. *PLoS One*, *8*(4), 1–9.

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.

Grossmann, I., Eibach, R. P., Koyama, J., & Sahi, Q. B. (2020). Folk standards of sound

judgment: Rationality versus reasonableness. *Science Advances*, *6*(2), 1–14.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to

moral judgment. *Psychological Review, 108*(4), 814–834.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and*

religion. New York: Pantheon Books.

Hardin, G. (1968). The tragedy of the commons. *Science, 162*(3859), 1243-1248.

Harris, S. (2010). *The moral landscape: How science can determine human values.* Free Press.

Horne, Z., & Powell, D. (2016). How large is the role of emotion in judgments of moral dilemmas? *PloS One, 11*(7), Article e0154780.

Haesevoets, T., Van Hiel, A., & Reinders, C. (2015). The underlying motives of different Mixed-motive games. *European Journal of Personality*, *29*(6), 580–584.

Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8658–8663.

Julmi, C. (2019). When rational decision-making becomes irrational: A critical assessment and re-conceptualization of intuition effectiveness. *Business Research*, *12*(1), 291–314.

Kahane, G., Everett, J., Earp, B., Farias, M., & Savulescu, J. (2015). 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition, 134*, 193-209.

Kahane, G., Everett, J., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131–164.

Kahn, P. H., & Lourenço, O. (2002). Water, Air, Fire, and Earth: A Developmental Study in Portugal of Environmental Moral Reasoning. *Environment and Behavior*, *34*(4), 405–430.

Kim, E., Iyer, R., Graham, J., Chang, Y.-H., & Maheswaran, R. (2013). Moral values from simple game play. *Lecture Notes in Computer Science, 56–64.*

Knez, I. (2016). Is climate change a moral issue? Effects of egoism and altruism on pro-environmental behavior. Current Urban Studies, 4(2), 157–174.

Korsgaard, M. (1996). *Kant: Groundwork of the metaphysics of morals*. Cambridge University Press.

Kreps, A., & Monin', B. (2014). Core values versus common sense: Consequentialist views appear less rooted in morality. *Personality and Social Psychology Bulletin*, *40*(11), 1529-1542.

Kundu, P., & Cummins, D. (2013). Morality and conformity: The Asch paradigm applied to moral decisions. *Social Influence, 8*(4), 268–279.

Hallsson, B. G., Siebner, H. R., & Hulme, O. J. (2018). Fairness, fast and slow: A review of dual process models of fairness. *Neuroscience & Biobehavioral Reviews, 89, 49–60.*

Markowitz, E. M., & Shariff, A. F. (2012). Climate change and moral judgment. *Nature Climate Change, 2(4),* 243–247.

Ning Ma, Nan Li, Xiao-Song He, De-Lin Sun, Xiaochu Zhang, & Da-Ren Zhang. (2012). Rejection of unfair offers can be driven by negative emotions, evidence from modified ultimatum games with anonymity. *PLoS One*, *7*(6), 1–6.

Noclain, M. (2018). In deontologists, we trust: The effect of gender and harm-type on inferences of trustworthiness based on other's moral judgments *Ghent University.*

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, *289*(5485), 1773–1775.

Patil, I., & Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in psychology*, *5*, 501.

Plunkett, D., & Greene, J. D. (2019). Overlooked evidence and a misunderstanding of what trolley dilemmas do best: Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, *30*(9), 1389–1391.

Pölzler T. (2017). Revisiting folk moral realism. *Review of philosophy and psychology*, *8*(2), 455–476.

Prentice, M., & Sheldon, K. M. (2015). Priming effects on cooperative behavior in social dilemmas: considering the prime and the person. *The Journal of social psychology*, *155*(2), 163–181.

Ramsey, W. (2019). Intuitions as evidence facilitators. *Metaphilosophy*, *50*(1-2), 76-99.

Randall, F. (1964). *Leviathan*. Washington Square Press.

Reuben, E. & Van Winden, F. (2005). Negative reciprocity and the interaction of emotions and fairness norms. *CEsifo Working Paper Series, 168.*

Rom, S. C., Weiss, A., & Conway, P. (2017). Judging those who judge: Perceivers infer the roles of affect and cognition underpinning others' moral dilemma responses. *Journal of Experimental Social Psychology*, *69*, 44–58.

Ross, D. Game theory, *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/game-theory/>.

Sacco, D., Brown, M., & Lustgraaf, C., & Hugenberg, K. (2017). The adaptive utility of deontology: Deontological moral decision-making fosters perceptions of trust and likeability. *Evolutionary Psychological Science*, *3*, 125-132.

Saltzstein, H. D. (1994). The relation between moral judgment and behavior: A social-cognitive and decision-making analysis. *Human Development*, 37(5), 299–312.

Schneewind, B. (1983). *An enquiry concerning the principles of morals*. Hackett
Publishing Company.

Shariff, A., Bonnefon, JF. & Rahwan, I. (2017). Psychological roadblocks to the adoption
of self-driving vehicles. *Nature Human Behavior 1*(10)**,** 694–696.

Sheldon, K. M., & McGregor, H. A. (2000). Extrinsic value orientation and "the tragedy
of the commons." *Journal of Personality*, *68*(2), 383–411.

Simpson, A. (2017). Moral foundations theory. In *Encyclopedia of Personality and
Individual Differences* (pp. 1–11).

Singer, P. (1972). Famine, affluence, and morality. *Philosophy & Public
Affairs, 1,* 229–243.

Sinnott-Armstrong, W., Young, L., & Cushman, F. (2010). Moral intuitions. In J. M.
Doris (Ed.) & Moral Psychology Research Group, *The moral psychology
handbook* (pp. 246–272). Oxford University Press.

Sunstein, C. (2005). Moral heuristics. *Behavioral and Brain Sciences.*

Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between
Judgment and Choice of Action in Moral Dilemmas. *Frontiers in psychology*, *4*,
250.

Thøgersen, J. (1996). Recycling and morality: A critical review of the
literature. *Environment and Behavior*, *28*(4), 536–558.

Troyer, J. (2003). *The classical utilitarians: Bentham and mill.* Hackett Publishing
Company.

Van Lange, P., Joireman, J., Parks, C., Dijk, E. (2013). The psychology of social

    dilemmas: A review. *Organizational Behavior and Human Decision Process*,

    120, 125-141.

Von der Pfordten, D. (2012). Five elements of normative ethics - A general theory of

    normative individualism. *Ethic Theory Moral Prac*, 15, 449–471.

Weaver, A. J., & Lewis, N. (2012). Mirrored morality: An exploration of moral choice in

    video games. *Cyberpsychology, behavior, and social networking*, *15*(11), 610–

    614.

Yamagishi, T., Horita, V., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009).

    The private rejection of unfair offers and emotional commitment. *Proceedings of*

    the National Academy of Sciences of the United States of America, *106*(28),

    11520–11523.

Zheng, Y., Cheng, X., Xu, J., Zheng, L., Li, L., Yang, G., & Guo, X. (2017). Proposers'

    economic status affects behavioral and neural responses to unfairness. *Frontiers*

    in psychology, 8, 847.

    uidelines with regards to font style, size, color, line spacing, etc.>

Table 1

Descriptive Statistics for Subject Variables and Initial Feelings about the Game Task Across Experiments One, Two, and Three.

		Study 1		Study 2		Study 3	
		M	SD	M	SD	M	SD
Subject variables	Handedness (Abs)	74.92	24.99	76.59	22.91	75.36	26.14
	Empathic Concern	3.86	.68	3.85	.71	3.84	.90
	Nature Relatedness	3.22	.89	3.23	.82	3.49	.94
	Conservatism	57.70	17.33	57.48	16.69	54.82	21.07
	Covid Seriousness	72.67	26.49	72.11	25.55	79.81	24.25
	Covid Vaccine	3.77	2.40	3.61	2.41	4.43	2.29
Apprehension about game	Deontologists	4.85	2.08	5.22	1.29	5.51	1.57
	Typical	4.92	1.69	5.42	1.31	5.92	1.26
	Utilitarian	5.00	1.86	5.38	1.38	5.47	1.37
	Total	4.94	1.79	5.37	1.33	5.61	1.42
Desire to profit in game	Deontologists	5.94	3.47	3.11	2.49	5.25	1.68
	Typical	6.33	3.35	3.29	2.54	5.62	1.29
	Utilitarian	6.78	3.75	3.36	2.60	5.36	1.52
	Total	6.43	3.51	3.29	2.55	5.40	1.52

Table 2

Opening Strategy (First Year Bid) Behavior Between Different Moral Responders Across Experiments One, Two, and Three.

		Experiment 1			Experiment 2			Experiment 3		
		N	M	SD	*N*	M	SD	*N*	M	SD
Opening Bids (0-10)	Deontologists	33	4.12	2.53	35	5.47	2.44	91	7.52	2.60
	Typical	135	4.42	2.47	120	5.29	2.35	73	7.22	2.85
	Utilitarian	87	5.14	2.27	93	5.39	2.57	99	6.55	2.78
	Total	255	4.63	2.43	248	5.35	2.44	263	7.07	2.76

Table 3

Results of the Hierarchical Linear Regression on Total Resources Extracted in Game Situations Across Experiments One, Two, and Three.

		Experiment 1			Experiment 2			Experiment 3		
		F	R^2	β	F	R^2	β	F	R^2	β
Step 1		43.05**	.26		86.91**	.42		66.06**	.34	
	Under			.27**			.16**			.21**
	Over			-.32**			-.54**			-.45**
Step 2		Δ4.82**	Δ.03		Δ3.10*	Δ.02		Δ2.60	Δ.01	
	U-Process			-.11*			-.12*			0
	D-Process			-.14**			-.03			-.12*
Step 3		Δ7.86**	Δ.06		Δ.99	Δ.01		Δ1.20	Δ.01	
	Nature Relatedness			-.14*			-.06			-.09
	Empathic Concern			-.09			0			0
	Conservatism			.14**			.09			.05
	Covid Seriousness			-			.04			-

Table 4

Partial Correlations Between Variables Across Experiments One, Two, and Three.

| Experiment 1	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Total Resources	-									
(2) Average Bid	.96**	-								
(3) U-Process	-.11	-.10	-							
(4) D-Process	-.14*	-.16*	-.15*	-						
(5) Handedness Abs	.04	.03	.01	-.01	-					
(6) Nature Relatedness	-.22**	-.23**	-.02	.07	-.09	-				
(7) Empathic Concern	-.21**	-.23**	.14*	.11	.05	.27**	-			
(8) Conservatism	.22**	.23**	.02	-.05	-.04	-.17**	-.25**	-		
(9) Covid Seriousness	-.14*	-.13*	.01	.13*	.16*	.21**	.29**	-.45**	-	
(10) Vaccine Mandate	-.03	-.03	.02	-.09	.15*	.11	.12	-.29**	.37**	-
Experiment 2	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Total Resources	-									
(2) Average Bid	.96**	-								
(3) U-Process	-.14*	-.18**	-							
(4) D-Process	-.04	-.05	-.04	-						
(5) Handedness Abs	0	.01	-.09	.08	-					
(6) Nature Relatedness	-.12	-.12	.05	.07	-.05	-				
(7) Empathic Concern	-.02	-.02	.13*	.13*	.07	.13*	-			
(8) Conservatism	.08	.07	-.11	.05	.09	-.05	-.16*	-		
(9) Covid Seriousness	0	-.01	.05	.02	.02	.04	.26**	-.50**	-	
(10) Vaccine Mandate	-.09	-.09	.12	-.18**	-.10	.05	.10	-.44**	.37**	-
Experiment 3	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Total Resources	-									
(2) Average Bid	.96**	-								
(3) U-Process	-.02	0	-							
(4) D-Process	-.14*	-.12	.16*	-						
(5) Handedness Abs	.05	.09	.09	.20**	-					
(6) Nature Relatedness	-.11	-.12	-.14*	.03	-.05	-				
(7) Empathic Concern	-.06	-.07	.04	.16**	.11	.40**	-			
(8) Conservatism	.05	.04	-.11	-.07	-.11	.15*	-.07	-		
(9) Covid Seriousness	-.01	0	-.03	-.05	.12	.19**	.23**	-.24**	-	
(10) Vaccine Mandate	.01	.03	-.02	-.29**	.03	.06	.08	-.21**	.48**	-
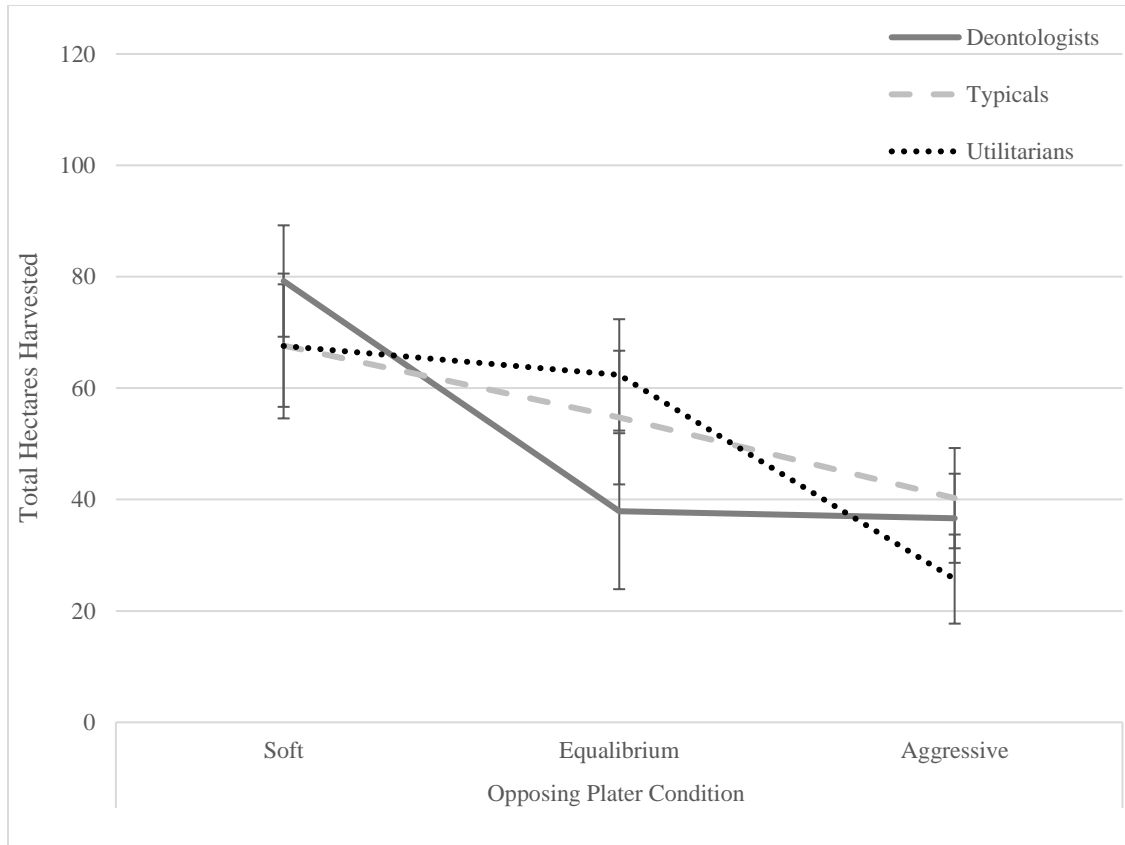
Figure 1. Bidding behavior (mean hectares harvested) separated by moral classification group and opposing player (strategy) condition in Experiment 1.
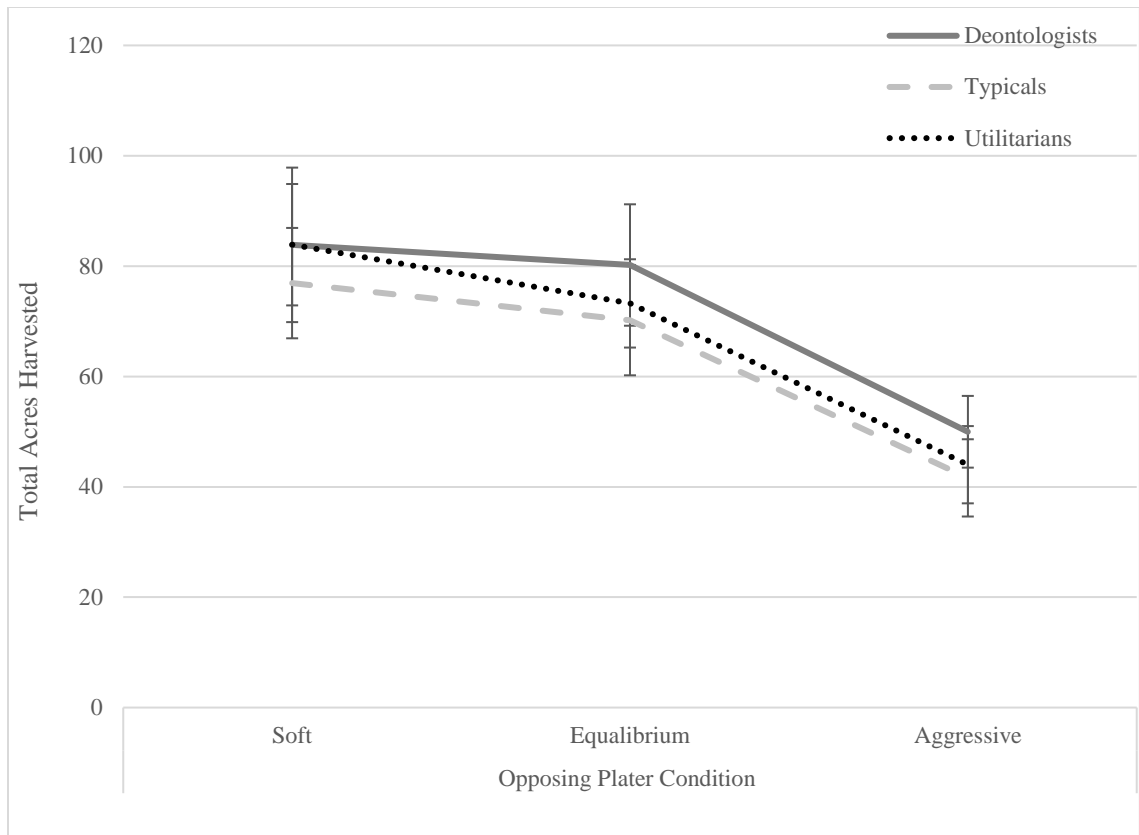
Figure 2. Bidding behavior (mean acres harvested) separated by moral classification group and opposing player (strategy) condition in Experiment 2.
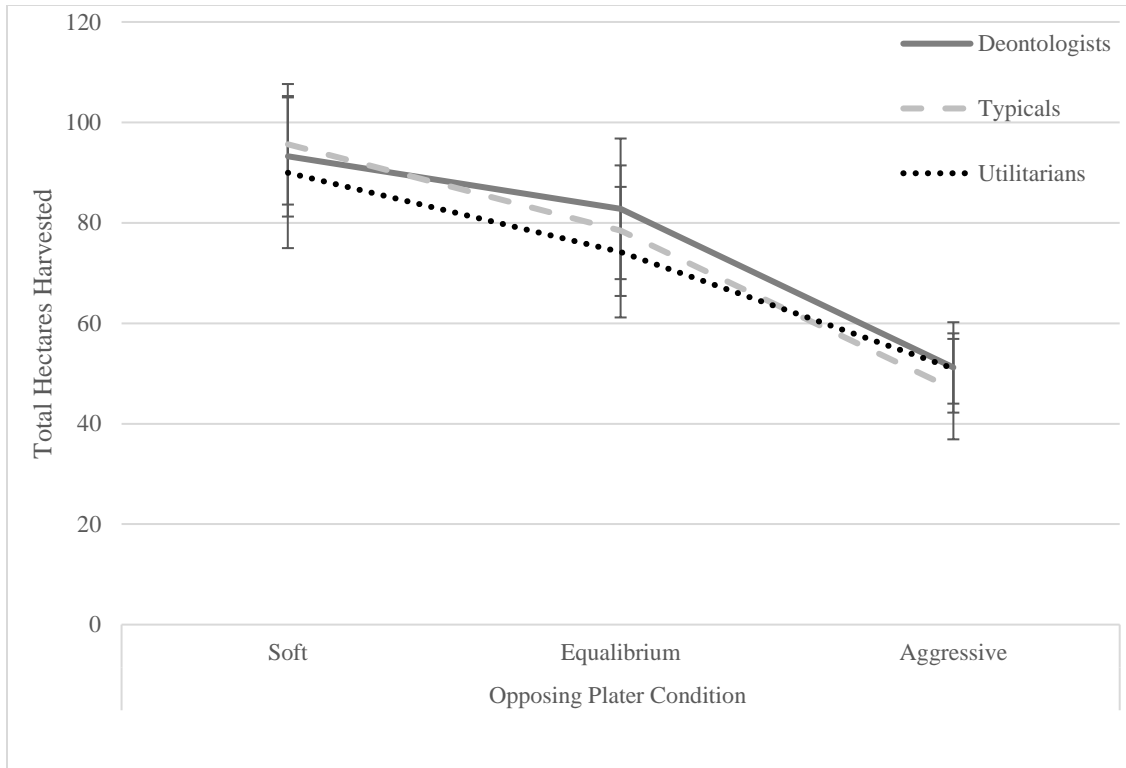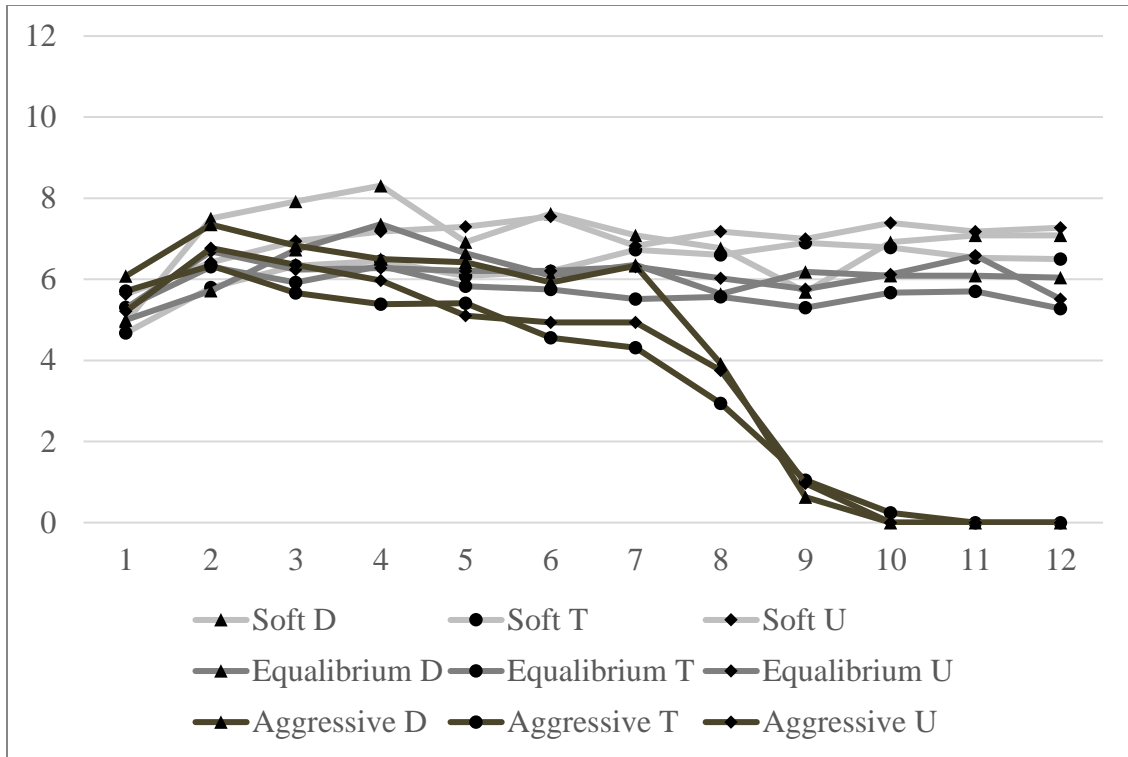
Figure 3. Bidding behavior (mean hectares harvested) separated by moral classification group and opposing player (strategy) condition in Experiment 2.
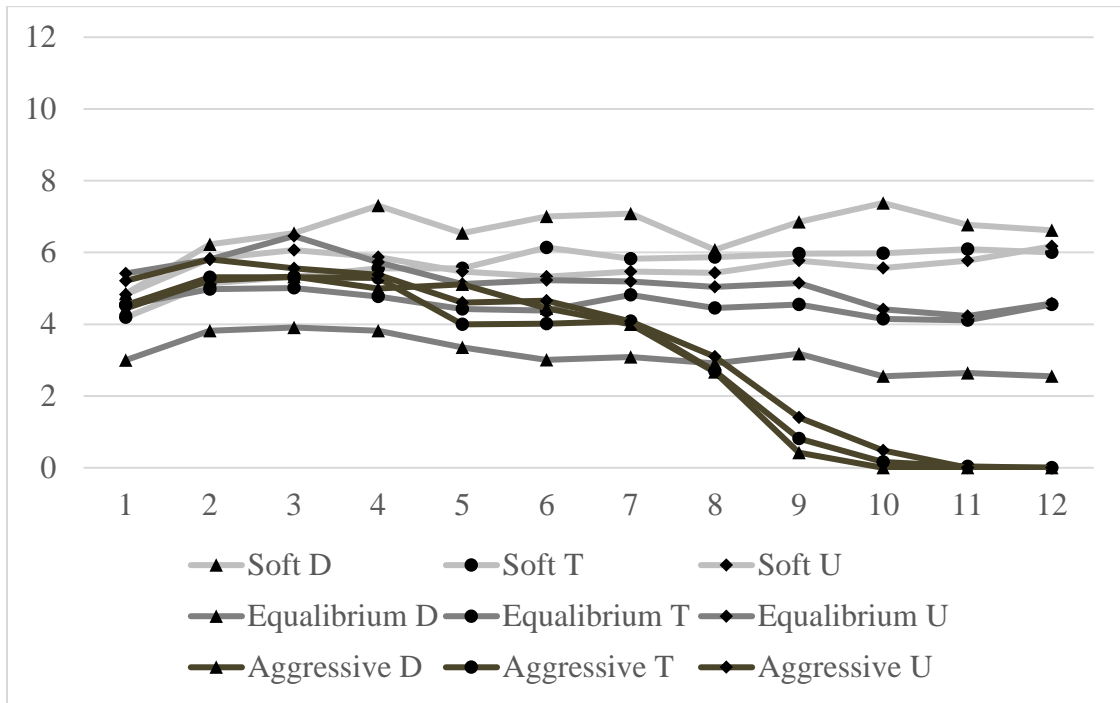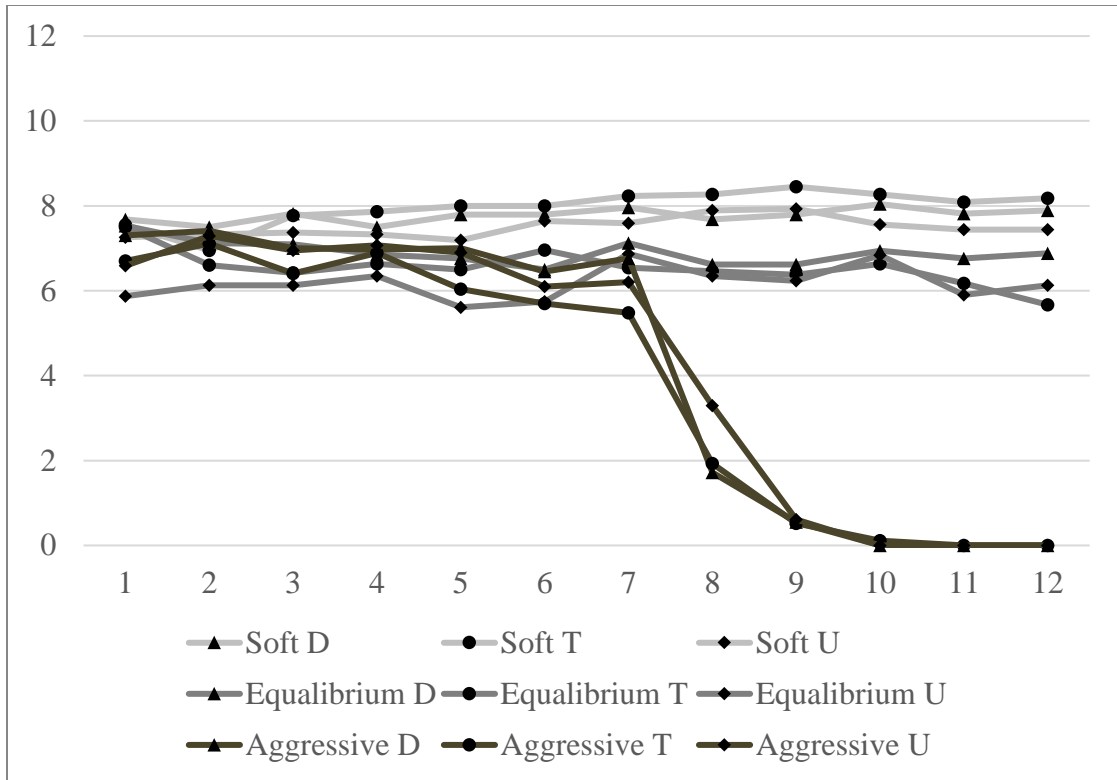
Figure 4. Extraction by round across game conditions and moral type in Experiment 1.

Figure 5. Extraction by round across game conditions and moral type in Experiment 2.

Figure 6. Extraction by round across game conditions and moral type in Experiment 3.

<div align="center">**Appendix A**</div>

<div align="center">**Process Dissociation Dilemmas and Technique**</div>

Incongruent Dilemmas.

  Time Machine. You find a time machine and travel back to the year 1920. While

checking into a hotel, you meet a young Austrian artist and veteran of the First World

War. You realize this is Adolf Hitler before his rise to power in Nazi Germany. He is

staying in the hotel room next to yours and the doors are not locked. It would be easy to

simply smother him with a pillow in his sleep and disappear, stopping the Second World

War and the Nazi party before they even start. However, he has not committed any

crimes yet and it seems wrong to hurt an innocent person. Is it appropriate for you to kill

an innocent young Hitler in order to prevent the Second World War?

  Car Accident. You are driving through a busy city street when all of a sudden a young

mother carrying a child trips and falls into the path of your vehicle. You are going too

fast to break in time; your only hope is to swerve out of the way. Unfortunately, the only

place you can swerve is currently occupied by a little old lady. If you swerve to avoid the

young mother and baby, you will seriously injure or kill the old lady. Is it appropriate to

swerve and hit the old lady in order to avoid the young mother and child?

  Hard Times. You are the head of a poor household in a developing country. Your

crops have failed for the second year in a row, and it appears that you have no way to

feed your family. Your sons, ages eight and ten, are too young to go off to the city where

there are jobs, but your daughter could fare better. You know a man from your village

who lives in the city and who makes sexually explicit films featuring girls such as your

daughter. In front of your daughter, he tells you that in one year of working in his studio

your daughter could earn enough money to keep your family fed for several growing seasons. Is it appropriate for you to employ your daughter in the pornography industry in order to feed your family?

Crying Baby. It is war time. Enemy soldiers have taken over your village. They have orders to kill all remaining civilians. You and some of your townspeople have sought refuge in the cellar of a large house. Outside you hear the voices of soldiers who have come to search the house for valuables. A baby with no parents begins to cry loudly. You cover her mouth to block the sound. If you remove your hand from the baby's mouth her crying will summon the attention of the soldiers who will kill you and the others hiding out in the cellar. To save yourself and the others you must smother the child to death. Is it appropriate for you to smother the child in order to save yourself and the other townspeople from being killed?

Relationship. You are divorced now and your former spouse lives far away. You don't expect to see your former spouse for a very long time. However, you have recently started dating a new person that is positively wonderful. Your new partner is head over heals in love for the first time and you feel equally strong about your partner. You want to do everything in your power to keep the relationship progressing smoothly to what you believe will be a second marriage. Unfortunately, your new partner has told you many times that dating someone who is divorced is totally unacceptable. It is expressly forbidden in your partner's religion. Your partner doesn't know about your ex-spouse. You consider simply pretending that you have never been divorced so that your current relationship can continue. Is it appropriate for you to leave your new partner in the dark

about your previous relationship in order to keep her/ him happy and the relationship alive?

Abortion. You are a surgeon. A young woman you know becomes pregnant, but her body reacts in an unusual fashion. She develops a severe case of preeclampsia, a dangerous syndrome that leads to rapid increases in blood pressure. The only treatment is to deliver the baby. Unless the baby is delivered soon, the mother will die. However, the baby is too young to survive on its own. If it is delivered, it will die. So, although it is very difficult for her, the mother asks you to abort the baby. Is it appropriate for you to perform an abortion in order to save the mother's life?

Torture. You are a police officer, and have recently caught a criminal you have been hunting for some time. He is allegedly responsible for rigging a series of explosive devices: some that have already gone off and some that have yet to detonate. He places explosives outside city cafes and sets them to go off at a time when people are drinking coffee on the patios. In this manner, he has injured many people and might injure many more. Now that the criminal is in custody, you want to know where the unexploded bombs are so you can defuse them. He refuses to talk, so you decide to use "aggressive interrogation techniques" like holding his head under water and beating him. Is it appropriate for you to use "aggressive interrogation techniques" in order to find and defuse the unexploded bombs?

Vaccine Policy. You are a doctor in a health clinic overrun by patients with a serious disease. You just received a shipment of drugs that can cure the disease but the drugs have their own severe sideeffects. If you administer the drugs to your patients, a

small number will die from the side effects but most will live. If you do not, most will die from the disease. Is it appropriate for you to administer the drug to your patients?

Animal Research. You have been hired by a pharmaceutical company to conduct research on their products. Since products must be fit for human use, they are first tried out on animals. Your job is to find out the effects various chemicals have on rats, pigeons, rabbits, and monkeys. Most chemicals have only minor effects on the animals, but some cause them discomfort or even permanent damage. The chemicals you are researching are slated to form part of a new AIDS drug cocktail that will give new hope to millions of AIDS victims around the world. You anticipate saving many lives with the chemicals. Is it appropriate to test these chemicals on animals?

Border Crossing. You are a soldier guarding a border checkpoint between your nation and one troubled by insurgent violence. You notice a young man in a cheap car approaching the checkpoint with a determined look on his face. You suspect he means to bomb the checkpoint, killing all the soldiers inside. He is rapidly approaching your station. Is it appropriate for you to shoot and kill the approaching man?

Congruent Dilemmas.

Time Machine. You find a time machine and travel back to the year 1920. While checking into a hotel, you meet a young petty criminal. You realize this is George Brackman, a man who later on abducted a child and held her for a week until her family paid him some ransom money. He is staying in the hotel room next to yours and the doors are not locked. It would be easy to simply smother him with a pillow in his sleep and disappear, stopping the abduction and ransom demands before they even start. However, he has not committed any crimes yet and it seems wrong to hurt an innocent person. Is it

appropriate for you to kill George Brackman in order to prevent him from taking a child hostage?

Car Accident. You are driving through a busy city street when all of a sudden a young mother carrying a child trips and falls into the path of your vehicle. You are going too fast to break in time; your only hope is to swerve out of the way. Unfortunately, the only place you can swerve is currently occupied by a group of children on their way to elementary school. If you swerve to avoid the young mother and baby, you will seriously injure or kill several of them. Is it appropriate to swerve and hit the schoolchildren in order to avoid the young mother and child?

Hard Times. You are the head of a poor household in a developing country. Your sons, ages eight and ten, work on the family farm every day in order to teach them the value of hard work and providing for others. Your daughter, however, cannot work on the farm due to the heavy lifting required. You worry that she will not learn the value of money if she does not work. You know a man from your village who lives in the city and who makes sexually explicit films featuring girls such as your daughter. In front of your daughter, he tells you that in one year of working in his studio your daughter could earn enough money to keep your family fed for several growing seasons. Is it appropriate for you to employ your daughter in the pornography industry in order to teach her the value of money?

Crying Baby. It is war time. Enemy soldiers have taken over your village. They have orders to capture all remaining civilians to make them work quarrying stone in a mine. You and some of your townspeople have sought refuge in the cellar of a large house. Outside you hear the voices of soldiers who have come to search the house for valuables.

A baby with no parents begins to cry loudly. You cover her mouth to block the sound. If you remove your hand from her mouth the crying will summon the attention of the soldiers who will capture you and the others hiding out in the cellar. To save yourself and the others from laboring in the mine you must smother the child to death. Is it appropriate for you to smother the child in order to save yourself and the other townspeople from being captured?

Relationship. You are divorced now and your former spouse lives far away. You don't expect to see your former spouse for a very long time. However, you have recently started dating a new person that is positively wonderful. Your new partner is head over heals in love for the first time and you feel equally strong about your partner. You want to do everything in your power to keep the relationship progressing smoothly to what you believe will be a second marriage. Unfortunately, your ex-spouse gave you a very serious and incurable Sexually Transmitted Disease (STD). Your partner does not know about it. You consider simply pretending that you don't know about the STD so that your current relationship can continue. Is it appropriate for you to leave your new partner in the dark about your STD in order to keep her/him happy and the relationship alive?

Abortion. You are a surgeon. A young woman you know becomes pregnant, but she is not yet ready for children. She has not finished high school, has no income, and was abandoned by the father. If she has the baby now, she will be stuck as a single mother on welfare for the rest of her life. This will make things very hard on her and the baby. She thinks that it would be smarter to wait and have children later. So, although it is very difficult for her, she asks you to abort the baby. Is it appropriate for you to perform an abortion in order to let the mother live a better life?

112

Torture. You are a police officer, and have recently caught a criminal you have been hunting for some time. He is allegedly responsible for rigging a series of explosive devices: some that have already gone off and some that have yet to detonate. He places explosives outside city cafes and sets them to go off at a time when no one is around. His explosives are inside paint cans so that they spray nearby objects with paint. In this manner, he has sprayed many cafes with paint and might spray many more. Now that the criminal is in custody, you want to know where the unexploded bombs are so you can defuse them. He refuses to talk, so you decide to use "aggressive interrogation techniques" like holding his head under water and beating him. Is it appropriate for you to use "aggressive interrogation techniques" in order to find and defuse the unexploded bombs?

Vaccine Policy. You are a doctor in a health clinic overrun by patients with the latest flu virus. You just received a shipment of drugs that can cure the flu but the drugs have their own severe side-effects. If you administer the drugs to your patients, a small number will die from the side effects but most will live. If you do not, most will continue to suffer from the effects of the flu virus for some time. Is it appropriate for you to administer the drug to your patients?

Animal Research. You have been hired by a pharmaceutical company to conduct research on their products. Since products must be fit for human use, they are first tried out on animals. Your job is to find out the effects various chemicals have on rats, pigeons, rabbits, and monkeys. Most chemicals have only minor effects on the animals, but some cause them discomfort or even permanent damage. The chemicals you are researching are slated to form part of a new acne facial cleanser that will give new hope

113

to people with pimples and greasy skin. You anticipate making many people feel better about their appearance with the chemicals. Is it appropriate to test these chemicals on animals?

Border Crossing. You are a soldier guarding a border checkpoint between your nation and one troubled by insurgent violence. You notice a young man in a cheap car approaching the checkpoint with a determined look on his face. You suspect he means to cross the border in order to work illegally inside your country. He is rapidly approaching your station. Is it appropriate for you to shoot and kill the approaching man?

Appendix B

Foresting Scenario

Please read the following hypothetical scenario carefully.

Now, we would like you to imagine that you are in a particular situation. Try to anticipate what you would actually do in this situation, being as realistic and honest as you can.

Here's the situation (please read carefully): you are the owner of a timber company. Your company and three other timber companies are all working within the same national forest. There are 200 hectares of timbered land within this forest (a hectare is 100 acres, if you're curious). Your goal is to cut as many hectares of trees as you can, so that your company will profit and thrive. Each year, each of the four companies makes bids regarding how many hectares it will cut that year, ranging from a minimum of 0 to a maximum of 10. None of the companies ever finds out what the other companies have bid; all they know is the total number of hectares cut each year. This total amount can range from 0 (if all four companies bid 0) to 40 (if all four companies bid 10). Thus, in a given year, the forest can be reduced by as much as 40 hectares (i.e., in year 1, the forest can shrink from 200 to 160 hectares). Of course the forest also regenerates, at a rate of about 10% per year.

Obviously, one danger is that the forest will eventually be wiped out, leaving all four companies "out in the cold." Thus, it may be to the four companies' collective advantage to make smaller bids. However, another danger is that a company will not do as well because it cuts less than the other three companies. Thus, it may be to each company's individual advantage to make larger bids.

What we are interested in is how you would think and behave in a situation such as this. Please answer the questions below, by imagining yourself in the above situation and anticipating your own responses. There are no "right" answers here; different people resolve such dilemmas in different ways.

Appendix C

Reefs Scenario

Here's the situation (please read carefully): you are the head of state of a country during the Covid-19 pandemic. Very recently, a shocking discovery was made that identified a cure for Covid-19 that is only found in a rare species of coral reefs, found only in a small area located in international waters. The international community is allowing countries to harvest these coral reefs to extract the cure and to distribute it amongst their populations. It is unclear how long it will take before a global vaccine will be ready, so, until that time, harvesting the cure from the coral reefs will be critical. Unfortunately, the supply of these all important cure containing coral reefs is very limited.

Thus, the international community has assigned your country and three others to share a small 200 acres plot of these coral reefs. Your goal is to harvest as many acres of coral reefs as you can, so that your country will be able to save the lives of its afflicted citizens. Currently, in your country, and in each of the three others, about 10,000 people are dying a month from Covid-19.

Each month, each of these four countries can make bids regarding how many acres of coral reefs they will harvest that month, ranging from a minimum of 0 to a maximum of 10. It is estimated that each acre of coral reefs will have enough of the cure to save 1000 people per month.

None of the countries ever finds out what the other countries have bid; all they know is the total number of acres harvested each month. This total amount can range from 0 (if all four countries bid 0) to 40 (if all four companies bid 10). Thus, in a given month, the coral reefs can be reduced by as much as 40 acres (i.e., in month 1, the coral reefs can

shrink from 200 to 160 acres). Of course the coral reef also regenerates, at a rate of about 10% per month.

Obviously, one danger is that the coral reefs will eventually be wiped out, leaving all four countries without a supply of the cure before the vaccine becomes available. Thus, it may be to the four countries' collective advantage to make smaller bids. However, another danger is that a country will not save as many of its own citizens lives because it did not harvest enough of the coral reefs. Thus, it may be to each countries individual advantage to make larger bids.

What we are interested in is how you would think and behave in a situation such as this. Please answer the questions below, by imagining yourself in the above situation and anticipating your own responses. There are no "right" answers here; different people resolve such dilemmas in different ways.

Appendix D

Switch and Footbridge Dilemmas

You will now be asked to make responses related to specific hypothetical situations.

Please read each scenario carefully and indicate your responses below.

Imagine that you are at the wheel of a runaway trolley quickly approaching a fork in the tracks. On the tracks going to the left is a group of five railway workmen. On the tracks going to the right is a single railway workman. If you do nothing, the trolley will proceed to the left, causing the deaths of the five workmen. The only way to avoid the deaths of these workmen is to hit a switch on your dashboard that will cause the trolley to proceed to the right, causing the death of the single workman.

Would you hit the switch in order to avoid the deaths of the five workmen?

Yes or No?

Imagine that a runaway trolley is heading down the tracks toward five workmen who will be killed if the trolley keeps going. You are on a footbridge over the tracks in between the approaching trolley and the five workmen. Next to you on this footbridge is a stranger who happens to be very large. The only way to save the lives of the five workmen is to push this stranger off the bridge and onto the tracks below where his large body will stop the trolley. The stranger will die if you do this, but the five workmen will be saved.

Would you push the stranger onto the tracks in order to save the five workmen?

Yes or No?

Appendix E

Moral Classification Quadrant

Switch Scenario

		Non-Utilitarian Response	Utilitarian Response
Foot Bridge Scenario	Non-Utilitarian Response	Deontological	Typical
	Utilitarian Response	Weird	Utilitarian