Title
What were they thinking?: The Impact of Mental State Explanations on Legal Decisions

Permalink
https://escholarship.org/uc/item/30f9m365

Author
Giffin, Carly

Publication Date
2018

Peer reviewed|Thesis/dissertation

What were they thinking?: The Impact of Mental State Explanations on Legal Decisions

By

Carly Giffin

A dissertation submitted in partial satisfaction of the

Requirements for the degree of

Doctor of Philosophy

In

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Tania Lombrozo, Chair
Professor Avani Mehta Sood
Professor Juliana Schroeder

Summer 2018

Copyright Permissions

Study in Chapter 2, "Wrong or Merely Prohibited: Special Treatment of Strict Liability in Intuitive Moral Judgment," used with permission from Wiley Publishing.

Study in Chapter 3, "An Actor's Knowledge and Intent are More Important in Evaluating Moral Transgressions than Conventional Transgressions," used with permission from the American Psychological Association.

Abstract

What were they thinking?: The Impact of Mental State Explanations on Legal Decisions

by

Carly Elise Giffin

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Tania Lombrozo, Chair

People have an intuitive sense of what makes a good explanation for behavior. Confronted with a particular action, most people believe they could provide explanations that would be mitigating, that would be *good* explanations for the behavior. The law also has expectations about the impact of explanations. The American legal system explicitly classifies certain explanations as defenses that if believed should reduce punishment or excuse the behavior entirely. Across three studies, this dissertation finds that the reality of people's behavior in response to an explanation often proves more complicated than laypeople or the law expect. Specifically, this dissertation considers the often unexpected ways that mental state explanations impact legal decisions.

Chapter 2 finds that ignorance of vital facts is not always a mitigating explanation. Across three experiments, we find that ignorance was a mitigating explanation for one class of crime, but not another, a cognitive distinction that mirrors a legal classification. This result runs counter to both intuitive expectation and scholarly argument that ignorance of the fact that makes an action illegal would always be a mitigating explanation. We also find statistical support for what may drive this distinction at a cognitive level, perhaps explaining how the crimes were classified by the legal system. We found that ignorance was a less mitigating explanation for crimes that appeared to have been set in an arbitrary manner, in that the line of illegality could have been drawn differently. Chapter 3 replicated the main findings of Chapter 2 in a non-legal setting, again finding that ignorance is a less mitigating explanation for some violations than others. In four experiments, we find that ignorance is a less mitigating explanation for violations of school rules that appear to have been set in an arbitrary manner. This study allowed us to manipulate arbitrariness experimentally, supporting our argument that it is an actor's belief about the arbitrariness of the rule, not whether it has actually be set in an arbitrary manner, that determines whether mental state explanations will be mitigating. Chapter 4 approaches mental state explanations differently, asking how people evaluate these explanations and whether these evaluation strategies impact legal outcomes. We find that which mental state explanation a defendant offers does change how that explanation is evaluated, and that evaluation strategies do impact legal decisions. Thus, this dissertation proposes that the relationship between mental state explanations and legal outcomes is consequential but complicated.

Dedication

To Professors Marilyn Boltz, Beth Burkstrand Reid, and Tania Lombrozo, who helped me
navigate at times when I felt quite lost and gave me courage to keep going.

And to my parents, Bruce and Kay Giffin, who answered all the "why" questions I asked without
resorting to "just because," and who listened to all my "did you know…" explanations without
rolling their eyes. You taught me that my questions were worth asking, and that I was capable of
finding the answers.

Table of Contents

"The juror's task involves the kind of uncertainty, ambiguity, and potential for situation/information interaction that is typical of our normal experience with the environment."
Pennington & Hastie (1981)

Chapter 1: Introduction

The opening quote from researchers Pennington and Hastie suggests that what a juror is asked to do in a trial is not actually much different from the kinds of judgments we all make every day, it is just done in a particular context with weightier outcomes. When I assess the explanation of a friend or coworker (is she telling me the truth about what she thought and intended?), my decision is based on some level of uncertainty. I cannot actually know what she was thinking, I have to decide what I *believe* she was thinking based on what I know of her actions and what she and others tell me. The task faced by many jurors is similar in type if not magnitude. A juror is frequently asked to assess the mental state of a defendant – what did they think, know, believe, or intend when they acted – and jurors are also faced with uncertainty.[1] What we learn about how jurors, or mock jurors,[2] do their jobs can tell us a great deal about how people evaluate explanations and make decisions in their everyday lives. While some insights are specifically legal, many others, as the studies described below will show, apply to explanation and decision making more broadly. Studying explanations and decision making in the legal system offers a concrete context in which to study processes of interest to both legal scholars and psychologists.

In a criminal trial, the jury has one large decision to make. The jury's ultimate job is to decide whether the defendant will be found guilty of the crime with which she has been charged. To make this decision, the jury will be presented with evidence and hear from witnesses, and they will also hear explanations of this evidence from both the prosecution and the defense. These explanations are crucial decision making tools because while neither side can change what has been seen or heard inside the courtroom, each can offer explanations for this material that encourage jurors to draw some inferences, but not others. The prosecution's explanation of the material will encourage the jury to come to the conclusion, beyond a reasonable doubt, that the defendant committed the crime. The defense's explanation will encourage entirely different inferences, and may address either – or both – of the two main elements of a crime.

The first main element of a crime is the *actus reus* which is defined as the acts or omissions that comprise the physical elements of a crime as required by statute (e.g. *Schad v. Arizona*, 1991). The *actus reus* of theft, for instance, would be the physical taking and moving of another person's property (California Penal Code, § 484). A defense explanation aimed at this

---

[1] It is not clear whether a juror faces more or less uncertainty than we do when making more prosaic assessments. The juror will often not be allowed to hear all testimony or see all evidence as some will be bared by rules of evidence that do not apply when judging a friend. However, these rules of evidence, and the fact that witnesses testify under oath with associated penalties for perjury, may increase the reliability of the information they do receive.

[2] The author is not aware of any experimental work conducted on actual jurors either during or after a trial. A few qualitative articles exist that detail interviews with jurors after the completion of a trial or examine archival information, but they do not manipulate the jurors during trial (*see, e.g.*, Diamond & Rose, 2005). In a few articles (i.e. Pennington & Hastie, 1986), participants were people who had been called to jury duty, but they performed the experiment rather than serve as jurors. All experiments in this Dissertation were conducted on mock jurors.

element would argue that the evidence presented does not prove that the defendant committed the crime – the taking of the property, and thus they should be found not guilty. In their explanation, the defense may note other suspects or the lack of direct evidence that the defendant committed these acts. The second main element of most crimes[3] is *mens rea*. *Mens rea* is Latin for "guilty mind" and is defined as the state of mind statutorily required in order to convict a particular defendant of a particular crime (e.g. *Staples v. United States*, 1994). The *mens rea* for theft requires that the defendant *knew* that the property was not his own when he carried it away (California Penal Code, § 484). An explanation aimed at *mens rea* argues that the defendant did not have the knowledge or intent required by law to hold him responsible for the crime. Perhaps they believed that the property they took was their own or had been abandoned. If that explanation is believed, then the defendant could not be convicted of theft because he did not know that the property belonged to someone else when he took it, meaning he did not have the requisite *mens rea* to be convicted of the crime.

It is this second class of explanations, those aimed at *mens rea* or mental states, that are the focus of this dissertation. This class of explanations acknowledges that the defendant committed the *actus reus*, but explain that the defendant acted without the mental state – the knowledge, beliefs, or intent – that would render him fully culpable for his behavior, and, therefore, he should be given some measure of leniency or found different of a lesser crime. I call these kind of explanations "yes, but…" explanations. These explanations argue that yes, the person committed the act, but his mental state should mitigate[4] our judgment of that behavior. While this argument appears straightforward, my research finds that "yes, but…" explanations do not always behave as the law or lay people expect. When are these explanations more or less mitigating? What strategies do people use to evaluate the mental state of a defendant? These questions are not only legal, but also psychological, touching on areas of psychological interest including explanation, mental states, decision making, and morality. This dissertation will begin to address these questions with studies that focus on the legal system and wrong doing, but the results will demonstrate that what is discovered in this specialized context is relevant to a more general understanding of the concepts.

The first section of Chapter One will consider the history of *mens rea*, including how mental states gained their importance in the American legal system. The second section of Chapter 1 reviews psychological research that has examined whether lay people's concepts of knowledge, belief, and intent coincide with the legal definition of these terms, the varying influence of mental states across different contexts, and what strategies people use to assess a defendant's mental state. This section will also preview the results of the original experiments contained in chapters two through four. Chapter Two describes a series of experiments that find ignorance of vital facts is not a mitigating explanation for some crimes, a finding which suggests a correspondence between legal and cognitive categories. Chapter Three details a set of experiments that replicate the main findings of Chapter Two in a non-legal context, and experimentally support a theory about why ignorance is not a mitigating explanation for some

---

[3] See later in this introduction and Chapter 2 for more detail about strict liability crimes – crimes that do not require *mens rea*.

[4] A lesser mental state could also mean that the defendant is guilty of a lesser crime. Charging a defendant with this lesser crimes is not mitigation, but would rather represent correctly charging the offense committed under the law. However, when I speak of mitigation here, I refer to all judgments of behavior, not only official judgments made within the legal system.

violations. These two chapters both find that "yes, but…" explanations that point to a defendant's ignorance that she was even committing a crime is not always mitigating, a finding that runs counter to legal and lay expectations. Chapter Four expands the discussion of "yes, but…" explanations with a series of experiments that seeks to discover the strategies people use to evaluate these explanations and whether different strategies has systematic impacts on legal outcomes.

Some of these findings may be troubling for the legal system which has clear expectations and rules about what information should be mitigating or exacerbating. However, gaining a better understanding of how these psychological concepts operate in the legal system is invaluable for any attempt at future reforms or restructuring. Additionally, as suggested earlier, this research is finding that the structure of the legal system may reflect the structure of our minds, offering a concrete starting point for considering how people evaluate the mental state of others and perhaps suggesting ways in which the law could – or should – be restructured.

The evolving role of *mens rea* in American criminal law

For most crimes, the statute specifies a *mens rea*. That is, for most crimes, the statute specifies what mental state the defendant must have had to be convicted of the crime. For instance, if a defendant started a small fire in their backyard, so that the fire would spread and burn their neighbor's house, that would be the crime of arson because she "willfully and maliciously" caused the neighbor's house to burn (California Penal Code, §451). However, if she started the fire merely to burn some brush and it spread to her neighbor's house, it might be the crime of "unlawfully causing a fire," if it was determined that the defendant acted "recklessly" (California Penal Code, §452). Arson is associated with greater punishment in the legal system and greater moral condemnation everywhere. Thus, a different mental state on the part of the defendant can change which crime she is charged with and the penalties she faces because she does or does not meet the specified *mens rea*. Broadly, *mens rea* is an attempt to carefully parse and codify notions of knowledge, belief, and intent that are interesting to scholars and, as this example illustrates, consequential to defendants. In fact, the majority of criminal statutes in the United States now require some level of *mens rea*,[5] but this was not always the case historically. The fact that *mens rea* was not always an integral part of the criminal law begs the question of why it is now? Or, perhaps, why it was not then?

The English system of law, on which our own is based, began with no consideration for mental states. Its purpose was merely to assure the swift compensation of the victim or victim's family to avert bloodshed (Gardner, 1993; Perkins, 1939; Sayre, 1932). Prior to the institution of a (more) formal system of law, a wronged party would seek compensation in the form of vengeance against the wrong doer and sometimes the wrong doer's entire family. It was these so-called Blood Feuds that English law was originally seeking to avoid. In that context, it did not matter whether a person killed her neighbor's sheep accidentally or on purpose, it was only important that the sheep be replaced before any vengeance was sought. However, more complicated cases were difficult to fit into this model. It is not clear how to assess the monetary fine for some actions, and for other actions it is not clear who ought to receive the money. What

---

[5] Although, especially at the federal level, the number of crimes which do not require *mens rea*, so called strict liability crimes, are proliferating in the United States (House Report No. 113–46 (2013)). As indicated in footnote 1, we will discuss this in more detail later in the chapter.

is the monetary worth of a particular human life? To whom do you pay that fine if the deceased had no family? The answers to those questions were not clear in the 12th century, they are arguably not clear in our own.

Thus, the law evolved to consider more than monetary fines, eventually imposing sentences that included imprisonment. As the emphasis shifted from compensating the victim to punishing the wrong doer, what the defendant knew or intended began to seem more important. In this new system, the king might award pardons in cases of clear accident or self-defense as an acknowledgment that the defendant lacked a fully culpable mental state (Gardner, 1993). While this growing focus on culpable mental states coincides with intuitive notions that ignorance or lack of intent should be mitigating, the shifting focus owes its roots to something more formal than intuition. The church was gaining influence in the 12th and 13th centuries, and it did not condone punishment without moral desert. The church argued that moral desert was lacking if a culpable mental state was not present. In fact, the first reference to *mens rea* as a consideration of English law is drawn from a canonical book (Sayre, 1932). As the church gained influence, the law had to respond to these concerns.

Thus, the inclusion of *mens rea* as an important element of first English and later American criminal law was a result of the church's belief that an act performed without a culpable mental state did not deserve punishment. A lack of *mens rea* was, from its first usage in law, believed to be a mitigating explanation for behavior. An actor who explained "yes I committed that act, but I did not know/intend the harm" was seen to be less deserving of punishment. This now foundational tenant of American law feels intuitively correct to most people as well, even to people who have no ties to organized religion. But do lay people's evaluation of mental states always – or even usually – coincide with legal categories and rules? What do we actually know about how people evaluate these different explanations and what factors influence these evaluations? The next section will explore psychological research that begins to address these questions, including the original work that is the focus of this dissertation.

What can research tell us about the evaluation of "yes, but…" explanations?

The church's insistence on culpable mental states to find a defendant morally and legally blameworthy presages the bulk of psychological findings. Studies typically support the expectations of both the church and the legal system: as knowledge and intent increase so too do moral condemnation and punishment (i.e. Giffin & Lombrozo, 2016; Mueller, Solan, & Darley, 2012; Shen, Hoffman, Jones, Greene, & Marois, 2011; Mikhail, 2009). A recent study has even found that actors are blamed for a particular mental state, "wicked desires," more than generally bad character (Inbar, Pizarro, & Cushman, 2012). However, the picture is not as uniform as we might intuitively imagine. Research finds striking instances of disassociation between mental states and condemnation, and these disassociations reveal a great deal about how people evaluate mental states in general, and "yes, but…" explanations in particular. First, we will review research conducted without reference to the legal system that bears on these issues. Second, we will consider literature that examines whether laypeople make – or are even capable of making – the kind of distinctions between mental states that are required by law. Finally, we'll consider what strategies people use to evaluate mental state explanations.

As alluded to above, the importance of an actor's knowledge and intent matters more for some actions than others. Young and colleagues have found that an actor's knowledge has less of an impact on later judgements of that behavior if the behavior violated purity rather than harm standards (Young & Tsoi, 2013; Young & Saxe, 2011).[6] For instance, people find the actor's knowledge less important if the act was incest – a purity violation, than if the act was poisoning – a harm violation. The authors suggest that this difference may be due in part to the fact that purity crimes tend to be "victimless" (Young & Saxe, 2011). They reason that people may care less what an actor intended or knew if the consequences of that action are only for themselves and other knowing participants. If two people decide to knowingly commit incest, they are, after all, only impacting themselves. It is not as important for others to assess their mental states because the behavior does not now, and is unlikely to ever, impact the evaluator.

However, more recent work laid out in full in Chapter Three suggests an additional explanation for why an actor's knowledge is more important for some violations than others: the perception that the rule which has been broken was set arbitrarily (Giffin & Lombrozo, 2017). This work was inspired by previous work we conducted that is described more fully in the next section and Chapter Two. In that study, participants were told either that the actor knew that his or her behavior violated a rule (knowingly wore an unacceptable shirt) or that he or she violated the rule unknowingly (innocently believed their shirt would be acceptable). Further, half of the rules were set in an arbitrary manner, in the sense that rule could reasonably have been different. For instance, the required shirt could have been any color.[7] The other half of the rules were not arbitrary – no reasonable alternative comes to mind for a rule against hitting other students. We found that the actor's ignorance had significantly less impact on later judgments of blame and punishment if the rule that was violated was arbitrary, such as our dress code example, rather than non-arbitrary, such as hitting. Participants did not give less censure or punishment to an actor who mistakenly wore the wrong shirt than they did to an actor who wore the wrong shirt knowingly. However, an actor who accidentally hit another person did receive significantly less censure and punishment than one who did so intentionally.

These findings suggest that part of the reason that performing an act knowingly seems intuitively more worthy of blame and punishment is because of the additional intentions that go along with this knowing action. Knowingly hitting another person has associated consequences – hitting causes pain and distress. Thus, when an actor knowingly hits another, they are also seen as intending to cause that pain and distress.[8] Conversely, when an actor knowingly wears the wrong shirt, it is not clear that they intend any further consequences. This helps explain why ignorance is significantly mitigating for actions like hitting: accidentally hitting someone means that the actor not only did not intend to break the rule, but also did not intend the additional consequences. That seems importantly different when assigning censure and punishment. An actor who knowingly violates a rule about dress code, though, does not seem to intend any additional consequences. Thus, the only difference between a knowing and accidental rule

---

[6] Young's work considers harm and purity violations to be separate kinds of violations, perhaps even traceable to different regions of the brain, but not all authors agree with them (*see e.g.* Schein & Grey, 2017).

[7] See the full experiment in Chapter 3 for a discussion of how a possible confound between harm and arbitrariness was addressed.

[8] At the very least, they would be aware that they are causing the pain and distress.

breaker is whether they knowingly broke the rule. This may not license a significant difference in censure and punishment (Giffin & Lombrozo, 2017).

The arbitrary and non-arbitrary rule distinction is reminiscent of the conventional and moral distinction common in the moral psychology literature (Turiel, 2008; Weston & Turiel, 1980). However, we find that even rules dealing with moral subject matter, causing harm, display this insensitivity to knowledge if participants *believe* the rule was set arbitrarily (Giffin & Lombrozo, 2017). These studies suggests that the mental state with which an actor acts is important to censure and punishment judgments to the extent that it signals something further about the actor's beliefs and intentions. This would also be consistent with the purity work discussed earlier – knowingly committing a purity violation does not indicate that you intend or believe anything further than committing this act.

The figure below, applicable to Chapters 2 and 3, offers a way to visualize how the arbitrariness of a law or rule impact the importance of knowledge, the rule itself, and the perceived associated consequences. Our results (Giffin & Lombrozo, 2017) suggest that as the arbitrariness of a rule or law increases, the perception that the actor intends to cause additional consequences that are wrong in themselves – such as physical harm – decreases. That is, an actor who knowingly violates an arbitrarily set law is not believed to knowingly cause an inherently bad outcome. Thus, an actor's knowledge or ignorance becomes less important to judging their behavior. Correspondingly, the importance of the rule or law to judging the actor's behavior increases as the arbitrariness increases because in the absence of intended, inherently bad consequences a greater amount of the censure rests on having violated a set rule or law.[9] Both strict liability laws and conventional rules tend to be more arbitrary, while *mens rea* laws and moral rules tend to be less arbitrary, see Figure 1-1.
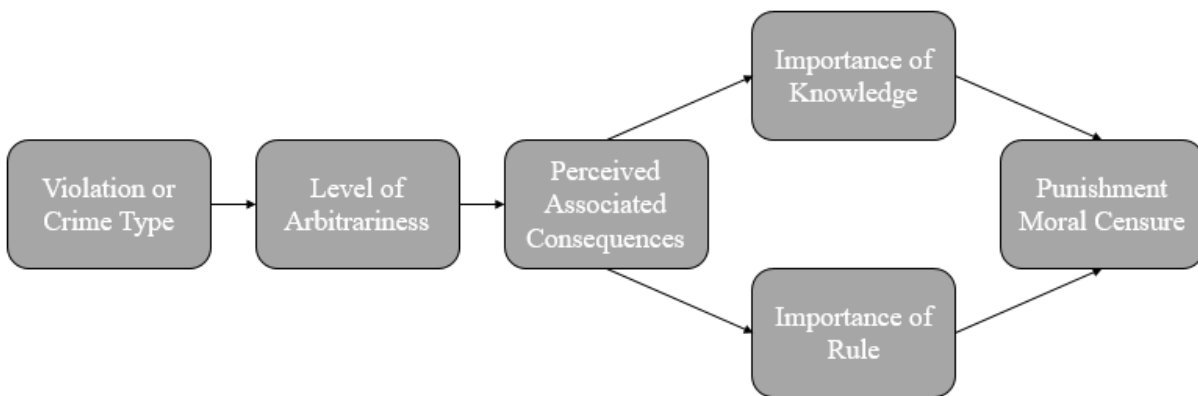
Figure 1-1: Impact of arbitrariness of law formation on importance of knowledge.

Do lay people and the law draw the same mental state distinctions?

The findings in the previous section support the argument that there is a link between mental states, moral censure, and punishment, but these studies were not focused on the kind of fine distinctions the legal system asks people to make. This is important because these fine

---

[9] The importance of a rule will be discussed in more detail in Chapter 3.

distinctions can have a major impact on the punishment a defendant receives. Research does show that people can and do make fine distinctions between beliefs, desires, and intent when assessing moral wrongness, blame, and punishment. Unfortunately, research suggests that the distinctions people make naturally may not be the ones the law expects them to make.

For instance, people blame and punish an actor more if he acted with intentionality, but intentionality is not a unitary concept. The concept of intention, used frequently in *mens rea* statutes, is composed of a belief that an action will cause a result and the desire to bring that result about. Intentionality, as defined by some psychologists, is composed of intention, skill, and awareness of fulfilling the intention while performing the action (Malle & Nelson, 2003).[10] Other research has found that these components are important to moral judgments, though, again, the type of judgment being made matters. For instance, research has found that judgments of moral wrongness depend more on the beliefs, desires, and actions of an individual, while punishment and blame[11] depend on those things as well as the outcome of the action (Cushman, 2015; Cushman, 2008). The Path Theory of blame also considers the outcome to be fundamental. This model argues that upon detecting an agent who caused an event, people assess blame by considering whether it was intentional – and if so consider reasons why the actor may have acted in this way – or unintentional – in which case people consider whether the actor had the duty and ability to behave differently (Malle, Guglielmo, & Monroe, 2014).

These findings track the general structure of the law. In the United States, people are not held criminally responsible for their thoughts (even conspiracy requires some act in furtherance), and the system considers defenses as well as duties. Thus, the finding that punishment depends in part on outcome and that people consider the reasons for acting are encouraging evidence that the law is representing a cognitive reality. However, *mens rea* statutes and judicial instructions are more precise than this general outline. Statutes and instructions tend to use "intention" or "intend" or "purposely" and "knowingly,"[12] when Malle and Nelson's work suggests that the word these statutes, arguably, ought to be using is intentionality, a word I have never encountered in a *mens rea* provision in either my legal education or statutory research. It is thus fair to say that people have a complex and detailed system for assigning wrongness, blame, and punishment judgments that follow the broad outlines of the law. Recent research suggests, however, that when distinctions get more precise, as they tend to in *mens rea* provisions, people's intuitive concepts and the law fall out of line.

Mens rea distinctions are set in statute and most states base their statutes on the Model Penal Code (MPC).[13] The MPC recognizes four *mens rea* levels: purposely, knowingly, recklessly, and negligently. For example, section 2.02(b) states that a defendant acted knowingly in regards to an element of the crime when, "(i) if the element involves the nature of his conduct

---

[10] Detailed and lengthy discussions about intention and intentionality exist in other fields, notably philosophy, but that discussion is beyond the scope of this dissertation.

[11] The definitions used in these psychological studies are not written or intended to match up with precise legal definitions.

[12] The exact terms used vary by statute and jurisdiction.

[13] The MPC is a document first drafted by the American Law Institute in 1962 to try to offer some clarity and uniformity to penal statutes, and the majority of states now base their codes, at least in part, on the MPC, making its definitions important starting points for discussion.

or the attendant circumstances,[14] he is aware that his conduct is of that nature or that such circumstances exist; and (ii) if the element involves a result of his conduct, he is aware that it is practically certain that his conduct will cause such a result." However, according to section 2.02(c), the defendant has acted recklessly if instead he, "… consciously disregards a substantial and unjustifiable risk that the material element exists or will result from his conduct. The risk must be of such a nature and degree that, considering the nature and purpose of the actor's conduct and the circumstances known to him, its disregard involves a gross deviation from the standard of conduct that a law-abiding person would observe in the actor's situation" (ALI, 1985). This type of extremely fine parsing is a particular specialty of the legal system, sometimes with good reason, but research casts doubt on whether or not lay people can or do make the same distinctions.

For instance, the difference between intentionally and knowingly committing an act has important consequences for moral judgments (Ames & Fiske, 2015), and while this distinction mirrors the MPC distinction of purposely and knowingly,[15] not all statutes differentiate these levels (Mikhail, 2009). In this case, it seems that people may make finer distinctions than some statutes. Newer research confirms that people can reliably differentiate between the *mens rea* categories of purposeful and knowing, as well as reckless and negligent – with or without jury instructions, indicating that these categories do track intuitive notions (Shen, Hoffman, Jones, Greene, & Marois, 2011). However, participants performed at just above chance levels when asked to differentiate between knowing and reckless, and this conflation is also seen in their punishment judgments (Shen, Hoffman, Jones, Greene, & Marois, 2011). Thus, while people can and do make fine distinctions between different mental states, they do not appear to be consistently making the same fine distinctions as *mens rea* statutes. These results have important implications for the adjudication of cases, implications that may lie in extremely specific wording choices. This question of whether lay notions of mental states coincide with legal classifications can, however, be taken one step broader: do people and the law agree about when a defendant's knowledge should matter at all? This question addresses a crime classification. To be found guilty of most crimes in the United States, the prosecution must prove that the defendant had the requisite *mens rea*. Most crimes, but not all.

Strict liability crimes are crimes for which the prosecution is not required to prove that the defendant had any *mens rea*. No unified theory explains why seemingly disparate crimes are all included in this category, but speeding is an accessible example. A person can get a speeding ticket regardless of whether she knew she was speeding. Even if she could prove that her speedometer was broken, through no fault of her own, she could still be forced to pay her ticket. Strict liability crimes have no *mens rea* requirement, so her faultless ignorance is no defense.[16]

---

[14] For the non-legal reader, attendant circumstances is a phrase that covers a large number of possible facts, such as a person's age, that might be legally relevant.

[15] The MPC defines conduct as performed "purposefully" when "(i) if the element involves the nature of his conduct or a result thereof, it is his conscious object to engage in conduct of that nature or to cause such a result; and (ii) if the element involves the attendant circumstances, he is aware of the existence of such circumstances or he believes or hopes that they exist" (MPC, § 2.02(a)).

[16] Some strict liability crimes, in some states, do allow for a reasonable mistake of fact defense. For instance, in some states it is a defense to a charge of statutory rape that the defendant reasonable believed his partner was of age, for instance meeting them in a bar where they were being served alcohol. However, in *Garnett v. State*, it was no defense to statutory rape that the defendant had an IQ of 52, and a mental age of approximately 11 or 12, younger

While speeding has fairly minimal punitive or reputational consequences, other strict liability crimes, such as statutory rape (where the prosecutor is not required to prove that the defendant knew the age of their sexual partner) carry the possibility of heavy punitive and reputational consequences. Legal scholars have long criticized strict liability, arguing that a lack of knowledge is an important mitigating fact when judging a defendant. Based on the findings reviewed above, we might have predicted they were correct.

Surprisingly, in the study detailed in Chapter Two, we found instead that participants agreed with the legal distinction (Giffin & Lombrozo, 2016). Participants found ignorance significantly less mitigating for strict liability crimes than crimes that required *mens rea*. Across several studies, ignorance did not significantly mitigate scores of moral wrongness or punishment. That is, participants gave just as much moral censure and punishment to the defendant who knowingly broke the speed limit or slept with an underage girl as they did to the defendant who did so with the completely reasonable false belief that they were driving the speed limit or that the girl was over 18. This was true for a range of strict liability crimes, from speeding to statutory rape to shooting a migratory bird. For crimes that did require *mens rea*, the picture was different. Participants gave higher censure and punishment when defendants intentionally took another's property than when they did so believing the property to have been abandoned, for example. These results mirror the law, with participants finding ignorance a mitigating explanation when the law required *mens rea*, but not for strict liability crimes. Our participants agreed with the law about when ignorance was a good explanation for behavior.

After experimentally ruling out several possible explanations for this disparity in treatment, we considered the legal distinction between *malum prohibitum* – wrong as prohibited – and *malum in se* – wrong in itself.[17] As noted in the figure above, the logic follows the same pattern as discussed above for violations of arbitrary and non-arbitrary rules. We noted that the precise line of illegality drawn for strict liability crimes did seem arbitrary, it could have been drawn at a different point, and was appeared *malum prohibitum*. While we may all agree it should be illegal for drivers to drive too fast, what constitutes "too fast" is not clear. The speed limit is 35 MPH on some streets, and 45 or 25 MPH on others, for reasons that are not clearly apparent. Thus, driving 35 MPH is not inherently wrong or punishable, but it is if the speed limit is 25 MPH. Conversely, crimes that require *mens rea* seem more *malum in se*. Theft is illegal because it deprives me of the use and enjoyment of my property. The line of illegality is quite clear: you have either taken my property or you have not. This relates to the mitigating impact of ignorance. The arbitrary nature of the line of illegality drawn for strict liability crimes means that it is not clear that a knowing offender intends any additional consequences. It may thus seem that a knowing violator of a strict liability law does not deserve much more censure or punishment than an accidental violator. Conversely, taking another's property has known consequences in addition to breaking the law, so a defendant who knowingly breaks these laws is held to be intending those additional consequences, making accidentally doing so importantly different.

Thus, the law and popular intuition appear to agree about when *mens rea* is important and the general idea of increasing moral culpability and punishment as defendants progress from

---

than the chronological age of his victim. The court ruled that in the state of Maryland the crime was strict liability so it was irrelevant that Garnett was perhaps not even capable of having the requisite *mens rea* (1993).

[17] No clear definition of what makes one crime *malum prohibitum* and another *malum in se* is offered. Rather, Justice Jackson, writing the opinion for the court, in *US v. Morissette*, seemed to believe that it would be an obvious matter to classify crimes as one or the other.

negligently to purposefully performing an act. People do not, though, consistently make the same distinctions *mens rea* does, even though they agree when it is important to do so. This and the prior section provide converging evidence, using legally relevant and more general stimuli, to show that the role of mental state explanations is not as predictable as we might have suspected. Their impact varies depending on the kind of act as well as how the rule has been set. This converging evidence is important not only because it speaks to the robustness of these findings, but also because it indicates that what we learn about cognition from studying the structure of the legal system, also generalizes to non-legal categories and decisions. However, this research does not address a fundamental question about mental state explanations. How do people evaluate them?

What strategies do people use to evaluate mental state explanations?

Determining how people evaluate mental states is not an isolated legal question. Evaluating the mental states of other people is a task we all perform, usually multiple times per day. *Why did she do that? Did she think it would be helpful? Was she just angry or not paying attention?* We spend a great deal of time asking these questions about other people, and most of us seem to believe that we are fairly good at answering them. This process of evaluating the thoughts, beliefs, and actions of those around you is referred to in psychology studies as mind reading (i.e. Eyal & Epley, 2010). Perhaps because we make these kinds of judgements so often, we generally do not give much thought to *how* we make them. What strategies do we use to determine why a person behaved the way she did or what she might have been thinking? While most people may not have given this process much thought, researchers have, discovering that people use more than one mind reading strategy, and the choice of which strategy seems to depend on an initial evaluation of the person whose mind we aim to read.

Research has found support for at least two mind reading strategies: projection and stereotyping. When using projection, a person uses herself as a basis for comparison (Ames, 2004a & 2004b). When determining why a coworker was rude to her, she might consider things she knows about the coworker – upcoming deadlines, family stress – and reason that she herself might be a little short and rude to people under those circumstances. Conversely, stereotyping uses a prominent group that the other person belongs to as a reference point (Fiske, 1998, Fiske & Neuberg, 1990). A person using stereotyping to decide why her coworker was rude might recall that her coworker was born and raised in New York City – where she has heard lots of people are gruff and rude – and reason that this explains her coworker's rudeness. Studies have found support for people using both of these strategies, and later work identified perceived similarity as a key moderator that helps to determine when a person uses one strategy rather than the other. This research finds that the more similar a person perceives themselves to be to the target whose thoughts or behavior they are evaluating, the more likely they are to use projection (Ames, Weber, & Zhou 2012; Ames, 2004a). This makes intuitive sense. If we do not believe that we are similar to another person, then we might not believe that our thoughts, beliefs, and feelings will be similar either, leading use to instead rely on prominent group membership to explain their behavior.

As this research shows, discovering the mind reading strategies people use, and when they use them, is not an isolated legal question. It is, however, an important legal question. One of the key tasks that a juror is asked to perform is to determine the mental state of the defendant

at the time the crime was committed. Jurors are asked, for instance, to determine whether the defendant was in genuine fear for her life when she short the victim or actually wanted to do the victim harm. Considering the primacy of these determinations in the legal system, it is surprising that none of the previous research on mind reading strategies was done with legally relevant stimuli. Many of the findings and principles from the prior research may translate to legal stimuli, but some characteristics about defendants and crimes may change how people utilize these strategies.

For instance, the prior research found that similarity helped to determine which mind reading strategy people use. When the target of mind reading is a person accused of a crime, perceived similarity may be too low to exert an influence. This may cause an overall dampening of the use of projection when evaluating defendants. A further complication is that a juror is not determining the defendant's mental state for an action they personally witnessed, they are evaluating competing explanations – from the prosecution and defense – for an action that has already occurred. This might be different from prior work that relied on participants' judgments of the actions of another actor they could see themselves or predictions about how another would act in the future (Ames, Weber, & Zhou 2012; Ames, 2004a; Ames, 2004b).

This prior work also did not consider what mind reading strategies people might use when evaluating the thoughts and actions of someone who may think in a fundamentally different way, such as a person with a mental illness. Yet, jurors are asked to make mental state evaluations even in such difficult cases, what mind reading strategy do they use? Based on prior research, we might predict that people would perceive themselves to be less similar to a defendant with a mental illness than one without a mental illness, leading to even less projection and, perhaps, more stereotyping. As noted above, however, it is possible that similarity will be so low for any defendant that it will cease to do the work it does in other mind reading contexts. It is also possible that when presented with a defendant with a mental illness, people may have so little relevant stereotype information that they default to using themselves as a basis again. That is, in the face of such uncertainty, people may disregard the person's mental illness entirely and resort to considering what they themselves would do. The prior literature cannot answer these questions. The prior literature also left unaddressed the punitive consequences of these mind reading strategies. If jurors more frequently use one mind reading strategy over another, does that impact how likely they are to find a defendant guilty?

The work detailed in Chapter Four begins to address these issues. The primary research questions were: 1) what mind reading strategy do participants use when judging a defendant with a mental illness? is it different from the one used for a defendant without a mental illness? 2) what, if any, are the punitive consequences of each mind reading strategy? is one more punitive than another? This study set out to determine how people evaluate mental state explanations, and whether different mental state explanations can impact legal decisions, even for identical actions.

Consistent with the prior research, we find that while perceived similarity is low for all defendants, it is significantly lower for defendants with a mental illness, and participants used significantly less projection when evaluating defendants with a mental illness. This was true even though the act the defendant was said to have committed was identical, varying only in whether the defendant is described as having a mental illness or not. We also find that similarity mediates the use of projection, partially in one experiment and fully in another. In these respects, our experiments replicate the prior findings while extending them to specifically legal questions and

stimuli. Participants' mind reading strategies were predicted by perceived similarity, despite similarity being low overall, and despite the different nature of the actions being judged (Giffin & Lombrozo, *under review*). More broadly, we find that different mental state explanations can change the strategy participants use to evaluate the defendant. In our experiments, the different explanations functionally changed perceived similarity, triggering different evaluation strategies.

Further, we find that the different mind reading strategies have different punitive impacts, separate and apart from the punitive impact of having a mental illness. Overall, we found that defendants without a mental illness were significantly more likely than defendants with a mental illness to be found not guilty. A mediation analysis found that levels of projection were a significant mediator of this effect. That is, participants who reported using more projection were more likely to vote not guilty, regardless of whether the defendant they read about was described as having a mental illness. Levels of projection, in turn, rest heavily on a participant's perceived similarity to the actor. When a participant reads about a defendant with a mental illness, they perceive themselves to be less similar to that defendant, leading to lower levels of projection and fewer not guilty verdicts. Figure 2-1 below provides a representation of how mental states are translated into mind reading strategies and then punitive outcomes.
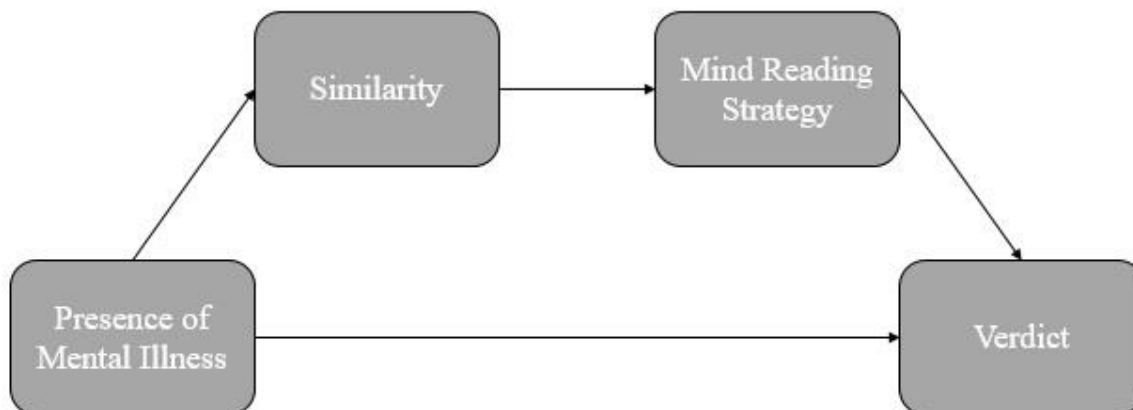
Figure 2-1: Impact of mental illness on mind reading strategy and verdict.

This result has troubling implications for defendants with mental illnesses, but a further experiment suggests that simple steps can be taken to mitigate this effect. We find that reading short, verbal prompts to put themselves in the defendant's place significantly increased reported projection levels for those participants who read about a defendant with a mental illness. That is, a simple prompt, a version of which could be included in opening and closing arguments, caused participants to use more projection, even though they initially perceived the defendant as dissimilar from themselves. The prompt did not have an effect on participants who read about a defendant without a mental illness, we believe because their reported projection levels were already quite high (Giffin & Lombrozo, *under review*). This work shows that what mental state explanation a defendant offers can change the way that explanation is evaluated and, ultimately, the way a defendant is treated. Though more work is needed to understand the mechanisms

behind these findings, this work suggests that some "yes, but…" explanations may be worse than explanation at all.

Outline

The next three chapters contain the complete studies discussed briefly above, each dealing with the impact of mental state explanations. Chapter Two describes a study which finds not only a surprising correspondence between lay opinion and the law on when mental state explanations are mitigating, but also statistical support for a theory that explains this correspondence. Chapter Three's study replicates the main findings of Chapter Two, but expands them into a non-legal context, and finds experimental support for a theory of when mental state explanations are less mitigating. The experiments in Chapter Four examine what strategies people use to evaluate mental state explanations, finding that which explanation a defendant offers changes the way the explanation is evaluated, as well as how the defendant is treated. Each chapter discusses the legal and psychological implications of its findings. These findings together paint a nuanced picture of the relationship between mental state explanations, law, and morality.

Chapter 2

Wrong or Merely Prohibited: Special Treatment of Strict Liability in Intuitive Moral Judgment

Abstract

Most crimes in America require that the defendant have "mens rea," Latin for "guilty mind." However, mens rea is not legally required for strict liability crimes, such as speeding, for which someone is guilty even if ignorant or deceived about her speed. In three experiments involving participants responding to descriptive vignettes, we investigated whether the division of strict liability crimes in the law reflects an aspect of laypeople's intuitive moral cognition. Experiment 1 (N = 396; 236 male, 159 female, 1 other; mean age = 30) found evidence that it does: ignorance and deception were less mitigating for strict liability crimes than for "mens rea" crimes. Experiments 2 (N = 413; 257 male, 154 female, 2 other; mean age = 31) and 3 (N = 404; 183 male, 221 female, mean age 35) revealed that strict liability crimes are not treated as pure moral violations, but additionally as violations of *convention*. We found that for strict liability crimes, ratings of moral wrongness and punishment were influenced to a greater extent by the fact that a rule had been violated, even when harm was kept constant, mirroring the legal distinction of *malum prohibitum* (wrong as prohibited) versus *malum in se* (wrong in itself). Further, we found that rules prohibiting strict liability crimes were judged more arbitrary than corresponding rules for "mens rea" crimes, and that this judgment was related to the role of mental states. Jointly, the findings suggest a surprising correspondence between the law and laypeople's intuitive judgments.

keywords: rule violation; knowledge; decision making; explanations; law; convention

Introduction

"The first requirement of a sound body of law," wrote Justice Oliver Wendall Holmes, "is that it should correspond with the actual feelings and demands of the community, whether right or wrong" (Holmes, 1881). Whether one believes that the law should conform to the feelings of the community, or vice versa, the idea of a correspondence between the legal system and laypeople's intuitive judgments continues to have popular and academic support (e.g., Alicke, 2008). To the extent this correspondence holds, the psychology of moral judgment has important implications for the law, and the law provides a rich source of information about human moral psychology (Hart & Honore, 1962; Mikhail, 2009).

One important correspondence between the law and intuitive judgments can be seen in the legal system's requirements regarding intent. Characteristically, a person cannot be found guilty of a crime in America unless that person had some intent to commit the crime or acted in a way that was negligent or reckless. Legally, this concept is referred to as "mens rea," Latin for "guilty mind," and has been a part of criminal law for centuries (Blackstone 4 Comm., 1769). Psychological research confirms that, in most cases, a person's intent is seen as a crucial variable in assigning moral responsibility and punishment (e.g. Malle & Knobe, 1997; Knobe, 2005). For instance, putting white powder in someone's coffee is judged considerably more harshly when the powder was known to be poison than when it was mistakenly thought to be sugar (Young & Saxe, 2011). Thus, law and psychology agree about the important role of mental states or mens rea, even if legal and intuitive concepts are not perfectly aligned (Malle & Nelson, 2003).

Notably, the law makes an exception concerning the role of mens rea for a specific class of crimes known as "strict liability" crimes. In criminal law, strict liability crimes are crimes for which the prosecution does not have to prove that the defendant had the requisite mens rea with respect to at least one element of a crime, and, in many cases, the defendant is not allowed to present evidence that she made a reasonable mistake or lacked relevant knowledge or intent (Levenson, 1992). For instance, a person who is speeding can be found guilty without the prosecution proving that she intended to speed or knew that she was speeding.

A variety of justifications for the imposition of strict liability have been offered. For example, some have argued that imposing strict liability increases the incentive for the actor to use care or to consider whether the activity should be engaged in at all (Singer, 1989; Manchester, 1977), or that strict liability appropriately apportions risk on the only person who could have benefited from the action (Keating, 2006). These positions are often tied to different justifications for punishment, such as the idea that punishment should *deter* the behavior, rather than being imposed in *retribution* (Craswell & Calfee, 1986; Stahlkopf et al., 2010).

Another explanation for strict liability comes from the Supreme Court case of *U.S. v. Morissette*, in which Justice Jackson noted that, historically, many strict liability crimes were primarily *malum prohibitum* (wrong *because* they were prohibited) rather than *malum in se* (wrong or evil in themselves). A crime that is wrong merely by law, it was argued, does not require mens rea because the violation of the law, regardless of intent, is part of the harm to be protected against (*US v. Morissette*, 1952). We will revisit this idea in Experiments 2 and 3.

In the present work, we investigate whether the legal distinction between strict liability crimes and mens rea crimes reflects a psychological distinction as well. One possibility is that strict liability crimes represent a departure from a general correspondence between the law and intuitive judgments. Indeed, many scholars point out how *un*intuitive the imposition of strict liability can seem (e.g. 113th Congress, 2013; Stanton-Ife, J, 2007). But a second possibility is

that the distinction between strict liability and mens rea crimes reflects a psychological distinction that has not been adequately recognized within the psychological literature. Based on our findings, this is what we will ultimately suggest.

In Experiment 1, our aim is to establish whether judgments concerning strict liability crimes are – on average – influenced less by the actor's knowledge and intent than are mens rea crimes, thus mirroring the law. Importantly, at no point are participants told or given any indication of which crimes are strict liability, nor do they come to the experiments with this knowledge (see online supplement). Nonetheless, Experiment 1 reveals a surprising degree of correspondence between the law's special treatment of strict liability crimes and laypeople's intuitive moral judgments.

In Experiments 1-3, we additionally explore potential bases for differentiation between strict liability and mens rea crimes. Specifically, we investigate whether views about punishment differ for strict liability and mens rea crimes, including the roles of retribution (Experiment 1 and S1, contact authors for full report), deterrence (Experiments 1 and S1), and incapacitation (S2, contact authors for full report). We also investigate whether strict liability and mens rea crimes tend to license different inferences about the actor's mental states (Experiments 1, 2, and S1), and whether relative to mens rea crimes, strict liability crimes are to a greater extent *malum prohibitum* (Experiments 2 and 3). We find support for only one of these proposals: that relative to mens rea crimes, strict liability crimes are treated as more *malum prohibitum*, and that this is because they tend to have more arbitrary elements.

Experiment 1

Experiment 1 investigated whether laypeople judge strict liability and mens rea crimes differently. Research in moral psychology suggests that mens rea is critical to moral judgment (e.g., Cushman & Inbar, 2012; Rai and Fiske, 2011), but there is also evidence that a person's knowledge and intentions are not equally influential in evaluating all types of transgressions (Barrett et al., 2016; Graham et al., 2011; Young & Saxe, 2011). Specifically, knowledge and intentions may be more important for "harm" violations than for "purity" violations, and many – but not all – offenses that have been historically treated as strict liability arguably involve some element of purity (e.g., statutory rape, incest, bigamy, food adulteration). We therefore consider both *public welfare* strict liability offenses that involve potential harm (speeding, selling drugs to minors) as well as *moral-purity* strict liability offenses that violate purity (statutory rape, incest). We compare judgments concerning these crimes against those for closely matched mens rea crimes that we likewise classify as public welfare (reckless driving, drug distribution) or as moral-purity (seducing a minor, first cousin marriage).

In order to investigate the role of mens rea in judging perpetrators of these crimes, we presented participants with vignettes involving a transgression and varied whether the actor knew the fact that rendered an action illegal – e.g., that he was driving over the speed limit, or that a sexual partner was underage. When the actor was ignorant, we additionally varied whether the ignorance resulted from inadvertent bad information or from intentional deception; based on prior work, we expected that deception would be more mitigating (Murray & Lombrozo, 2015; Phillips & Shaw, 2015). Finally, we also included a condition in which mental states were unspecified.

One primary hypothesis (H1) was that judgments concerning moral censure and punishment would depend on the actor's mental state, with the harshest judgments when the

actor transgressed knowingly. We also anticipated that being deceived would be more mitigating than mere ignorance (H1a), and that these effects would be found for all crimes (H1b).

Our secondary hypotheses concerned possible reasons why mens rea and strict liability crimes could receive different treatment. We tested the following predictions: (H2) that the mental state manipulation would influence judgments concerning public welfare crimes more strongly than those concerning moral-purity crimes, (H3) that deterrence would be regarded as a more central reason for punishing strict liability crimes than for punishing mens rea crimes, and (H4) that when an agent's mental states were not specified, participants would more often infer that the transgressor knew the key fact for strict liability crimes than for mens rea crimes.

Methods

Participants. Three-hundred-and-ninety-six adults (236 male, 159 female, 1 other/prefer not to specify, mean age = 30, *SD* = 10) participated in the study through Amazon Mechanical Turk. An additional 140 participants were tested, but were excluded for failing catch questions (92) and due to a data collection error (52).

In all studies, participation was restricted to workers with IP addresses in the United States and with an approval rating of 95% or higher on previous Mechanical Turk tasks. Only U.S. citizens above the age of 18 were allowed to participate, to better mimic the composition of an actual jury. Participants were given thirty cents for their participation.

Materials & Procedure. Participants were randomly assigned to one of thirty-two distinct vignettes, about which they answered several questions (detailed below).

Eight different crimes were chosen, four of which are strict liability crimes in a majority of states (statutory rape, incest, speeding, and selling drugs to minors) and four of which are mens rea crimes (seducing a minor, first cousin marriage, reckless driving, and drug distribution). We chose mens rea crimes that were as similar as possible to the corresponding strict liability crimes to make it more likely that differences, if found, would not reflect idiosyncratic properties of the selected examples. Four of the crimes involved moral purity (statutory rape, seducing a minor, incest, first cousin marriage) and four involved public welfare (speeding, reckless driving, selling drugs to minors, drug distribution).

Table 1-1: Crimes used in Experiment 1. For each crime, participants were assigned to one of four mental state levels: *knowing*, *unspecified*, *unknowing*, or *deceived*. Each crime is followed in parentheses by the relevant knowledge that was varied across the mental state manipulation.

	Strict Liability Crimes	**Mens Rea Crimes**
Moral	Statutory Rape *(Age of sexual partner)*	Seducing a Minor *(Age of person receiving material)*
	Incest *(Relatedness of sexual partner)*	First Cousin Marriage *(Relatedness of spouse)*
Public Welfare	Speeding *(Vehicle speed)*	Reckless Driving *(Speed & attendant circumstances)*
	Drugs to Minors *(Age of customer)*	Drug Distribution *(Nature of substance)*

Each of the eight crimes was presented to a participant in one of four versions, which varied the principal actor's mental states. In the *knowing* condition, participants were told that the actor knew the relevant fact that made a crime illegal (e.g., the age of an underage sexual partner). In the *unspecified* condition, participants were not given any information about the actor's knowledge. In the *unknowing* condition, participants were told that the actor received false but well-intentioned information about the relevant fact (e.g., in the statutory rape vignette, another person honestly believed the underage victim to be an 18-year-old college student and said so). In the *deceived* condition, the actor was deliberately deceived by another person (e.g., the victim lied about her age).

Examples of the different mental state conditions for speeding are excerpted below (see online supplement for all materials):

> *Knowing*: "…. Alan knew the speed limit was 55 miles per hour and that he was driving over the speed limit, but Alan didn't slow down as the speedometer crept up to 70 miles per hour."
> *Unspecified*: "Alan was driving along a two lane highway… A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic…the officer… noticed that Alan was going 70 miles per hour."
> *Unknowing*: "… Alan knew the speed limit was 55 miles per hour and was careful to check the speedometer periodically to make sure the car wasn't going faster than the speed limit…Alan had never had any reason to suspect the speedometer was broken but later found out that it was."
> *Deceived*: "…. Alan knew the speed limit was 55 miles per hour and was careful to check the speedometer periodically to make sure the car wasn't going faster than the speed limit…Alan was surprised to hear [he had been speeding]…Alan had never had any reason to suspect the speedometer was broken but later found out that his mechanic had intentionally tampered with it after Alan disputed a bill."

After reading the assigned vignette, it was removed and participants answered questions. First, participants rated our variables of interest, moral censure and punishment, in a randomized order:

>*Moral censure*: "How morally wrong were [Actor's] actions?" Participants indicated their answer on a scale from 1 (not at all morally wrong) to 7 (very morally wrong).
>*Punishment:* "How much punishment does [Actor] deserve?" Participants indicated their answer on a scale from 1 (none at all) to 7 (very much).

We did not ask participants to evaluate guilt, as in each vignette it was stipulated. Second, to insure that strict liability and mens rea vignettes were well matched along other dimensions that could impact judgments, participants rated the following questions on a separate screen, in randomized order, from 1 (not at all/none) to 7 (very/a great deal):

>*Disgust*: "How disgusting did you find [Actor's] actions?"
>*Anger*: "How angry did [Actor's] actions make you?"
>*Harm:* "How much harm did [Actor's] actions cause?"

Third, to test whether differences in the evaluation of strict liability and mens rea crimes stem from commitments about why punishment is appropriate for each crime type (H3), participants rated the statements below in a randomized order, from 1 (completely agree) to 7 (completely disagree):

>*Deterrence*: "Punishing people who [commit this crime] is critically important to preventing acts of [this crime] in society in general."
>*Retribution*: "Even if it won't prevent future [instances of this crime], a person who [commits this crime] deserves to be punished for this action."

Fourth, to insure that participants were attending to the task and reading materials carefully, participants answered two "catch questions," easy true/false questions that assessed having read the vignette. Participants who answered either question incorrectly were excluded from further analyses. This screen included a third question about the actor's knowledge (e.g., "whether "Alan knew he was driving over the speed limit"), which served as a mere comprehension question for participants in the *knowing*, *unknowing*, and *deceived* conditions, where knowledge or ignorance of the vital fact – speed of the car, etc. – was stipulated. It also allowed us to test for differences in baseline inferences across crime types for the *unspecified* condition (H4).

Finally, participants justified their reaction to the story, received another attention check based on Oppenheimer, Meyvis, and Davidenko (2009), and answered demographic questions. (Neither the justifications nor the demographic information were analyze for this paper.)

Results

Disgust, anger, and harm. To ensure that our strict liability and mens rea vignettes did not differ systematically in judged harm or elicited anger and disgust, we performed a series of one-way ANOVAs with legal category (strict liability, mens rea) as a between-subjects variable and with each rating as a dependent variable. No significant differences were found for harm,

$F(1,394)= 0.20$, $p < .66$, $d = .04$, 95% CI [-.296, .469], anger, $F(1,394)=1.74$, $p < .19$, $d = .13$, 95% CI [-.126, .635], or disgust, $F(1,394)= 1.35$ $p < .25$, $d = .12$, 95% CI [-.169, .658], suggesting that the effects of legal category we report below are unlikely to result from these potential confounds.

Moral censure and punishment. Do mental states play different roles in intuitive judgments concerning strict liability and mens rea crimes? And if so, does the relationship hold across moral-purity and public welfare crimes? To address these questions we performed a series of 2 (legal category: strict liability, mens rea) x 2 (moral classification: moral, public welfare) x 4 (mental state: *knowing*, *unspecified*, *unknowing*, *deceived*) ANOVAs for moral censure and punishment.

These analyses revealed a significant main effect of mental state for both moral censure, $F(3,380) = 29.34$, $p <.000$, $\eta p^2 = .188$, and punishment, $F(3,380) = 20.52$, $p <.000$, $\eta p^2 = .139$ (see Figure 1-2). Independent sample t-tests revealed that ratings for each of these dependent variables were significantly higher in the *knowing* condition than in any of the other three conditions ($ps < .003$, $ds > .43$), consistent with most previous work on the effects of intent on moral judgments, and supporting hypothesis H1. We also found that ratings for the *unspecified* condition were significantly higher than for the *unknowing* and *deceived* conditions ($ps < .09$, $ds > .38$), but that ratings for the *unknowing* and *deceived* conditions were not significantly different from each other for either dependent variable ($ps > .30$, $ds < .15$), challenging hypothesis H1a. This suggests that misinformation had a mitigating effect, but that the source of the misinformation was not relevant.

Figure 1-2: Effect of mental state on moral censure and punishment. Higher values correspond to greater amounts of moral censure and punishment. Error bars correspond to one SEM in each direction.
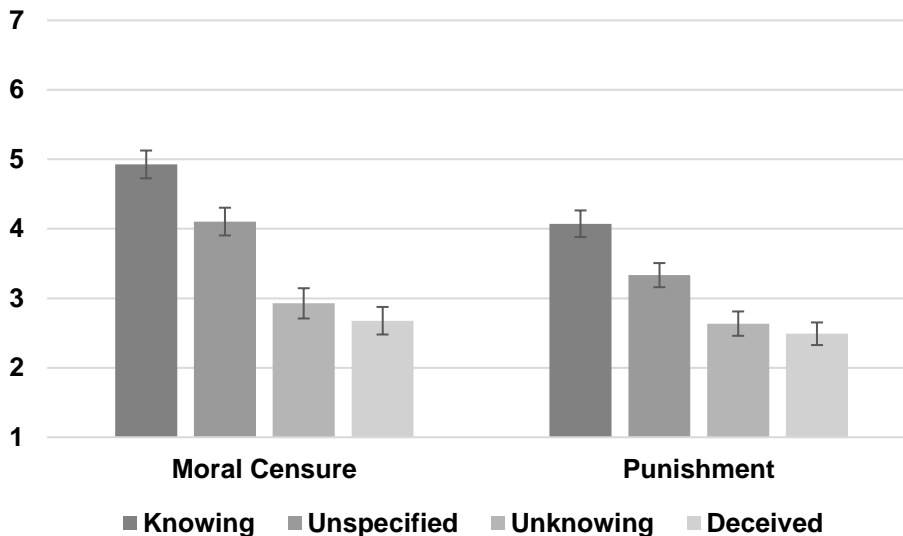
We also found significant effects of legal category for both moral censure and punishment, with higher ratings for strict liability crimes (Mc = 3.87, SDc = 2.20; Mp = 3.38,

SDp = 1.81) than for mens rea crimes (Mc = 3.44, SDc = 2.16; Mp = 2.89, SDp = 1.84) (see Table 2-2).

Experiment	Dependent Measure	Knowing	Unspecified	Deceived	Unknowing
	Moral Censure				
	Strict Liability	4.78 (1.96)	4.50 (1.78)	2.82 (2.12)	3.44 (2.34)
	Mens Rea	5.14 (1.93)	3.71 (2.06)	2.51 (1.67)	2.43 (1.80)
Exp. 1	Punishment				
	Strict Liability	3.96 (1.64)	3.60 (1.55)	2.71 (1.87)	3.28 (1.94)
	Mens Rea	4.27 (2.07)	3.06 (1.82)	2.16 (1.18)	2.10 (1.30)
	Moral Censure				
	Strict Liability	4.04 (2.00)	4.04 (1.83)	3.50 (2.08)	3.71 (2.18)
	Mens Rea	6.02 (1.30)	4.50 (1.81)	3.04 (2.20)	3.54 (2.12)
Exp. S1	Punishment				
	Strict Liability	4.29 (1.66)	4.21 (1.76)	3.44 (1.90)	4.15 (1.74)
	Mens Rea	4.92 (1.57)	3.69 (1.53)	2.58 (1.67)	3.21 (1.88)

Table 2-2: Means as a function of legal category. Standard deviations in parenthesis.

We hesitate to draw strong conclusions from these differences, as they could simply reflect our choice of stimulus materials rather than the classes of crimes. Moreover, our primary predictions concern an *interaction* between mental state and legal category, and this main effect was indeed qualified by an interaction for punishment, $F(3, 380) = 3.40$, $p <.018$, $\eta p^2 = .026$: ignorance of the vital fact was more mitigating for mens rea crimes than for strict liability crimes. To confirm this interpretation statistically, we repeated our analysis with only the *knowing* and *unknowing* conditions (see Figure 2-2), which revealed that for mens rea crimes, punishment ratings were significantly higher for the *knowing* condition than for the *unknowing* condition, $t(80) = 6.24$, $p < .000$, $d = 1.40$, 95% CI [1.48, 2.86] (corrected for violating Levene's), but that for strict liability crimes, there was no significant difference between the two conditions, $t(98) = 1.89$, $p < .061$, $d = .38$, 95% CI [-.033, 1.39]. The corresponding interaction for moral censure was not significant, $F(3, 380) = 2.33$, $p <.074$, $\eta p^2 = .018$, but trended in the same direction. These findings challenge prediction H1b: contrary to our initial expectations, and contrary to common assumptions in legal scholarship, our data support an intuitive basis for the special treatment of strict liability crimes. Moreover, this pattern was robust across crime pairs: for each of our four pairs, and for both moral censure and punishment, the difference between the *knowing* and *unknowing* conditions was greater for the mens rea crime than for the strict liability

crime, with a single exception: for moral censure, the difference between ratings in the *knowing* and *unknowing* conditions was the same for statutory rape as for seducing a minor.

Figure 2-2: Interaction between mental state and legal category for moral censure (not significant) and punishment (significant). Error bars correspond to one SEM in each direction.
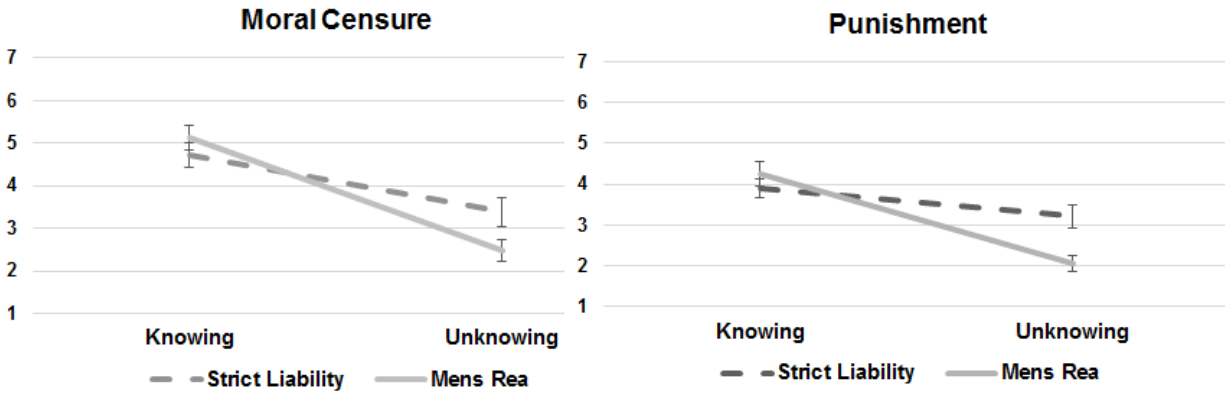
Finally, for moral censure, there was a two-way interaction between mental state and moral classification, $F(3, 380) = 3.12$, $p < .026$, $\eta p^2 = .024$: for moral-purity crimes, moral censure did not differ across the *knowing* and *unspecified* conditions, $t(94) = .548$, $p < .585$, $d = .11$, 95% CI [-.546, .963], nor across the *unknowing* and *deceived* conditions, $t(98) = .548$, $p < .548$, $d = .11$, 95% CI [-.853, .455], but was significantly higher for both of the former conditions than for either of the latter conditions. In contrast, for public welfare crimes, moral censure was significantly higher for the *knowing* condition than all others, $p < .000$, $d > .75$, which did not differ from each other, ps ranged from .09 to .53, ds ranged from .13 to .35.

Bases for punishment: deterrence and retribution. To test whether punishment was justified differently across legal categories, deterrence and retribution ratings were each subjected to a 2 (legal category: strict liability, mens rea) x 4 (mental state: *knowing*, *unspecified*, *unknowing*, *deceived*) ANOVA. For deterrence, this analysis revealed a main effect of legal category, $F(1, 388) = 9.50$, $p = .002$, $\eta p^2 = .24$, with higher ratings for mens rea crimes ($M = 4.82$, $SD = 1.88$) than for strict liability crimes ($M = 4.22$, $SD = 1.95$), contrary to hypothesis H3 and what one might expect on the basis of legal scholarship, which sometimes justifies the imposition of strict liability on utilitarian grounds related to deterrence. This surprising result is unlikely to be a simple artifact of our stimulus materials: recall that moral censure and punishment were both *lower* for mens rea crimes than for strict liability crimes, and we did not find main effects for anger, disgust, or harm across legal category. No significant results were found for retribution (p's > .215; ηp^2's < .008).

Inferring knowledge and intent. Responses to the true/false knowledge question were checked to confirm that participants correctly responded "true" for the *knowing* condition and "false" for the *unknowing* and *deceived* conditions. Accuracy was high, ranging from 93% for the *deceived* condition to 95% for the *knowing* condition and 99% for the *unknowing* condition.

For the *unspecified* condition, we were interested in whether assumptions about knowledge would differ for mens rea and strict liability crimes. We therefore conducted a chi-squared test comparing rates of inferred knowledge. This analysis revealed no difference, challenging prediction H4: in both cases, participants inferred knowledge 54% of the time.

Discussion

In Experiment 1, we sought to establish whether moral and punitive judgments concerning strict liability crimes reflect their treatment by the legal system, according to which knowledge is irrelevant. While we found that mental states did have a significant impact on judgments (supporting H1), with ignorance having a mitigating effect (regardless of its source, challenging H1a), the effect of mental states was not uniform across legal category. Contrary to our prediction (H1b), participants did differentiate between strict liability and mens rea crimes: ignorance was significantly less mitigating for strict liability crimes than for closely matched "mens rea" crimes.

We also tested, but failed to find support for, three hypotheses about why mental states might have different effects across legal categories. First, we considered the possibility (H2) that mental states would be less important for intuitive judgments of moral-purity crimes relative to public welfare crimes, consistent with both psychological data (Barrett et al., 2016; Russell & Giner-Sorolla, 2011) and the historical association between strict liability and purity (*US v. Morissette*, 1952). While we did find that mental state interacted with moral classification – with the *unspecified* conditions patterning differently – we did not find the anticipated difference between the *knowing* and *unknowing* conditions, nor an interaction with legal category. It thus appears that violating purity is neither necessary nor sufficient to explain the attenuated role for mental states observed in strict liability cases.

Second, we tested the idea (H3) that for strict liability crimes, punishment may be warranted for reasons of deterrence, if not retribution. However, we found no global differences in retribution ratings, and *lower* deterrence ratings for strict liability crimes than for mens rea crimes. This is surprising given that punishment ratings were *higher* for strict liability crimes than for mens rea crimes, even though crime types did not differ in perceived harm. This is not to say that deterrence might not still furnish some justification for the imposition of strict liability, but our data do suggest that for the crimes tested, laypeople do not associate punishment with deterrence more strongly for strict liability crimes than for mens rea crimes.

Finally, we also failed to find support for the idea (H4) that inferences about mental states might differ across legal category, with mental states being more likely to be inferred (or perhaps more difficult to ascertain) for strict liability: in our *unspecified* condition, rates for inferred knowledge did not differ across strict liability and mens rea crimes.

In sum, the results of Experiment 1 suggest that laypeople's intuitive moral judgments mirror the law in reserving strict liability crimes for special treatment, but don't shed light on why this might be. Before moving on to Experiments 2 and 3, in which we find support for one possibility, it's worth addressing some open questions from Experiment 1, which we investigated in supplementary experiments.

First, one concern is that the results from Experiment 1 could be driven by the particular crimes that we happened to test. Moreover, by selecting crimes that were so closely matched *across* legal categories (e.g., speeding paired with reckless driving), we could have inadvertently selected crimes that are not representative of strict liability or mens rea crimes in general. To address this possibility, we ran an additional experiment which we refer to as Experiment S1

(contact authors for a full report). In Experiment S1, we selected strict liability and mens rea crimes based on how frequently they are charged, as represented by FBI crime statistics (FBI, 2012) and consultation with attorneys. The strict liability crimes selected were statutory rape, selling drugs within 1,500 feet of a school, speeding, and driving under the influence. The mens rea crimes selected were sexual battery, burglary, theft, and (non-sexual) battery. Other than crime selection, the methods and procedure mirrored those of Experiment 1. Our key findings from Experiment 1 were replicated with these new crimes: ignorance significantly mitigated moral censure and punishment ratings for mens rea crimes, but had no significant effect for strict liability crimes. Ratings for deterrence were again higher for mens rea crimes.

A second concern is that the judgments concerning punishment in Experiment 1 may have been insufficiently fine-grained. In supplementary Experiment S2 (contact authors for full report), participants separately specified the appropriate levels for *fines* and for *jail time*. We also included a third possible basis for punishment: incapacitation. S2 supported the interpretation of Experiment 1: we found the same pattern of responses concerning the role of mental states across legal categories, and we found that incapacitation was rated a significantly better basis for punishing mens rea crimes than strict liability crimes.

In conjunction with the findings from Experiment 1, these supplementary experiments provide support for the idea that people mirror the law in their treatment of mens rea versus strict liability crimes, but the findings do not (yet) reveal the basis for this special treatment.

Experiment 2

Experiment 1 found support for the idea that laypeople's intuitive judgments mirror the law in the following way: for mens rea crimes, judgments are sensitive to differences in mental states (*knowing* versus *unknowing*), while for strict liability crimes, judgments are less sensitive (or even indifferent). The data also revealed high levels of punishment assigned to strict liability offenders, despite the fact that strict liability offenses were judged no worse than mens rea crimes (in terms of overall harm or elicited anger and disgust), and despite the fact that three common justifications for punishment – retribution, deterrence, and incapacitation (tested in S2) – were not endorsed more strongly for strict liability crimes than for mens rea crimes.

Experiment 2 considers an explanation for these results with roots in the Supreme Court case of *United States v. Morissette*. In *Morissette*, Justice Jackson noted that historically many strict liability crimes were regulatory in nature, with the aim of improving public safety and welfare. He further suggested that strict liability crimes were therefore primarily *malum prohibitum* rather than *malum in se*, meaning that transgressions were considered wrong *because they were prohibited* (i.e., violated a regulation), not wrong or evil *in themselves*. For instance, while it isn't intrinsically wrong to drive a car at 50 miles per hour, it is wrong to do so in a 25-mile-per-hour zone. To quote Justice Jackson's Supreme Court Opinion:

> "While such offenses do not threaten the security of the state in the manner of treason, they may be regarded as offenses against its authority, for their occurrence impairs the efficiency of controls deemed essential to the social order as presently constituted. In this respect, whatever the intent of the violator, the injury is the same.... Hence, legislation applicable to such offenses…does not specify intent as a necessary element" (Morissette v. United States, 1952).

The legal distinction between *malum prohibitim* and *malum in se* mirrors a cognitive distinction explored in developmental research. Studies have found that children as young as six are able to distinguish between actions that are "rule contingent," or wrong because an authority figure or rule says they are wrong, and actions that are *intrinsically* wrong, and would therefore be wrong even if no rule or authority prohibited them (Nucci & Turiel 2009, Turiel 2008). For example, if there is a rule against wearing pajamas to school, then doing so might be wrong and merit detention, no matter that the action would be perfectly acceptable on pajama day. Similarly, it could be that participants judged our strict liability crimes wrong and deserving of punishment in large part because they violated rules, even though they did not regard the actions themselves as more harmful, angering, or disgusting than the mens rea crimes, and even though retribution, deterrence, and incapacitation were rated more appropriate reasons for punishing commissions of mens rea crimes.

To test these ideas, participants evaluated vignettes written in the style of our *unspecified* condition from Experiment 1. Then, subsequent to reading and evaluating the original vignettes, participants were told that the relevant statute had been changed. For example, participants in the statutory rape vignette were told, after initial judgments, that the age of consent had been lowered to 15, meaning that the action described in the vignette was no longer a crime. Participants then provided a second set of ratings for moral censure and punishment, allowing us to assess how much the change in statute affected these judgments, and therefore how "rule contingent" (or *malum prohibitum*) they regarded the original wrong.

Finally, we also added a question about whether the offender had a responsibility to acquire the knowledge relevant to the crime (e.g., the speed at which he was driving). If people believe that strict liability crimes have higher demands in this regard, then ignorance may be treated as a form of negligence, deserving of punishment despite an absence of bad intent.

In sum, Experiment 2 was designed to test two hypotheses. The central hypothesis (H5) was that punishment and moral censure ratings would decrease after participants learned of the statute change (such that the action was no longer prohibited), and that this decrease would be greater for strict liability crimes than for mens rea crimes. A secondary hypothesis (H6) was that participants would believe offenders had a greater responsibility to avoid ignorance for strict liability crimes than for mens rea crimes.

Methods

Participants. Four-hundred-thirteen adults (257 male, 154 female, 2 other/prefer not to specify, mean age = 31, SD = 10) participated in the study through Amazon Mechanical Turk. An additional 77 participants were tested, but were excluded for failing catch questions. Only US participants who indicated they were older than 18 were allowed to participate, to approximate a real jury pool. Participants were given fifty cents for their participation.

Materials & Procedures. In an effort to view strict liability and mens rea crimes broadly, we included all the crimes used in Experiments 1 and S1 and added three new strict liability and three new mens rea crimes. The new strict liability crimes were illegal dumping, possession of an unregistered firearm, and hunting a migratory bird. The new mens rea crimes were disturbing the peace, public drunkenness, and minor in possession of alcohol (see Table 3-2). As in Experiment S1, these new crimes were chosen as the next most frequently charged

according to FBI statistics and consultation with attorneys. The resulting design had 20 conditions, one for each crime.

Table 3-2: Crimes used in Experiments 2 and 3. Each crime is followed in parentheses by the relevant knowledge that was varied across the mental state manipulation in Experiment 3.

Strict Liability	Mens Rea
Illegal Dumping	Disturbing the Peace
(Distance from waterway)	*(Unlawful to fight in public)*
Unregistered Firearm	Public Drunkenness
(Registration status of firearm)	*(Intoxication level prevented taking care)*
Hunting Migratory Bird	Minor in Possession
(Bird's protected status)	*(Beverage contained alcohol)*
Selling Drugs within 1,500 Feet	Sexual Battery
(Distance from nearest school	*(Consent for the touching)*
DUI	Burglary
(Level of intoxication)	*(Owner's intent to keep the vase)*
	Theft
	(whether saw had been discarded)
	Battery
	(Whether friend was genuinely in danger)

After reading a vignette describing a crime, participants were asked the moral censure, disgust, harm, and anger questions used in Experiment 1. Participants were also asked two punishment questions: fines and jail time:

> *Jail Time*: "How much jail time should [Actor] receive?" Participants indicated their answers on a continuous slider that ranged from 0 to twice the maximum penalty that could actually be assigned for the most harshly punished crime, assuming a first offense. The maximum value for jail time was 36 years.
> *Fines*: "How much should [Actor] be fined?" Participants indicated their answers on a continuous slider that ranged from 0 to twice the maximum penalty that could actually be assigned for the most harshly punished crime, assuming a first offense. The maximum value for fines was $100,000.

We did not ask the deterrence, retribution, or incapacitation questions in this experiment.

We added two new dependent variables about the actor's knowledge and responsibility to acquire that knowledge. Examples from the speeding story are below:

> *Inference*: "Did Alan know that he was driving over the speed limit?" Rated from 1 (no, definitely not) to 7 (yes, definitely).
> *Responsibility*: "Did Alan have a responsibility to check whether he was going over the speed limit?" Rated from 1 (no, definitely not) to 7 (yes, definitely).

After participants answered these questions, we presented them with a statute change that stated that the law had been changed such that the actor's action was no longer illegal. In each case, we were clear that the statutory change was due to a technicality, not due to any change in the harm caused by the action. The participants then provided moral censure and punishment ratings in light of the new statutory information. An example from the migratory bird story is below:

> "After Alan's arrest, the officer was informed that the statute had been changed. Due to a technicality (not to a change in the bird species' status), the bird Alan shot was no longer on the list of protected birds. Alan's actions were therefore no longer technically a crime. Please answer the following questions based on the change in statute."

One concern with including this statute change as a within-subjects factor is that participants' initial ratings (before the statute change) could anchor or otherwise influence their second ratings (after the statute change). We were not overly worried by this possibility because our key hypothesis (H6) concerned a *differential* drop in ratings across strict liability versus mens rea crimes, not the values of the ratings themselves. But if anything, this methodological choice works against our hypotheses.

Results

Moral censure and punishment: ratings under statutory change. To test hypothesis H5, that the wrong associated with strict liability crimes is more contingent on the presence of a rule than that of mens rea crimes, we created difference scores for moral censure, fines, and jail time by subtracting each participant's rating after the statute change from that participant's corresponding initial rating. An independent samples t-test on moral censure difference scores revealed a significant difference in the predicted direction, $t(411) = 3.16$, $p < .002$, $d = .31$, 95% CI [.155, .665]: ratings dropped an average of .914 points (SD = 1.41) on a 7-point scale for strict liability crimes, but only .504 points (SD = 1.23) for mens rea crimes.

To analyze punishment ratings, we transformed fines and jail time differences into z-scores and computed a repeated measures ANOVA with punishment type as a within-subjects factor and legal category as a between subjects factor. While the main effect of legal category was not significant, $F(1,398) = 2.26$, $p = .103$, $\eta p^2 = .007$, there was a nearly significant interaction between punishment type and legal category, $F(1,398) = 3.70$, $p = .055$, $\eta p^2 = .009$. For fines only, the predicted effect was significant, $t(303) = 2.36$, $p < .019$, $d = .28$, 95% CI [449.9, 4969] (corrected for violating Levene's): for strict liability crimes, there was a fine drop of $4,844 (SD =$13,595), versus only $2,134 (SD = $8,637). Jail time also dropped a larger amount for strict liability crimes ($M = 0.364$, $SD = 1.09$) than for mens rea crimes ($M = 0.339$, $SD = 1.35$), but not significantly, $t(411) = .198$, $p < .843$, $d = .03$, 95% CI [-.217, .266].

Inferred knowledge and responsibility. Responses to the question about the actor's responsibility for obtaining reliable information differed significantly between strict liability ($M = 5.23$, $SD = 2.04$) and mens rea crimes ($M = 5.61$, $SD = 1.83$) crimes, $t(378) = 1.98$, $p < .049$, $d = .20$, 95% CI [-.759, -.002] (corrected for violating Levene's), with higher responsibility in the mens rea case. If anything, this goes counter to our expectation (H6) that people might be held

more responsible in strict liability cases, rendering ignorance akin to a form of negligence deserving punishment.

We also found that participants were significantly more likely to believe that the principal actor possessed the relevant knowledge for mens rea crimes ($M = 3.74$, $SD = 2.04$) than for strict liability crimes ($M = 2.61$, $SD = 1.62$), $t(410) = 6.27$, $p < .000$, $d = .62$, 95% CI [-1.48, .774]. This differs from what we found in Experiment 1 in relation to hypothesis H4; in that experiment we found no differences in inferred knowledge across mens rea and strict liability crimes. It is possible that the difference revealed here is the result of a more sensitive, continuous measure, or to the inclusion of additional crimes. Whatever the cause, it raises the question of whether the greater drop in ratings for strict liability crimes than for mens rea crimes after the statutory change was a consequence of differences in inferred knowledge, or an independent effect. To test this, we computed partial correlations between legal category and each of our three difference scores (moral censure, fines, and jail time), controlling for inferred knowledge. The relationship between legal category and moral censure remained significant, $r = -.165$, $p < .001$, as did that for legal category and fines, $r = -.126$, $p = .012$.

Analyses across crimes. Not surprisingly, there was considerable variation across crimes for all of our measures (see Appendix A for additional data by crime). Figure 3-2 plots the moral censure difference scores for each crime. The figure makes it clear that while the means for mens rea and strict liability crimes differed, their distributions were overlapping. In the general discussion we return to some of the outlying cases.

Figure 3-2: Moral censure difference scores (rating before statute change minus rating after statute change) by crime in Experiment 2.
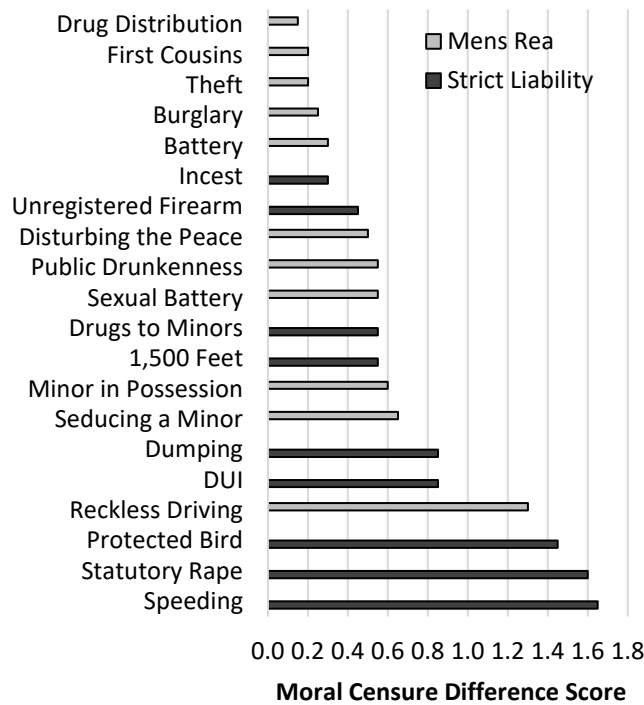
Discussion

Experiment 2 had several aims. Most importantly, we found support for one account of what differentiates strict liability crimes from mens rea crimes. For the crimes we tested, the mere act of violating a prohibition played a larger role in judgments of moral censure and punishment for strict liability crimes than for mens rea crimes. Specifically, a change in statute led to a significantly greater decrease in moral censure and fines for strict liability crimes than for mens rea crimes, supporting hypothesis H5. This is consistent with the idea that strict liability crimes are wrong to a greater extent *because* they are prohibited; they are more *malum prohibitum* than mens rea crimes, which are, to a greater extent, *malum in se*.

Experiment 2 also found a difference in the knowledge imputed to offenders of different crimes: the modal response for strict liability crimes was to infer an absence of knowledge; the modal response for mens rea crimes was the midpoint of our scale. This difference does not appear to have driven the greater "rule contingency" of strict liability crimes. However, it may nonetheless be related. To the extent that strict liability crimes hinge on the precise specification of a rule (e.g., whether a bird or substance is on an official list), ignorance may be plausible, yet still fail to mitigate: because the crimes are *malum prohibitum*, the guilty mind is less relevant.

Experiment 3

Experiment 1 found that ignorance was less mitigating for strict liability crimes than for mens rea crimes, but failed to identify a potential basis for this differential effect. In Experiment 2, we found support for one relevant difference between strict liability and mens rea crimes: on average, the former were treated as more "rule contingent" in the sense that censure and punishment dropped to a greater extent after a rule change. This is consistent with the suggestion that strict liability crimes are somewhat akin to conventional violations, or *malum prohibitum*. A remaining question, however, is *why* strict liability are treated in this way. The historical examples cited by Justice Jackson in *United States v. Morissette* were largely public safety regulations, but many contemporary strict liability crimes address other domains of human behavior. What is it, then, about strict liability crimes that renders their evaluation more contingent on the presence of a rule, and less contingent on mens rea?

While not all contemporary strict liability crimes involve public safety or welfare regulations, many seem to involve a somewhat *arbitrary* element. Returning to an earlier example, it is not in itself wrong or illegal to drive a car 50 miles per hour, but it is illegal to do so in a 25-mile-per-hour zone. But why is the zone's speed limit designated as *25* miles per hour, and not 24 or 26? Similarly, it is not in itself wrong or illegal to have sex, but in our country it is illegal for an adult to do so with a person who is below a certain age. But why is that age 18 in some states, and 16 in others? Of course, there are very good reasons why driving too quickly or having sex with a child could be considered wrong in themselves – the rules are not merely a matter of convention. But the specific line that is drawn between legal and illegal actions – the specific age, speed, distance, etc., specified in a prohibition – is somewhat arbitrary and may therefore render strict liability crimes a hybrid of *malum in se* and *malum prohibitum*, whereas most mens rea crimes are more cleanly *malum in se*.

To test this idea, Experiment 3 investigated whether participants consider strict liability crimes to be more arbitrary than mens rea crimes. Our first prediction (H7) was that, on average, the rules corresponding to strict liability crimes would be judged more arbitrary than those corresponding to mens rea crimes. We also obtained a measure reflecting the role of mental

states for each crime by presenting *unknowing* and *knowing* conditions to each participant. We predicted that the effect of mental states would be smaller for strict liability crimes than for mens rea crimes. We also predicted an association between arbitrariness ratings and the magnitude of the effect of mental state (H8): the more arbitrary the rule, the smaller the influence of mens rea.

Methods

      Participants. Four-hundred-and-four adults (183 male, 221 female, mean age = 35, SD = 12) participated in the study through Amazon Mechanical Turk. An additional 80 participants were tested, but were excluded for failing catch questions. Only US participants who indicated they were older than 18 were allowed to participate, to approximate a real jury pool. Participants were given sixty-five cents for their participation.

      Materials & Procedures. Participants were randomly assigned to one of twenty conditions, representing each of the 20 crimes from Experiment 2. Each participant initially read the *unknowing* vignette of the assigned crime. Most of the *unknowing* version had been used in previous experiments, but for the crimes added in Experiment 2, new *unknowing* versions were created (see online supplement for the six new vignettes). After reading their *unknowing* vignette, participants were asked the moral censure, fine, and jail time questions used in Experiment 2.

      Next, participants were presented with a knowledge change. Participants were asked to imagine that the actor had in fact known the vital fact, and they then provided the moral censure, fine, and jail ratings again. This manipulation allowed us to obtain a within-subjects measure of the effect of mental states. An example from the migratory bird vignette is presented below:

> "Suppose that Alan had actually realized, when he shot the bird, that his actions broke the law – that is, that he was shooting a protected bird. And suppose that he decided to shoot the bird anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)"

      Next, all participants, regardless of which crime they initially evaluated, rated the arbitrariness of all twenty crimes. The crimes were preceded by the following instructions:

> "Below you will see a list of laws. We'd like you to give your intuitions about how arbitrary each law seems to you. That is, do you believe that there's a good reason for the law to draw the line where it does in terms of which actions are legal versus illegal? Or does it seem like the law is somewhat arbitrary in the sense that it could reasonably have drawn the line differently?
> Please do not consult any outside resources, like other people or websites. We are interested in your own intuitions, whether or not they correspond to the current legal system. Even if you think all laws are somewhat arbitrary or not at all arbitrary, please take note of which seem more or less arbitrary and respond accordingly."

      The twenty crimes were then presented in a randomized order. Each crime was presented as true to its statutory language as possible, with minor adjustments in wording to facilitate

comprehension. Providing the full statutory language insured that participants were not operating with different, common sense definitions each crime. Below is the text for migratory bird:

> "Migratory Birds: it is illegal for anyone to pursue, hunt, take, capture, or kill any migratory bird, or any part, nest, or egg of any such bird." Participants indicated their agreement on a scale of 1 (not at all arbitrary) to 7 (completely arbitrary).

Finally, participants answered the demographic and catch questions used in previous experiments.

Results

Arbitrariness ratings. A primary aim of Experiment 3 was to test whether strict liability crimes are considered more arbitrary than mens rea crimes (H7). To do so, we computed the average arbitrariness rating for strict liability crimes and for mens rea crimes for each participant. We then ran a paired-samples t-test on mean arbitrariness ratings as a function of legal category. Consistent with our prediction, participants rated strict liability crimes significantly more arbitrary (M = 2.99, SD = 1.61) than mens rea crimes (M = 2.93, SD = 1.59), $t(403) = 1.87$, $p <$ .031, $d = .04$, 95% CI [-.122, .003] (one-tailed test). (Mean arbitrariness ratings for each crime can be found in Appendix B.)

Moral censure and punishment: ratings under mental state change. This experiment was the first in which we manipulated mental states within subjects. Thus, we were able to create difference scores for each participant for moral censure, fines, and jail time by subtracting each participant's initial, *unknowing* ratings from the rating given after participants were asked to assume that the perpetrator acted knowingly. An independent samples t-test comparing moral censure difference scores as a function of legal category failed to find a significant effect, $t(402) = .126$, $p < .899$, $d = .01$, 95% CI [-.356, 406]. This departure from our previous findings could be due to the within-subjects design or to the addition of six crimes that were not included in Experiment 1 or in the supplementary experiments, and which were much less familiar than those used in our initial studies. We ran the same independent samples t-test on difference scores for the original 14 crimes from Experiments 1 and S1 and found a significantly greater effect of mental state for mens rea than for strict liability crimes on moral censure, $t(273) = 4.08$, $p < .000$, $d = .49$, 95% CI [.475, 1.36] (corrected for violating Levene's), replicating the original effect.

For punishment, however, we did find the predicted effect when considering all 20 crimes. A repeated measures ANOVA with legal category as a between subjects factor, and z-scores of difference scores for fines and jail time as a within subjects fact, revealed a significant effect of legal category, $F(1,398) = 3.88$, $p = .050$, $\eta p^2 = .008$.

Correlation between mental state effects and arbitrariness of assigned story. Finally, we tested whether the arbitrariness of a crime would relate to the magnitude of a mental state effect for that crime. If part of what differentiates strict liability crimes from mens rea crimes is their more conventional or arbitrary nature, and this difference is partially responsible for the relative unimportance of mens rea, then we would expect arbitrariness ratings to be negatively correlated with the magnitude of a mental state effect. We ran a bivariate correlation between each participant's *knowing-unknowing* difference score and her arbitrariness rating for the

specific crime she read. Consistent with our prediction (H8), we found that moral censure scores were significantly negatively related to arbitrariness, r (404) = -.13, $p < .010$. We found no significant correlations with either punishment measure.

Discussion

Experiment 3 confirmed our prediction (H7) that strict liability crimes are considered more arbitrary than mens rea crimes. This prediction was motivated in part by arguments developed in *Morissette* and the findings from Experiment 2: if an action is wrong merely because it is prohibited, then the prohibition must not be grounded in intrinsic harm or other moral considerations. To be clear, our contention is not that strict liability crimes are not wrong in themselves, but merely that many also have a conventional element, and that the role for such arbitrary or conventional elements is greater, on average, for strict liability crimes than for mens rea crimes.

We also found a significant, though small, correlation between participants' arbitrariness ratings and the extent to which their judgments varied across knowing and unknowing violations. Specifically, the perpetrator's mental states (mens rea) had a greater impact on ratings the less arbitrary the rule. We return to this relationship in the general discussion.

<center>**General Discussion**</center>

In the work reported here, we set out to test whether the legal distinction between strict liability and mens rea crimes reflects (and potentially reveals) aspects of intuitive moral judgment. Here we summarize and discuss our key findings.

Experiment 1 revealed that judgments concerning mens rea crimes were in fact more sensitive to the actor's knowledge and intent than were those concerning strict liability crimes. When an actor transgressed *knowingly*, punishment and censure were greater for mens rea crimes than for strict liability crimes, but when the actor transgressed *without knowledge and intent* (whether merely ignorant or deceived), this pattern was reversed. As a result, judgments for mens rea crimes were strongly influenced by the actor's knowledge and intent, whereas those for strict liability crimes were about the same. Put differently, ignorance was treated as a mitigating factor for mens rea crimes, but not for strict liability crimes.

In Experiments 2 and 3, we found support for a possible explanation: that strict liability crimes have more features that are *malum prohibitum,* or wrong because prohibited, while *mens rea crimes* are largely *malum in se*, or wrong in themselves. Experiment 2 found that judgments for strict liability crime were more dependent on the presence of a rule: when a statute was revoked, ratings for moral censure and punishment decreased more for strict liability crime than for mens rea crimes. These findings have analogues in social and cognitive development, where research differentiates between moral and "conventional" wrongdoing (Nucci & Turiel, 2009; Turiel, 2008). Experiment 3 focused on another aspect of conventionality: the potentially arbitrary nature of the rule involved. We found that rules prohibiting strict liability crimes were on average judged more arbitrary than rules for mens rea crimes, and that the more arbitrary the rule, the more attenuated the role of mental states.

The relationship between strict liability and arbitrariness can help explain why strict liability crimes involve an attenuated role for mental states. First, the more arbitrary nature of strict liability crimes means that the intentions associated with a knowing violation are not necessarily *intrinsically* bad. For instance, intentionally driving at 55 miles per hour is not itself

wrong; it only becomes so in the context of a law that specifies an area as a 50-mile per hour zone. When a law is *not* arbitrary, by contrast, the intention is reprehensible even absent a law: intentionally hitting another person is unacceptable whether or not a law is in place. Relatedly, crimes that involve arbitrary thresholds are likely to impose a looser coupling between the legal status of an action and the "wrongness" of the act. Dumping a chemical a few feet outside a recognized zone, for example, could be regarded as only *slightly* less harmful than dumping the chemical just inside the zone. Yet crossing the threshold to illegal status will make the act wrong not just by an incremental increase in harm, but by virtue of having violated a prohibition. For mens rea crimes, in contrast, harm may track legal status more closely: taking something that doesn't belong to you is both wrong and illegal; not doing so is neither. For these reasons, the evaluation of strict liability crimes –relative to mens rea crimes – could be less sensitive to mental states, and also more dependent on the presence of a rule.

In light of these considerations, we can revisit our specific crimes. Reckless driving, seducing a minor, and minor in possession – all mens rea crimes – also involve a potentially arbitrary threshold or contingency. And in fact, in Experiment 2, reckless driving had the greatest drop in moral censure after the statute change of any mens rea crime, while seducing a minor had the second greatest drop in moral censure and the highest for fines and jail time of any mens rea crime (see Figure 3-2 and Appendix A). Minor in possession had the third greatest reduction in moral censure and a complete erasure of fines. In Experiment 3, minor in possession was also judged the second most arbitrary crime. On the other hand, the strict liability crimes of incest and unregistered firearm showed the smallest drops in moral censure in response to a statute change, and both arguably lack an arbitrary threshold beyond which they become legal.

Considering these "exceptions to the rule" helps make a valuable point. We do not mean to suggest that strict liability crimes form a special psychological kind, or that they have clear-cut necessary and sufficient conditions. Rather, a variety of different factors that come in degrees may be more likely to be found in strict liability crimes, and these features can potentially help explain their legal origins as well as the way they're evaluated by both experts and laypeople. The surprise, perhaps, is that the strict liability designation has *any* counterpart in intuitive moral judgments; that the match is imperfect is to be expected, especially given the recognized heterogeneity of strict liability crimes within the law (Levenson, 1992; Sayre, 1933).

One lingering concern is that our findings could reflect American citizens' knowledge of their own legal system, and not their untutored moral intuitions. Arguing against this possibility, an additional study confirmed that participants sampled from the same population as that in our studies were no more likely to think that knowledge was required for the conviction of our mens rea crimes than of our strict liability crimes (see online supplement for details).

Implications for the Law

Finding that the legal category of strict liability mirrors a cognitive distinction is exciting, but also potentially surprising: some of strict liability's most vocal criticism is rooted in the belief that defendants' mental states *should* and *would* have an effect on their legal outcomes. That is, there is often an implicit assumption that jurors care about a defendant's knowledge or intent, and that presenting evidence for the absence of either would mitigate a defendant's penalty in strict liability cases. For instance, Justice Jackson in *Morissette* said:

> "The contention that an injury can amount to a crime only when inflicted by intention is no provincial or transient notion. It is as universal and persistent in mature systems of law as belief in freedom of the human will and a consequent

33

ability and duty of the normal individual to choose between good and evil." (*US v. Morissette*, 1952, p. 250).

In contrast, we found that our participants' judgments were surprisingly consistent with the law, finding mental states less relevant for strict liability crimes than for other crimes. We are not arguing that mental states play *no* role for strict liability crimes, but our results do suggest that the (putative) counterintuitiveness of this aspect of strict liability has been overstated, and might not be the most compelling basis for arguments by its opponents. Indeed, recent work in the realm of torts has found that laypeople may endorse strict liability more strongly than scholars (Sanders et al., 2014).

Our studies also bear on more general questions about the role of explanations in legal judgment. We found evidence that offering an *explanation* for an actor's mental states can be more mitigating than the mere absence of knowledge. Specifically, Experiments 1 and S1 found that *unspecified* participants gave significantly higher ratings on every dependent measure than either of the conditions that offered explanations for false beliefs: *unknowing* or *deceived*. Experiment S1 also found that deception was more mitigating than bad information. Our studies thus go beyond prior work about the importance of knowledge and intent by showing that while mental states are undeniably important, *explanations* for those mental states are as well.

Limitations and Future Directions

One major concern with the present research is its lack of ecological validity. Certainly, our participants were making judgments in a much less formal and demanding environment than actual jurors judging a real case. Moreover our participants did not engage in group deliberation, and had less information than would be available at trial. However, these limitations do not undercut the central aim of the present research, which was to investigate whether laypeople's intuitive moral judgments discriminate between strict liability and mens rea crimes in the absence of specific instructions to do so, and if so, why.

Our results support the idea that the relative arbitrariness of a crime, and therefore its status as *malum prohibitum*, is intuitively recognized and influences judgments. However, open questions remain concerning how participants were interpreting and evaluating arbitrariness. In addition, the relationship between arbitrariness and the role of mental states was significant, but small. This suggests that a variety of additional factors influence the relative importance of mental states in legal judgment, and identifying these factors is an important direction for future research.

Another concern is the number of participants who failed our attention check questions. Using an Internet sample necessitates stringent attention checks, as participants cannot be monitored as they would be in a lab. Despite our multiple checks, however, our exclusion rates were comparable to those for other studies run on Mechanical Turk (e.g., Downs, Holbrook, & Sheng, 2010). One advantage of using a Mechanical Turk sample is that it is more representative of a jury than a sample of university undergraduates. That said, extending this line of research to jury-eligible participants who engage in deliberation over more detailed and realistic cases has clear value.

Conclusion

In sum, our studies are the first (to our knowledge) to relate the strict liability designation to laypeople's intuitive moral judgments. We find that strict liability and mens rea crimes differ in two potentially related ways: mens rea crimes are more sensitive to the presence or absence of

relevant knowledge, but strict liability crimes are more contingent on the presence of a rule. Our studies suggest that people find strict liability crimes to be more arbitrary and *malum prohibitum,* while mens rea crimes appear to be, to a greater extent, *malum in se*. This is good news for those, like Justice Holmes, who believe that the law should reflect community standards, but it may still be troubling for those who believe the law should maintain the highest standard of justice, even when the community does not demand it.

Chapter 3

An actor's knowledge and intent are more important in evaluating moral transgressions than conventional transgressions

Abstract

An actor's mental states – whether she acted knowingly and with bad intentions – typically play an important role in evaluating the extent to which an action is wrong and in determining appropriate levels of punishment. In four experiments, we find that this role for knowledge and intent is significantly weaker when evaluating transgressions of conventional rules as opposed to moral rules. We also find that this attenuated role for knowledge and intent is partly due to the fact that conventional rules are judged to be more arbitrary than moral rules; whereas moral transgressions are associated with actions that are intrinsically wrong (e.g., hitting another person), conventional transgressions are associated with actions that are only contingently wrong (e.g., wearing pajamas to school, which is only wrong if it violates a dress code that could have been otherwise). Finally, we find that it is the perpetrator's *belief* about the arbitrary or non-arbitrary basis of the rule – not the reality – that drives this differential effect of knowledge and intent across types of transgressions.

Introduction

Both laypeople's intuitions and the law accord a prominent role to a perpetrator's knowledge and intentions when it comes to assessing the severity of a transgression and how it should be punished (Young & Tsoi, 2011; Mikhail, 2009; Cushman, 2008). For example, serving someone a cup of coffee sprinkled with poison is deemed quite a bit worse when it was done *intentionally* – with full knowledge that the white powder added to the coffee was poison – than when it resulted from the false belief that the white powder was sugar (Young et al., 2007). To take a legal example, determinations of whether a defendant should be charged with murder versus manslaughter depend, in large part, on whether the killing was intentional.

Although knowledge and intent often play a central role in moral judgment, recent findings reveal that these mental states are not equally influential in evaluating transgressions of all types (Barrett et al., 2016; Chakroff et al., 2015; Hawley-Dolan & Young, 2013; Russell & Giner-Sorolla, 2011; Young & Saxe, 2011). This variation is also reflected in the law: for crimes classified as strict liability, such as speeding or statutory rape, the American legal system does not require the presence of *mens rea* – "a guilty mind" – for a defendant's conviction. Moreover, laypeople's intuitive moral judgments mirror this aspect of the law: Giffin and Lombrozo (2016) found that ignorance was less mitigating for strict liability crimes compared to "mens rea" crimes, such as theft or battery. For instance, people treated ignorance that one was speeding (even if the speedometer was broken) as less mitigating than ignorance that one committed theft (due, for example, to a false belief about an object's ownership).

Here we explore the prediction that knowledge and intent are also differentially important when it comes to evaluating transgressions of moral rules (such as hitting another person) versus conventional rules (such as violating a dress code). Specifically, we test the novel prediction that these mental states have a larger impact in moral cases relative to conventional cases when evaluating the "wrongness" of a transgressor's actions and how a transgressor should be punished. In the sections that follow, we briefly review prior work that motivates this prediction, and we provide an overview of the four experiments that follow.

From Strict Liability to the Moral / Conventional Distinction

One basis for our prediction that knowledge and intent may be less important in evaluating transgressions of conventional rules relative to moral rules comes from recent work in the psychology of law. Giffin and Lombrozo (2016) found an attenuated role for knowledge and intent in evaluating strict liability crimes (relative to other crimes), and explained this finding by appeal to a legal distinction between transgressions that are *malum in se*, or wrong in themselves, versus those that are *malum prohibitum*, or wrong because prohibited (*US v. Morissette*, 1952). Battery[18], to take one example, is arguably wrong in itself: it is wrong to hit another person whether or not there are rules against harmful physical contact. Driving at 50-miles per hour, in contrast, is not *inherently* wrong, but becomes wrong if the speed limit is 35-miles per hour. The knowledge and intent that underlie the commission of crimes that are considered *malum in se* versus *malum prohibitum* could vary accordingly. In the former case, the intention to violate a rule (e.g., prohibiting battery) is inextricably bound with an additional intention that is itself reprehensible: an intention to harm. In the latter case, the intention to violate a rule (e.g.,

---

[18] Battery is defined in California penal code 242 as the "(1) unlawful and willful; (2) application of force or violence; (3) upon the person of another." Hitting or applying force to another in self-defense, defense of another, defense of property, or as part of your official duties would not constitute battery.

prohibiting speeding) is only *contingently* bound to a bad intention: were the rule against speeding not in place or the speed limit raised to 50-miles per hour, intentionally driving 50-miles per hour down a well-lit, empty street would be perfectly fine. We might therefore expect that, in general, the intentions that accompany the knowing transgression of a rule that is *malum in se* will be more reprehensible than those that accompany the knowing transgression of a rule that is *malum prohibitum.*

Giffin and Lombrozo (2016) found empirical support for this proposal in two forms. First, they found that when a statute prohibiting a particular act was repealed, judgments about how wrong the corresponding transgression was went down significantly more for strict liability crimes than for mens rea crimes. This suggests that the strict liability offenses were regarded to a greater extent as wrong *because they were prohibited* (i.e., as *malum prohibitum).* Second, Giffin and Lombrozo found that, on average, participants judged strict liability laws more arbitrary than those governing mens rea crimes, suggesting that participants regarded the former as somewhat regulatory, and the latter as linked more intrinsically to non-arbitrary matters of harm – that is, to acts that are *malum in se*, or wrong in themselves.

The distinction between *malum prohibitum* and *malum in se* finds a counterpart in research on moral psychology, where scholars have differentiated between transgressions of moral and conventional rules. In classic experiments, Turiel and colleagues presented children with stories in which an actor violated a rule. The children were asked to judge how bad the actor's behavior was, both with the rule in place and in a situation in which the rule did not apply (Turiel, 2008; Weston & Turiel, 1980). They found that children as young as six judged transgressions of moral rules (such as hitting another child) wrong, even when no explicit rule was in place, suggesting that they found the act to be wrong in itself. However, transgressions of conventional rules (such as a dress code) were only judged wrong when the rule was in place, suggesting the act was wrong merely because it was prohibited.

The importance of the moral/conventional distinction is further supported by more recent developmental work, which finds that children are sensitive to the fact that people can *choose* to opt out of conventional rules, in a way that is not appropriate when the rule is moral (Josephs & Rakoczy, 2016). Moreover, the distinction is reflected in a variety of behaviors that emerge before the age of 6: 5-year-old children tattle and have stronger emotional reactions to moral transgressions than to conventional transgressions (Hardecker, Schmidt, Roden, & Tomasello, 2016), and 3- and 4-year-olds differentiate between moral and conventional rules insofar as they place more emphasis on freedom of action in the moral domain (Josephs, Kushnir, Grafenhain, & Rakoczy, 2016). Specifically, Josephs, Kushnir, Grafenhain, and Rakoczy (2016) found that children were more likely to protest when an agent violated a moral rule if the agent did so when alternative actions were available (as opposed to being constrained in their choice of actions), but that the availability of alternatives had a smaller impact on protests directed towards agents who violated conventional rules.

While scholars differ in how they conceptualize the moral/conventional distinction (e.g., Nichols, 2008; Turiel 2008b; Weston & Turiel, 1980), one important element seems to be the *arbitrariness* of a prohibition: like the rules against strict liability evaluated by participants in Giffin and Lombrozo (2016), conventional rules are somewhat arbitrary in the sense that they could have been different (e.g., the *color* specified by a dress code, the specific side of the plate on which a fork is placed), even when there are good reasons for having some sort of regulation in place. In Josephs, Kushnir, Grafenhain, and Rakoczy (2016), for example, the conventional

rules specified which color marble should be placed in which box – a rule that could easily have been different.

The results from Giffin and Lombrozo (2016), combined with the developmental work on the moral/conventional distinction, generate a previously unexplored prediction: that knowledge and intent should have a greater impact on how people evaluate the "wrongness" of transgressions involving moral rules versus conventional rules, with corresponding effects for the levels of punishment deemed appropriate in each case. Put differently, the gap between how wrong it is to hit someone knowingly versus accidentally should be greater than the gap between how wrong it is to break the dress code knowingly versus accidentally. Across four experiments, we find support for this prediction and test competing hypotheses about what drives differential effects of knowledge and intent across different types of transgression.

Overview of Experiments

In Experiment 1, we find support for our key prediction. While transgressions of both moral and conventional rules are judged more harshly when an actor transgressed knowingly as opposed to unknowingly, the effect of knowledge is greater for transgressions of moral rules than of conventional rules. In Experiment 2, we test two candidate explanations for this effect: that it is driven by greater expected or actual harm in moral cases, or that the critical difference instead lies in the fact that moral transgressions are *malum in se*, whereas transgressions of convention are often *malum prohibitem*, as they involve the transgression of a somewhat arbitrary convention. We find support for the latter possibility, and in Experiment 3 go on to test the role of arbitrariness experimentally. Finally, in Experiment 4, we investigate whether it is the reality (whether a rule is *in fact* arbitrary) or the transgressor's beliefs (whether she *believes* that the rule is arbitrary) that drives judgments.

Experiment 1

In Experiment 1, we test the prediction that a perpetrator's knowledge and intent moderate how wrong a transgression is perceived to be and how much punishment is deemed appropriate, but that this role is greater in evaluating transgressions of moral rules (hereafter referred to as "moral transgressions") than of conventional rules (hereafter referred to as "conventional transgressions"). To do so, we compare judgments of "wrongness" (i.e., how wrong an act is judged to be) and punishment across vignettes involving transgressions of a stipulated moral or conventional rule, where the transgression is committed knowingly or unknowingly (i.e., due to an accident or false belief concerning something other than the rule itself). In other words, we test the prediction that the evaluation of moral transgressions is more "knowledge dependent" than the evaluation of conventional transgressions. For the sake of continuity with previous research, we also attempt to replicate the well-established finding that judgments concerning conventional transgressions tend to be more "rule dependent" (i.e., contingent on the presence of a rule) than those concerning moral transgressions.

Methods

Participants

Two-hundred-and-forty adults (105 female, 134 male, 1 other/prefer not to specify, mean age = 32, *SD* = 15) participated in the study through Amazon Mechanical Turk for monetary

compensation. An additional 28 participants were tested, but were excluded for failing catch questions (27) or to ensure even numbers in all conditions (1). Participation was restricted to workers with IP addresses in the United States and an approval rating of 95% or higher on previous tasks.

Materials and procedure

The experimental stimuli consisted of 12 distinct vignettes, six of which concerned conventional transgressions and six of which concerned moral transgressions. Six of the vignettes (Teacher's Title, Greeting, Baseball, Dollar, Pushing, and Embezzler) were based on vignettes originally presented to children by Davidson, Turiel, and Black (1983). The Embezzler vignette was additionally modified to take place in a school setting (see Supplementary Material for full stimuli).

Each participant read only one of the 12 vignettes, leading to 12 conditions. Participants first read the *unknowing* version of their assigned vignette and answered wrongness and punishment questions. For instance, the *unknowing* Baseball vignette read:

> Jack is a boy who likes to play games, especially baseball…There's a rule that all the students on the Blue Jays' team wear blue shirts with the school logo on the back to baseball practice. The school takes this rule and respect for the school and its logo very seriously.
>
> One day, Jack was getting ready for a baseball practice. He was in a hurry to get to the bus on time, so he dressed quickly and left. Jack didn't realize he had grabbed the wrong blue shirt. So Jack went to the practice wearing a blue shirt that did not have the school logo on the back, in violation of the long-standing school policy.
>
> Because Jack was wearing the wrong shirt, he was allowed to practice that day, but was sent to the principal's office after practice.

Participants were then asked, in random order:

> *Wrongness*: "How wrong was [Actor's actions]?" Participants indicated their answer on a scale from 0 (not at all wrong) to 6 (very wrong).
> *Punishment*. "Students who break a rule at [Actor's] school are given school service hours during which they clean classrooms, organize supplies, and pick up trash on the grounds. How many hours of school improvement service should [Actor] get?" Participants indicated their answer on a scale from 0 to 6 hours.

Participants were then asked to imagine that the actor had instead violated the rule knowingly, and again rated wrongness and punishment.[19] Below is the *knowledge change* prompt from the Baseball story:

---

[19] In a pilot version of Experiment 1 with 160 participants, the *knowing* and *unknowing* versions of each vignette were presented to different participants in a between-subjects design. This study revealed the same effects of knowledge on judgments of wrongness as those presented here. For punishment, however, the pilot study asked about hours of detention, which seemed to generate a floor effect: participants were disinclined to assign detention hours for any of our transgressions. This motivated the change to "service hours" as the measure of punishment in Experiment 1. See Supplementary Material 2 for a full report and Supplementary Material 3 for full stimuli from that experiment.

> *Knowledge Change.* "Suppose that Jack had actually realized, while he was dressing, that the shirt he was about to put on for practice violated the rule – that is, that it didn't have the logo on the back. And suppose that he decided to wear it anyway. In this case, where Jack knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)"

In no case (for either moral or conventional transgressions) did the vignette specify that the actor violated the rule with the intention to cause harm or disruption; in most cases the motivation resulted from a desire for change (e.g., being tired of following the rule).

Next, participants were told to imagine that the rule was not in place, and answered the evaluative questions a final time. These ratings were solicited to replicate prior findings that conventional transgressions are more rule dependent than moral transgressions, and thus as verification that our vignettes successfully presented transgressions that were "moral" versus "conventional" in the relevant sense. The wording of these questions (again presented in random order) was identical to that above, but preceded by a rule change.[20] Below is the *rule change* prompt from the Baseball story:

> *Rule Change.* "Finally, suppose that Jack's school had no rule prohibiting wearing a shirt without the school logo to practice, and Jack knowingly wore a shirt without the school logo to practice. In this case, with no rule about how to dress for practice in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)"

Finally, on a separate screen, participants were presented with a true/false comprehension question relating to the vignette they had just read, and answered one additional catch question, modeled after Oppenheimer, Meyvis, and Davidenko (2009). These questions were used to assess whether participants had read the vignette and the instructions carefully; those who answered either question incorrectly were excluded from further analyses. To conclude, participants answered demographic questions about their age and gender.

Results

Participants responded to the wrongness and punishment questions three times: for the initial *unknowing* transgression (with a rule in place), the subsequent *knowing* transgression (with a rule in place), and the final *no rule* transgression (with no rule in place). We first present the results from the first two sets of judgments as a measure of "knowledge dependence," and we then consider the final two sets to evaluate "rule dependence."

---

[20] This "no rule" rating was always solicited last. We feared that if we initially told participants that no rule was in place, and then later told them a rule had been instituted, they might make unwarranted assumptions about the reasons for this change. For instance, participants might assume that the actions had actually caused harm or disruption, making the rule necessary. These kind of assumptions could artificially inflate wrongness and punishment ratings. Moreover, because the primary aim of the experiment was to assess the knowledge dependence of moral and conventional violations, it seemed wise to solicit those judgments first.

Knowledge dependence

To test the prediction that judgments regarding conventional transgressions are less "knowledge dependent" than those regarding moral transgressions, we performed mixed ANOVAs with knowledge status as a within-subjects variable (2: *knowing*, *unknowing*), transgression domain as a between-subjects variable (2, conventional, moral), and either wrongness or punishment ratings as the dependent variable (see Figure 1-3). We expected to find a main effect of knowledge status, with higher ratings for knowing transgressions than for unknowing transgressions, qualified by an interaction between knowledge status and transgression domain, with a larger effect of knowledge status for moral transgressions than for conventional transgressions.

As predicted, this analysis revealed a main effects of knowledge status for both wrongness, $F(1,238) = 465.33$, $p < .001$, $\eta_p^2 = .66$, and punishment, $F(1,238) = 382.16$, $p < .001$, $\eta_p^2 = .62$; in both cases, ratings were higher for knowing than for unknowing transgressions (see Figure 1-3). However, this main effect was qualified by the predicted interaction between knowledge status and transgression domain, for both wrongness, $F(1,238) = 12.84$, $p < .001$, $\eta_p^2 = .05$, and punishment, $F(1, 238) = 33.06$, $p < .001$, $\eta_p^2 = .12$. Independent-samples t-tests comparing the average difference between participants' ratings for the *knowing* and *unknowing* vignettes confirmed that the knowledge effect was greater for moral than conventional transgressions for both wrongness ($M_M = 3.01$, $SD_M = 1.74$; $M_C = 2.11$; $SD_C = 1.64$), $t(238)= 4.12$, $p< .001$, $d = .53$, and punishment ($M_M = 2.68$, $SD_M = 1.81$; $M_C = 1.45$; $SD_C = 1.34$), $t(221)= 5.97$, $p< .001$, $d = .80$ (corrected for violating Levene's test).

Finally, there was also a main effect of transgression domain for both wrongness, $F(1,238) = 36.51$, $p < .001$, $\eta_p^2 = .13$, and punishment, $F(1,238) = 73.47$, $p < .001$, $\eta_p^2 = .24$, with higher ratings for moral transgressions than for conventional transgressions.

Figure 1-3: Ratings for wrongness and punishment for all vignette versions as a function of transgression type. Error bars correspond to one SEM in each direction.
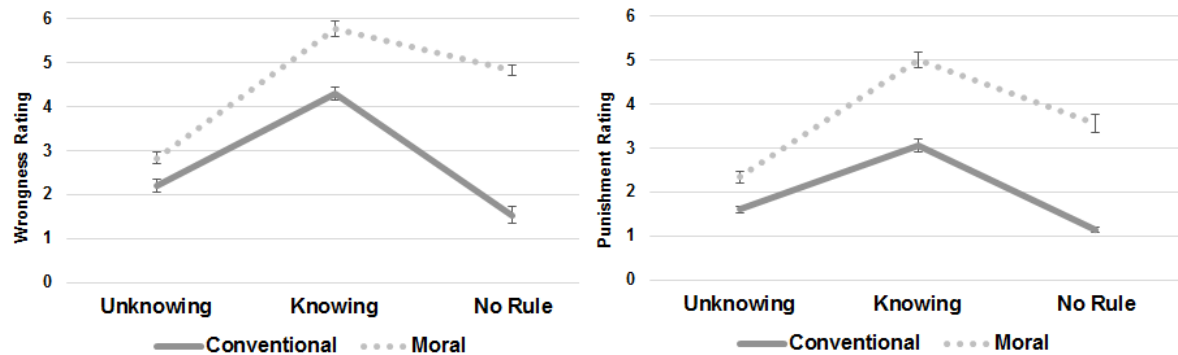
Rule Dependence

Contingency on the presence of a rule ("rule dependence") was measured by subtracting participants' wrongness and punishment ratings after the rule change from their corresponding scores for the *knowing* vignettes. Independent t-tests were performed on these difference scores, and we predicted a greater difference for conventional relative to moral transgressions.

As predicted, these analysis found that rule dependence was significantly greater for conventional than moral transgressions for wrongness (1.91 versus 1.43), $t(238)= 6.93$, $p < .001$,

d = .90, and nearly so for punishment (2.76 versus .942), *t*(221) = 1.89, *p* < .06, *d* = .25 (corrected for violating Levene's test).

Discussion

Experiment 1 found that relative to judgments concerning conventional transgressions, those concerning moral transgressions were more sensitive to whether the actor transgressed knowingly versus unknowingly. This was the case for judgments of wrongness and also for our measure of punishment (service hours). To our knowledge, this is the first demonstration that a transgressor's mental states differentially influence judgments of wrongness and punishment when it comes to moral versus conventional transgressions. We also found that relative to judgments concerning moral transgressions, those concerning conventional transgressions were more contingent on the presence of a rule. This is a familiar finding from the literature on the moral/conventional distinction, but helps confirm that our vignettes varied appropriately along this dimension.

While Experiment 1 supports our key prediction that knowledge and intent have a greater impact in evaluating moral versus conventional transgressions, the findings do not reveal why. An important question thus remains: what is it about the domain of the moral versus the conventional that drives the differential role of knowledge and intent? One possibility is that the moral transgressions in Experiment 1 were simply more severe than the conventional transgressions: if they involved greater harm, the associated intentions may have been regarded as more "wrong" and deserving of punishment. Indeed, moral transgressions do tend to be more severe (Smetana, 1995), and Experiment 1 found a main effect of domain, with harsher judgments for moral vignettes than for conventional vignettes.

On the other hand, past work has found that severity is insufficient to fully account for differences in judgment across the moral and conventional domains (Tisak & Turiel, 1988), though scholars have offered competing accounts of what the additional difference might be (Nichols, 2008; Turiel 2008b; Weston & Turiel, 1980). One account argues that conventions are fundamentally *arbitrary* (Turiel, 2008a), whereas moral norms stem from non-arbitrary considerations of harm (Turiel, 2008b). For instance, it's important that we have *some* convention about which side of the road to drive on, but the specific rule (to drive on the right versus the left) is arbitrary, varying from country to country. In contrast, moral rules are less likely to be regarded as arbitrary matters of preference (Goodwin & Darley, 2012; see also Sarkissian, et al., 2011).

This suggests a second possibility: that the relevant difference between the moral and conventional domains stems from the fact that moral rules are non-arbitrary and intrinsically linked to harm. As a result, knowingly violating a moral rule is intrinsically linked with an intention to harm, whereas knowingly violating a conventional rule is only contingently associated with bad intentions: there's nothing wrong with the intention to wear a shirt without a school logo; it only becomes wrong in the context of a rule prohibiting this behavior. On this view, it is not the *degree* of harm caused that differentiates the moral from the conventional, but rather *how* the harm is linked to the transgressive act. Reflecting this difference, Table 1-3 represents how attributions of knowledge and intent might shift at each step of Experiment 1: when the transgression is committed unknowingly, when it occurs knowingly, and when it occurs with no rule in place.

Table 1-3. Representation of the knowledge and intent one might attribute to the perpetrator at each step of Experiment 1, as a function of domain.

	Moral	**Conventional**
Unknowing Transgression	Knowledge of rule	Knowledge of rule
Knowing Transgression	Knowledge of rule Intention to break rule Intention to cause harm	Knowledge of rule Intention to break rule
No Rule	Intention to cause harm	

As the table reveals, the difference in mental states across the knowing and unknowing transgressions is greater for moral transgressions than for conventional transgressions: the former involves an additional "intention to cause harm" in the knowing case. We propose that this explains the greater knowledge dependence of moral transgressions relative to conventional transgressions. However, it's the conventional cases that involve a greater shift from the knowing transgression to the scenario without a rule: for the moral transgressions, the intention to cause harm is preserved across both cases, but for conventional cases no knowledge or intent remains. We propose that this explains the greater rule dependence of conventional transgressions relative to moral transgressions.

Of course, the violation of a conventional rule could *also* be accompanied by an intention to harm. For instance, someone could drive on the wrong side of the road with the intention to harm other drivers, or violate a dress code with the intention of hurting someone's feelings. Our contention is not that such intentions do not occur, but rather that that are not entailed by the intentional rule-breaking itself. In contrast, it's difficult to imagine a moral transgression that is *not* accompanied by an intention to harm, even if such an intention is unspecified (as was the case for our vignettes). Except in very rarefied cases, an intention to hit or steal involves an intention to harm.

If this analysis is correct, then a key difference between moral and conventional transgressions stems from the fact that conventional rules are somewhat arbitrary – this is what renders the intention to violate a conventional rule only contingently associated with additional bad intentions. Experiment 2 tests this possibility while controlling for the overall severity of the transgressions.

Experiment 2

In Experiment 2, we sought to replicate the key finding from Experiment 1: that whether a transgression is committed knowingly versus unknowingly has a greater impact on the evaluation of moral transgressions relative to conventional transgressions. However, we also sought to identify what it is about moral and conventional transgressions that drives this differential effect.

To test the possibility that the moral vignettes simply involved greater harm than the conventional vignettes, we more closely matched our *knowing* moral and conventional

transgressions along a variety of dimensions, including the amount of harm that a third person would have expected from the transgression, the amount of harm the actor foresaw, and the amount of harm that actually occurred. To accomplish this we pretested our stimuli, and additionally verified that these dimensions were matched in a post-test. If differences along these dimensions are responsible for the greater knowledge dependence of moral transgressions relative to conventional transgressions, then we should expect the domain differences observed in Experiment 1 to disappear in Experiment 2.

A second possibility is that knowledge dependence is driven by more than just actual harm, and instead depends on the nature of the perpetrator's intentions. Because conventional rules are somewhat arbitrary and not intrinsically linked to harm, the intentions associated with knowingly breaking a conventional rule are only bad contingently. These ideas generate two pairs of predictions. First, conventional rules should be regarded, on average, as more arbitrary than moral rules, in the sense that they could reasonably have been specified differently. Moreover, perceived arbitrariness should be a negative predictor of the magnitude of the knowledge effect. Second, the intentions associated with knowing moral transgressions should be regarded, on average, as worse than the intentions associated with conventional transgressions, and perceived badness should be a positive predictor of the magnitude of the knowledge effect. We test these predictions in Experiment 2.

Methods

Participants

Two-hundred-and-eighty adults (155 female, 124 male, 1 other/prefer not to specify, mean age = 35, $SD = 11$) participated in the study through Amazon Mechanical Turk as in Experiment 1. An additional 42 participants were tested, but were excluded for failing catch questions (36) or to ensure even numbers in all conditions (6).

Materials & Procedure

The experimental stimuli consisted of eight vignettes, five of which were used in the previous experiment (Dress Code, Lunch Table, Tardy, Throwing, and Pushing), and three of which (Hall Monitor, Candy, Burner) were created for this experiment after pre-testing items to better equate harm across transgression domains (see Supplementary Material 4 for full stimuli).

Each participant read only one of the eight vignettes, leading to eight distinct conditions. As in Experiment 1, participants first read the *unknowing* version of their assigned vignette and answered the wrongness and punishment questions used in Experiment 1. Participants were then asked to provide the same ratings assuming the actor had violated the rule knowingly, using the same knowledge change prompt as in Experiment 1. The order of the wrongness and punishment questions was randomized in each case.

After these ratings, participants rated the event, assuming the actor did know the relevant facts, along a series of dimensions related to harm and intent. The questions were presented in randomized order, and all were rated on a scale of 1 (none at all) to 7 (a great deal). The questions are provided below:

> *Expected Harm.* "How much harm could [Actor's] actions have been expected to cause?"
> *Foreseen Harm.* "How much harm do you think [Actor] believed his/her actions would cause?"

Actual Harm. "How much harm did [Actor's] actions actually cause?"
Intentions. "How wrong was [Actor's] intention when she behaved this way?"

Next, participants provided a rating of how arbitrary they found each of the rules used in Experiment 1 as well as the new rules added for Experiment 2, even though they had only read one rule. Participants used a 1 (not at all arbitrary) to 7 (completely arbitrary) rating scale. Below are the instructions that preceded the arbitrariness judgments:

"Below you will see a list of school rules. We'd like you to give your intuitions about how arbitrary each rule seems to you. That is, do you believe that there's a good reason for the rule to draw the line where it does in terms of which actions are allowed versus not allowed? Or does it seem like the rule is somewhat arbitrary in the sense that it could reasonably have been formulated a little differently, with a slightly different set of actions ruled in versus ruled out?

Please do not consult any outside resources, like other people or websites. We are interested in your own intuitions. Even if you think all the rules are somewhat arbitrary or not at all arbitrary, please take note of which seem more or less arbitrary and respond accordingly."

One concern in equating our vignettes in terms of various facets of harm is that in doing so, we may have inadvertently blurred or erased the boundary between moral and conventional transgressions. In addition to the question about arbitrariness, we therefore introduced an explicit judgment concerning the transgression domain. Participants made judgments about all rules used in Experiments 1 and 2 using a rating scale from 1 (clearly about morality) to 7 (clearly about conventions/norms). Below is the paragraph that preceded these judgments:

Morality of Action. "Actions can be wrong in multiple ways. Some are morally wrong, such as vandalizing a car. Others are wrong because they violate a group convention or norm, such as driving on the wrong side of the road.
Please rate the rules below on the following scale, from (1) clearly a rule about something moral to (7) clearly a rule about a group convention or norm. Please do not consult any outside resources, like other people or websites. We are interested in your own intuitions."

The relatively minor offense of vandalizing a car was used to help participants focus on the intended dimension (moral versus conventional) rather than on severity.

Finally, participants answered the same catch and demographic questions used in Experiment 1.

Results

Harm Ratings

First, we verified that we succeeded in equating the moral and conventional vignettes along various dimensions of harm. We ran independent samples t-tests comparing the three harm ratings as a function of domain (2: moral, conventional). These tests revealed that for Expected Harm and Believed Harm, there were no significant differences across domains, ps > .174. For

ratings of Actual Harm there was a significant difference, $t(229) = 4.08$, $p < .001$, $d = .54$ (corrected for violating Levene's test), but ratings were higher for the *conventional* vignettes than for the moral vignettes ($M_C = 2.84$, $SD_C = 1.97$; $M_M = 2.04$, $SD_M = 1.19$). These findings confirm that we succeeded in making the moral transgressions no worse than the conventional transgressions along these dimensions of harm.

Domain Ratings

In light of the ways in which our vignettes were matched in terms of harm, it was important to test whether the moral and conventional vignettes were still clearly differentiated in terms of their perceived domain. Averaging judgments for the four moral transgressions versus the four conventional transgressions, a paired-samples t-test revealed that our moral rules were found to be based significantly more on moral precepts than were our conventional rules ($M_M = 2.63$, $SD_M = 1.06$; $M_C = 5.98$, $SD_C = 1.01$), $t(279) = 31.71$, $p < .001$, $d = 3.80$. The same difference was observed if we considered a single rating from each participant corresponding to the moral basis of the rule in the vignette that they evaluated, and performed an independent samples t-test with domain as a between-subjects factor ($M_M = 2.61$, $SD_M = 1.83$; $M_C = 5.96$, $SD_C = 1.36$), $t(257) = 17.32$, $p < .001$, $d = 2.16$.

Knowledge Dependence

Having verified that harm was no greater for the moral vignettes than the conventional vignettes (despite preserving a difference in perceived domain), we went on to test the prediction that judgments about conventional transgressions should still be less knowledge dependent than corresponding judgments about moral transgressions. Mirroring Experiment 1, we performed mixed ANOVAs with knowledge status as a within-subjects variable (2: *knowing*, *unknowing*), transgression domain as a between-subjects variable (2: conventional, moral), and either wrongness or punishment ratings as the dependent variable (see Figure 2-3). As in Experiment 1, we expected to find a main effect of knowledge status qualified by an interaction between knowledge state and transgression domain, with a larger effect of knowledge for moral than conventional transgressions.

Mirroring Experiment 1, the analysis revealed a significant main effect of knowledge status for both wrongness, $F(1,278) = 366.20$, $p < .001$, $\eta_p^2 = .568$, and punishment, $F(1,278) = 374.73$, $p < .001$, $\eta_p^2 = .574$, with higher ratings for the *knowing* vignettes. As predicted, these main effects were qualified by significant interactions between knowledge status and transgression domain for both wrongness, $F(1,278) = 7.42$, $p < .007$, $\eta_p^2 = .03$, and punishment, $F(1,278) = 11.45$, $p < .001$, $\eta_p^2 = .04$. Independent samples t-tests on the average difference score between the *knowing* and *unknowing* ratings revealed that the knowledge effect was significantly greater for moral than conventional transgressions for both wrongness ($M_M = 2.44$, $SD_M = 1.95$; $M_C = 1.83$, $SD_C = 1.77$), $t(278) = 2.73$, $p < .007$, $d = .41$, and punishment ($M_M = 1.80$, $SD_M = 1.35$; $M_C = 1.26$, $SD_C = 1.30$), $t(278) = 3.38$, $p < .001$, $d = .41$.

The main effect of transgression domain was not significant for wrongness, $F(1,278) = 2.72$, $p < .431$, $\eta_p^2 = .002$, but was significant for punishment, $F(1,278) = 8.90$, $p < .003$, $\eta_p^2 = .031$, with punishment being higher for conventional ($M_{UK} = 1.53$, $SD_{UK} = 1.64$; $M_K = 2.79$, $SD_K = 1.87$) than moral ($M_{UK} = .76$, $SD_{UK} = .89$; $M_K = 2.56$, $SD_K = 1.59$) transgressions.

Figure 2-3: Ratings for wrongness and punishment for both knowledge states as a function of transgression type. Error bars correspond to one SEM in each direction.
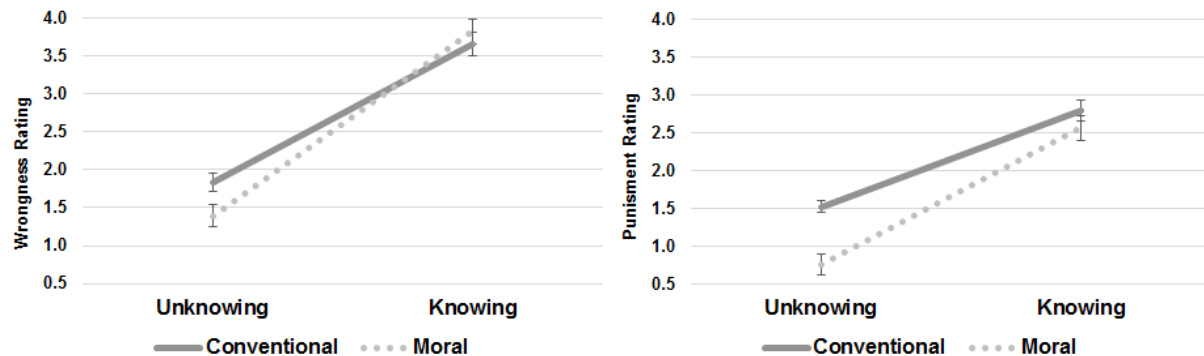
Arbitrariness Ratings

To assess whether there were domain differences in the perceived *arbitrariness* of different types of transgressions, we computed a pair of averages for each participant: the average arbitrariness ratings for the four rules corresponding to our moral vignettes versus those corresponding to our conventional vignettes. A paired samples t-test revealed that moral rules were rated significantly less arbitrary than conventional rules (M_M =2.65, SD_M = .019; M_C = 3.85, SD_C = .076), $t(279) = 11.87$, $p < .001$, $d = 1.42$. This same pattern was observed if we considered a single rating from each participant corresponding to the arbitrariness of the rule in the vignette that was evaluated, and performed an independent samples t-test with domain as a between-subjects factor (M_M =2.62, SD_M = 1.95; M_C = 3.81, SD_C = 1.97), $t(278) = 5.06$, $p < .001$, $d = .61$.

Intention Ratings

We also predicted that the perpetrator's intention would, on average, be perceived as more wrong in the moral vignettes than in the conventional vignettes. However, an independent samples t-test comparing intention ratings as a function of domain did not reveal a significant difference, $t(278) = .561$, $p < .575$, $d = .07$. It may be that in equating foreseen harm we effectively erased domain differences in intent, or that by asking specifically about the wrongness of the perpetrator's *intentions*, we missed out on other relevant mental states.

Predicting Knowledge Effects

As in Experiment 1, we found that judgments concerning moral transgressions were more sensitive to knowledge than corresponding judgments about conventional transgressions. We also found, as predicted, that participants rated conventional rules as significantly more arbitrary than moral rules. However, participants did not judge the intention of the actor to be significantly worse in the moral vignettes, as we had predicted. Regardless, we still predicted that the arbitrariness of the rule and the wrongness of the intention would be significant predictors of the knowledge effect – with arbitrariness as a negative predictor and wrongness of intention as a positive predictor. Thus, we ran simultaneous linear regressions for both wrongness and punishment with arbitrariness and intention ratings as predictors of the difference in ratings between the *knowing* and *unknowing* vignettes.

Table 2-3: Regression results for the wrongness knowledge effect.

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Standard Error	Beta	t	Significance
Constant	1.862	.292		6.374	.000
Arbitrariness	-.146	.053	-.158	-2.724	.007
Intention	.199	.051	.225	3.885	.000

As predicted, the arbitrariness of the rule was a significant, negative predictor while the wrongness of the actor's intention was a significant, positive predictor for both the wrongness knowledge effect, $R^2 = .078$, $F(2, 277) = 12.77$, $p < .001$, and the punishment knowledge effect, $R^2 = .062$, $F(2, 277) = 10.19$, $p < .001$ (see Tables 2-3 and 3-3).

Table 3-3: Regression results for the punishment knowledge effect.

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Standard Error	Beta	t	Significance
Constant	1.358	.211		6.444	.000
Arbitrariness	-.094	.039	-.142	-2.435	.016
Intention	.129	.037	.203	3.471	.001

Discussion

Experiment 2 replicated both the punishment and wrongness knowledge effects found in Experiment 1. We found that judgments concerning moral transgressions were more sensitive to knowledge than corresponding judgments concerning conventional transgressions, despite the fact that none of the harm measures were significantly higher for moral transgressions.

Further, Experiment 2 offers support for our proposal that the knowledge effect is tracking the arbitrariness of the rule that is transgressed, as well as the intentions of the actor. We found that participants rated conventional rules to be more arbitrary, and that arbitrariness was a negative predictor of the magnitude of knowledge effects.[21] We also found that the wrongness of the intention was a positive predictor of both knowledge effects.

---

[21] Notably, the relationship between arbitrariness and knowledge effects held not only across domains, but also within domains: further analyses revealed negative correlations between both knowledge effect scores and how arbitrary a participant rated the rule in the vignette they read, whether analyses were restricted to only moral or conventional vignettes. For moral violations, the correlations between arbitrariness score and the wrongness knowledge effect, $r = -.160$, $p < .059$, and punishment knowledge effect, $r = -.149$, $p < .079$, were both negative and marginally significant. For conventional violations the correlations between arbitrariness and the wrongness knowledge effect, $r = -.133$, $p < .116$, and the punishment knowledge effect, $r = -.084$, $p < .324$, were still negative, though farther from significance.

One of our predictions was not confirmed: the actor's intentions were not perceived as significantly more wrong for knowing moral transgressions than for knowing conventional transgressions. One possible explanation is that our intention question was simply too narrow to capture the full range of mental states that participants considered. For instance, participants might have believed that while intentions were not significantly worse in the moral vignettes, the actor should have taken more care.

In sum, Experiment 2 succeeded in replicating the domain differences observed in Experiment 1 while ruling out the possibility that these differences were driven by greater anticipated, foreseen, or actual harm in the moral vignettes. The findings also supported three of our four novel predictions: we found that the conventional rules were judged to be more arbitrary than the moral rules, and that arbitrariness was associated with a smaller knowledge effect. We also found that the wrongness of the intention associated with a knowing transgression predicted the magnitude of the knowledge effect, but we did not find that intentions were perceived to be more wrong in the moral vignettes than the conventional vignettes.

Experiment 3

In Experiments 1 and 2 we found support for the prediction that relative to conventional transgressions, the evaluation of moral transgressions is more sensitive to the perpetrator's knowledge and intent. In Experiment 2, we found that this difference remains when moral and conventional transgressions are matched across a variety of dimensions of harm. Experiment 2 also suggested that a relevant difference between moral and conventional transgressions is the extent to which the violated rule is perceived to be *arbitrary* in the sense that it could reasonably have been otherwise. In Experiment 3 we pursue this association experimentally by manipulating whether otherwise-matched actions violate rules that are arbitrary or non-arbitrary.

All rules in Experiment 3 involved the prevention of harm, and thus fall under the moral domain. In order to effectively manipulate whether such rules were perceived to be more or less arbitrary, we developed vignettes in an alien world for which we could stipulate associated harms without challenging participants' preconceptions about what was or was not arbitrary. Each rule specified some threshold – for example, that (alien) students were not permitted to watch gory movies below the age of 15, because younger students were more likely than older students to suffer negative emotional consequences from doing so. Thus, the presence of *some* rule was not arbitrary, but the exact threshold of 15 years *could* be. Half the participants were told that the threshold was based on a discrete developmental change that occurs at exactly age 15, and was therefore not arbitrary. The remaining participants were told that there was a more gradually changing continuum, so the exact choice of 15 was arbitrary in that 14.5 or 15.5 could reasonably have been chosen instead. We could thus test the prediction that judgments would be more sensitive to mental states for non-arbitrary rules relative to arbitrary rules using actions that were otherwise identical.

Methods

Participants

Two-hundred-and-forty adults (103 female, 135 male, 2 other/prefer not to specify, mean age = 33, *SD* = 11) participated in the study through Amazon Mechanical Turk as in Experiments

1-2. An additional 90 participants were tested, but were excluded for failing catch questions (46) or to ensure even numbers in all conditions (44).

Materials & Procedure

The experimental stimuli consisted of six distinct vignettes (Movies, Humming, Vitamins, Dumping, Laser Gun, and Speeding). The stimuli from Experiment 3 involved an alien school on an alien planet. The shift to an alien world, while regrettable in some respects, was deemed necessary as a way to exert experimental control over participants' beliefs about the relationship between different actions and possible harms.

Each vignette had two variants, one in which the rule had been set arbitrarily, and one in which the rule had been set non-arbitrarily, with arbitrariness as a between-subjects factor. For example, in the Humming vignette, all participants learned about a rule concerning where they could practice loud and potentially disruptive alien humming: they were not allowed to do so within 15 feet of an occupied classroom. The relevant excerpt from the arbitrary and non-arbitrary versions of the Humming vignette are presented below (see Supplementary Material 5 for full stimuli):

> *Arbitrary.* "…The school administration chose the distance of 15 feet somewhat arbitrarily. There isn't any reason to suspect that the noise would be much more disruptive at 14.5 feet than at 15.5 feet, but they had to choose some distance cutoff, and decided on 15 feet…"
> *Non-Arbitrary.* "…The school administration chose the distance of 15 feet after careful research on the unique auditory capabilities of their species. Aliens of their species can hear the noises clearly if they are made within exactly 15 feet. Beyond that distance, however, the noises fall below the critical threshold for the alien's auditory system, so they are barely audible and no longer distracting…"

Participants first read the *unknowing* version of their assigned vignette and answered the questions about wrongness and punishment from Experiments 1-2. As in Experiments 1-2, participants were then asked to imagine that the actor had knowingly violated the rule and to answer the questions once again. The question order was randomized in each case.

Finally, as in Experiment 1, participants were asked to imagine that the rule was not in force and to answer the wrongness and punishment questions a third and final time. In this experiment, we took care to inform the participant that the rule change was not due to a change in the harm caused by the action. We took this extra step because those participants who read an arbitrary version of a vignette might be more likely to believe that the rule had been changed because the action was not actually causing harm at the arbitrarily selected threshold. That is, as the rule was set arbitrarily, a rule change might be taken as an indication that the rule was ill conceived and unnecessary from the start. Below is a sample from the Humming vignette:

> *Rule Change.* "Finally, suppose that due to a clerical error, the rule against humming within 15 feet of a classroom had never been put in the school charter and therefore could not be enforced. That is, suppose that Bernice's school EFFECTIVELY had no rule prohibiting humming within 15 feet of an occupied classroom, and Bernice practiced her humming at her locker. In this case, with no rule about humming in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)"

51

After participants read all three versions of their assigned vignette (*unknowing*, *knowing*, and *no rule*), they answered catch questions and demographic questions similar to those in Experiments 1-2.

Results

Knowledge Dependence

To test our prediction that knowledge and intent play a greater role in evaluating transgressions of non-arbitrary rules than of arbitrary rules, we performed mixed ANOVAs with rule type (2: arbitrary, non-arbitrary) and vignette (6: Movies, Humming, Vitamins, Dumping, Laser Gun, and Speeding) as between subjects' factors, knowledge state (2: *unknowing*, *knowing*) as a within-subjects measure, and either wrongness or punishment (service hours) as the dependent variable (see Figure 3-3). We predicted a greater knowledge effect for the non-arbitrary transgressions, mirroring the greater knowledge effect seen for moral transgressions. This would manifest as an interaction between rule type and knowledge state.

As predicted, we found a significant interaction between knowledge state and rule type for both wrongness, $F(1,228) = 6.96$, $p < .009$, $\eta_p^2 = .03$, and punishment, $F(1,228) = 13.31$, $p < .001$, $\eta_p^2 = .06$: the increase in ratings from the *unknowing* condition to the *knowing* condition was greater for non-arbitrary transgressions than for arbitrary transgressions, both for judgments of wrongness ($M_{NA} = 2.37$, $SD_{NA} = 1.70$; $M_A = 1.81$, $SD_A = 1.60$), $t(238) = 2.62$, $p < .009$, $d = .34$, and for punishment ($M_{NA} = 2.28$, $SD_{NA} = 1.55$; $M_A = 1.58$, $SD_A = 1.38$), $t(238) = 3.65$, $p < .000$, $d = .47$.

The analysis also revealed a main effect of knowledge state, with higher ratings in the *knowing* condition than the *unknowing* condition for both wrongness, $F(1,228) = 388.96$, $p < .001$, $\eta_p^2 = .63$, and punishment, $F(1,228) = 414.22$, $p < .001$, $\eta_p^2 = .65$. We did not find a main effect of rule type, for either wrongness, $t(238) = 0.55$, $p > .580$, $d = .07$, or punishment, $t(238) = 0.27$, $p > .787$, $d = .04$. There were no main effects nor interactions involving vignette.

Figure 3-3: Ratings for wrongness and punishment for all mental states as a function of transgression type. Error bars correspond to one SEM in each direction.
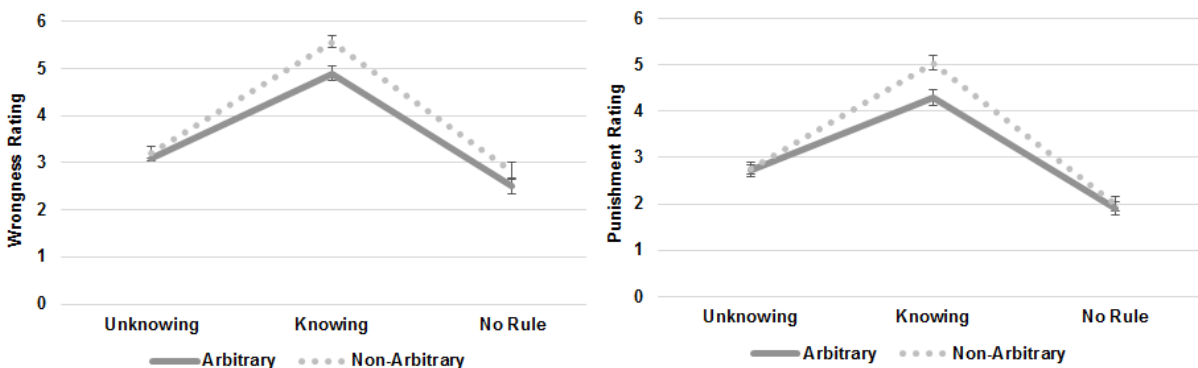
Rule Dependence

Finally, might transgressions of *arbitrary* rules mimic conventional transgressions in their contingency on a rule? To answer this question, we created a measure to reflect the effect of the rule, following the same procedure as in Experiment 1. We performed a series of 2 (rule type:

arbitrary, non-arbitrary) x 6 (vignette: Movies, Humming, Vitamins, Dumping, Laser Gun, and Speeding) ANOVAs to assess the rule effect. Contrary to our expectations, we found no significant difference between arbitrary and non-arbitrary transgressions for wrongness ($p > .19$), but we did see a significant difference for school service hours $F(1,228) = 6.31$, $p < .013$, $\eta_p^2 = .027$.

Discussion

Experiment 3 confirmed our prediction that knowledge dependence would be greater for non-arbitrary transgressions relative to arbitrary transgressions. We hypothesized that part of what drives the knowledge effect for moral rules is that they are typically non-arbitrary. That is, because moral rules prohibit actions that are intrinsically tied to harm, the rules could not reasonably have been otherwise (i.e., they are non-arbitrary), and a knowing transgression constitutes a knowing commission of harm. Experiment 3 tested this by setting all violations in the moral domain, but manipulating the arbitrariness of the rules. The key prediction was that a knowing violation of a non-arbitrary rule would be deemed more wrong and more deserving of punishment than a knowing violation of an arbitrary rule, resulting in an interaction between arbitrariness and knowledge. This is precisely what we found.

Participants' ratings for both wrongness and punishment (school service hours) were less sensitive to knowledge and intent when the threshold specified by the rule was arbitrary. Thus, we were able to replicate the pattern of knowledge dependence seen in moral and conventional transgressions entirely within the moral domain by manipulating whether the rule was set arbitrarily. This finding supports our findings from Experiment 2 that the importance of these mental states is not solely a function of whether a norm involves harm, as all the transgressions considered in Experiment 3 were moral in nature and tied to a real harm, but is also influenced by the arbitrariness of the rule itself – that is, whether the rule could reasonably have been specified otherwise. These results leave open, however, whether the critical issue is *reality* (whether the rule is in fact set arbitrarily or not) or the actor's *belief* about the rule; we return to this issue in Experiment 4.

Experiment 3 failed to find significantly greater effects of a rule change for arbitrary relative to non-arbitrary rules for wrongness, although this predicted effect was observed for punishment. We suspect that this failure to replicate the pattern observed on Experiment 1 comes from the smaller effect sizes associated with rule type in the present experiment, and the corresponding loss in statistical power.

Experiment 4

While Experiments 2 and 3 found support for the idea that mental states are more influential when a rule is non-arbitrary, it's not clear what was driving this effect: that rules were *believed* by the actors to have arbitrary versus non-arbitrary bases, or that the *reality* was such that some transgressions crossed an arbitrarily set threshold, while others crossed a non-arbitrary threshold. In other words, was the effect driven by the knowledge and intentions themselves, or by their differential consequences in the world?

Based on our predictions and findings in Experiments 1-2, we would expect a critical factor to be the knowledge and intent of the actor. In Experiments 1-2, we suggested that moral transgressions involve a greater role for knowledge and intent because some desires and

intentions are themselves inextricably bound to the harm that the action will cause, thereby influencing the perceived magnitude of wrongdoing and the appropriate level of punishment. While Experiment 2 found that intention ratings significantly predicted the magnitude of the knowledge effect, the intentions of moral transgressors were not rated significantly worse.

Thus, Experiment 4 aims to take a closer look at the role of intentions and beliefs in driving knowledge effects. If it's a perpetrator's mental states that drive such effects, then the effects found in Experiment 3 should be eliminated by stipulating that all transgressors *believed* the rules to be set arbitrarily, while varying the reality: for half of participants the rule was arbitrary (i.e., it could reasonably have been specified differently); for the other half it was not (i.e., their transgression involved crossing a non-arbitrary threshold). If instead judgments for wrongness and punishment are affected by the differential *consequences* of violating a rule that was or was not set arbitrarily (i.e., intrinsically tied to *actual* harm or not), then specifying that a transgressor always believes a rule to be arbitrary should be insufficient to eliminate the effect.

Methods

Participants
Two-hundred-and-forty adults (117 female, 122 male, 1 other/prefer not to specify, mean age = 32, SD = 12) participated in the study through Amazon Mechanical Turk as in Experiments 1-3. An additional 185 participants were excluded for failing catch questions.

Materials & Procedure
The experimental stimuli consisted of three of the stories used in Experiment 3 (Movies, Vitamins, and Speeding). Each vignette had two versions, one in which the actor believed the rule had been set arbitrarily, and in fact it had been, and one in which the actor believed the rule had been set arbitrarily, but in fact the rule had been set after careful, but secret, research. The three stories chosen were those that could be credibly adapted to meet these constraints. The relevant excerpts of the arbitrary and non-arbitrary version of the Movies vignette are presented below. The relevant rule was that students could not lend particular movies to any student under age 15 due to their gory content (for full stimulus materials, see Supplementary Material 6):

> *Arbitrary*: "The school administrators chose the age of 15 based on the recommendation from the studio. The studio chose the age of 15 somewhat arbitrarily. There isn't any reason why seeing the gory content would harm a 14.5-year old alien much more than a 15.5-year old alien, but they had to choose some age cutoff, and decided on 15.
> Neither the school administrators nor the students have any idea how the studio chose the age and believe the age was selected arbitrarily. In fact, Bernice once asked a school administrator why the age was 15, and was told that it was an arbitrary cut-off determined by the movie studio."
> *Non-Arbitrary*: "The school administrators chose the age of 15 based on the recommendation from the studio. The studio chose the age of 15 after careful research on alien development. Aliens of their species undergo an important developmental change such that younger aliens are susceptible to the negative effects of gory and violent content, while older aliens are not. The developmental change is linked to an increase in hormone levels that occurs, like clockwork, at age 15. The studio feared that releasing this

information would damage their sales and reputation so they have carefully kept it from the public. Only the top executives and a few scientists even know the study was conducted. Neither the school administrators nor the students have any idea how the studio chose the age and believe the age was selected arbitrarily. In fact, Bernice once asked a school administrator why the age was 15, and was told that it was an arbitrary cut-off determined by the movie studio."

Participants were randomly assigned to read one of the six vignettes. As in Experiments 1-3, they read an *unknowing* version followed by wrongness and punishment questions in a randomized order, and then answered the questions again after seeing a knowledge change prompt similar to those used in Experiments 1-3. Next, participants were shown the rule change manipulation from Experiments 1 and 2. We used the rule change from Experiments 1 and 2, in case the absence of a significant rule change effect for wrongness in Experiment 3 stemmed from the way the question was asked. Finally, participants answered catch questions as in previous experiments.

Results

Knowledge Dependence

If the differential effects of knowledge and intent found in Experiments 1-3 were driven by differences across conditions in the mental states attributed to transgressors, then we should not find differential knowledge effects here: in all cases, the transgressor had the belief that a given norm was arbitrary. To test this prediction, we performed mixed ANOVAs with rule type (2: arbitrary, non-arbitrary) as a between subjects' factor, knowledge state (2: *unknowing*, *knowing*) as a within subjects' factor, and either wrongness or punishment (service hours) as the dependent variable. Consistent with our prediction, we did *not* find an interaction between rule type and knowledge state for either wrongness, $F(1,238) = 0.770$, $p = .381$, $\eta_p^2 = .003$, or punishment, $F(1,238) = 1.24$, $p = .267$, $\eta_p^2 = .005$ (see Figure 4-3). There was, unsurprisingly, a main effect of knowledge state for both wrongness, $F(1,238) = 440.62$, $p < .000$, $\eta_p^2 = .649$, and punishment, $F(1,238) = 360.62$, $p < .000$, $\eta_p^2 = .602$.

To compare the results of Experiments 3 and 4 directly, we ran an additional analysis in which we treated experiment (Experiment 3 vs. Experiment 4) as a between-subjects variable. To make the two experiments as comparable as possible, we included only the three vignette pairs corresponding to those used in Experiment 4. We then performed 2 (experiment: 3, 4) x 2 (basis for rule: arbitrary, non-arbitrary) ANOVAs with the knowledge effect difference score as the dependent variable (i.e., rating for *knowing* transgression minus rating for *unknowing* transgression). This analysis revealed a significant interaction for punishment (service hours), $F(1, 356) = 5.40$, $p < .021$, $\eta_p^2 = .015$, and a marginally significant interaction for wrongness, $F(1, 356) = 3.11$, $p <= .079$, $\eta_p^2 = .009$. In both cases, the knowledge effect was greater in Experiment 3 than in Experiment 4, which supports the idea that it's an agent's beliefs about the harm associated with rule breaking, and not the actual harm caused, that partially or wholly determines differential effects of mental states across types of rule breaking.

Figure 4-3: Ratings for wrongness and punishment (service hours) for all vignette variants as a function of transgression type. Error bars correspond to one SEM in each direction.
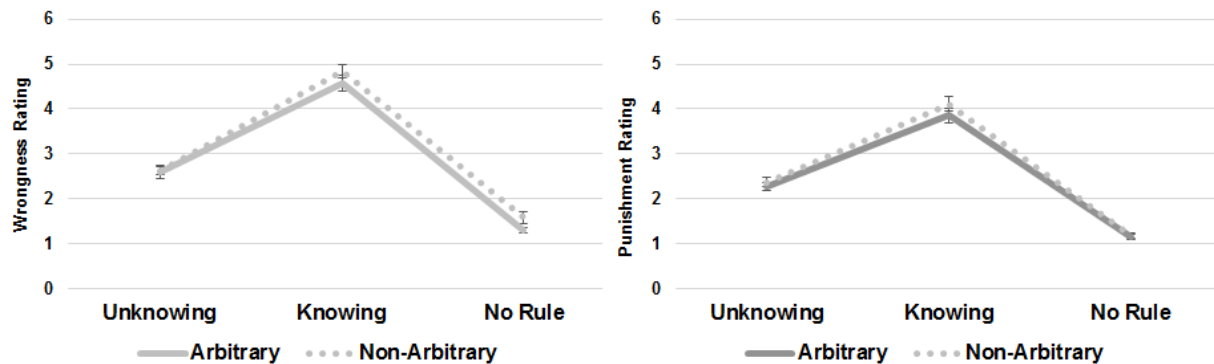
Rule Dependence

Difference scores were created as in previous Experiments to test for rule dependence. We did not find any significant difference between the arbitrary and non-arbitrary conditions (all p's > .35).

Discussion

Experiment 4 supports our prediction that it is an actor's *belief* that a rule is arbitrary that participants use to evaluate wrongness and ascribe punishment, not the actual consequences of transgressing an arbitrary or non-arbitrary rule: the differential knowledge effects across conditions completely disappeared when the actor's knowledge was held constant across different realities. This finding supports our broader contention that what drives the knowledge effect in moral and conventional (or non-arbitrary and arbitrary) transgressions is an actor's knowledge and intent concerning his or her actions.

Experiment 4 also supports Experiment 2 in suggesting that the presence, absence, or degree of *actual* or *potential* harm are not what really drive our knowledge effects. If harm were driving the differential knowledge effect, we would have expected to see a greater knowledge effect when a transgression violated a rule that was in fact non-arbitrary relative to cases in which the rule was in fact arbitrary, regardless of the perpetrator's beliefs. This, along with the careful matching of harm in Experiment 2, provides compelling support that it's a transgressor's mental states that are predominantly responsible for the magnitude of the knowledge effect.

General Discussion

In Experiments 1 and 2 we found evidence that the evaluation of moral transgressions is more knowledge dependent than the evaluation of conventional transgressions – that is, "knowing what you're doing" is more damning, and ignorance is more mitigating, when the rule that's violated is moral as opposed to conventional. We also found that the evaluation of conventional transgressions is more rule dependent than that of moral transgressions, replicating findings from past literature. In Experiments 2-4, we found evidence that part of what drives differential effects of knowledge and intent is the typically non-arbitrary nature of moral rules, and especially a transgressor's *beliefs* about the basis for a violated rule.

The finding that relative to moral transgressions, the perceived severity of conventional violations is contingent on the presence of a rule is not new; however, to our knowledge, this is the first demonstration that judgments concerning the wrongness of moral transgressions are more sensitive to knowledge and intent than those concerning conventional transgressions, with corresponding effects for punishment. That said, our findings nicely complement recent work with children by Josephs, Kushnir, Grafenhain, and Rakoczy (2016), who found that another aspect of an actor's actions – whether they were chosen freely or under constraint – had a greater impact on children's protests concerning violations of moral versus conventional rules. Our study differs from this recent work in manipulating actors' mental states directly (rather than manipulating constraints on their choices), and in considering explicit judgments of wrongness and punishment (rather than children's "protests"). Our study is thus the first to systematically vary whether a transgressor "knew what she was doing" in violating a rule, and to do so across moral and conventional cases.

Despite the novelty of our result, the relationship between knowledge dependence and the conventionality of a rule is consistent with prior work on the evaluation of strict liability crimes. In Giffin and Lombrozo (2016), we argued that knowledge is less important in laypeople's judgments concerning strict liability crimes because such crimes tend to involve the violation of a rule with somewhat arbitrary – and therefore arguably "conventional" – elements. For example, speeding is a strict liability crime, and it involves the violation of a somewhat arbitrary speed limit. Driving 40 miles per hour is not inherently wrong, but it is wrong when it occurs in a 35-mile zone, and the designation of 35 miles (as opposed to 34 or 36.5) as the limit is somewhat arbitrary. It is not clear that the harm or consequence is much different from 35 to 40 or 40 to 45, even though there's a good reason for specifying *some* speed limit. In other words, the rule is arbitrary in the sense that it could reasonably have been specified differently. This feature of "arbitrariness" could help explain why strict liability crimes and conventional violations behave similarly.

Experiment 2 found support for an association between arbitrariness and knowledge dependence, and Experiment 3 set out to manipulate the role of arbitrariness directly. We hypothesized that if part of what makes conventional violations less sensitive to a transgressor's knowledge and intent is their more arbitrary specification, then rules that are expressly characterized as somewhat arbitrary – even when they involve harm – should also show less sensitivity to these mental states. This is precisely what we found: judgments concerning the violation of arbitrarily set rules depended less strongly on the actor's knowledge and intent, whether participants were evaluating the severity of wrongdoing or the number of service hours that should be required as punishment. Finally, in Experiment 4 we found evidence that what matters is the transgressor's *belief* that a rule has been set arbitrarily; not necessarily the actual consequences of breaking a rule that involves an arbitrary threshold.

Our findings raise a number of important questions for future research. First, how do our findings relate to prior work (Barrett et al., 2016; Chakroff et al., 2015; Hawley-Dolan & Young, 2013; Russell & Giner-Sorolla, 2011; Young & Saxe, 2011), which documents a weaker effect of knowledge and intent in evaluating purity violations, such as incest, relative to harm violations, such as battery? Like conventional transgressions, purity violations differ from moral transgressions in the primary locus of harm. In a typical harm violation, such as poisoning, the victim is another person (Gray & Wegner, 2009, 2012; Gray, Young, & Waytz, 2012). This is also the case for the moral transgressions considered in our experiments. However, in the case of purity violations, such as incest, the victim is often the self (Young & Tsoi, 2013; Young &

Saxe, 2011). Conventional transgressions could mirror purity violations not in the focus on self per se, but in lacking another person as an identifiable victim. Supreme Court Justice Morissette argued that legal wrongs that are *malum prohibitum* are offenses against *the authority of the state*: even in the absence of certain harm to others, disregarding convention is a potential harm to society or social order. Individuals may be harmed downstream, but they are not typically identifiable as individuals at the time of the transgression. It could be that the central importance of knowledge and intent in moral transgressions stems from the role of an identifiable victim other than the self. In such cases, it may be especially important to track the mental states of the perpetrator to evaluate moral character and prevent future harms. Potentially consistent with this line of thought, Josephs, Kushnir, Grafenhain, and Rakoczy (2016) explain their results by appeal to the idea that violations of moral norms involve a focus on the *actor*, whereas violations of conventional norms shift the focus from the actor to the *consequences* of the violation.

Second, in what sense must rules be arbitrary for the role of mental states to be attenuated? If our account is right, then the critical factor is whether a rule is linked to intrinsic harm, and thus constrained such that it couldn't reasonably have been otherwise. In our experimental manipulation of arbitrariness (Experiments 3-4), we created rules that were arbitrary because they set a threshold that could reasonably have been otherwise. It was useful to adopt this approach to arbitrariness because all vignettes were set in the moral domain and involved some level of harm. It's worth noting, though, that this differs from the way in which conventional rules are often arbitrary: rather than identifying a critical threshold, many specify conditions for group coordination (e.g., driving on the right rather than the left). We expect that comparable effects would be found in such cases, but this is at present an untested prediction. A further question concerns the relationship between conventionality and arbitrariness and whether all forms of conventionality involve some inherently arbitrary elements.

Third, might there be important boundary conditions on the effects we report here? It's instructive to consider conventional transgressions accompanied by truly bad intentions. For example, consider a person who intentionally drives on the wrong side of the road, not only with the intention of violating traffic rules, but also with the intention to generate a harmful collision. Although the anticipated harm stems from a violation of convention, the addition of the intention to cause harm seems to shift the offense from the domain of the conventional to the domain of the moral: the primary offense isn't in violating rules of traffic, but in violating rules against aggravated battery or perhaps vehicular homicide. We take it as evidence for our distinction between conventional and moral transgressions that it is so difficult to construct instances of purely conventional transgressions that involve genuinely bad intentions. It appears that once an action is accompanied by genuinely bad intentions, it shifts domains from a conventional to a moral transgression.

Fourth, how might culture moderate the effects we report? Prior work has found that the relative importance of intent in moral evaluations can differ across cultures (Barrett et al., 2016). Moreover, there is documented cultural and individual variation in the perceived boundary between moral and conventional wrongdoing (Hauser, Cushman, & Young, 2008). Our vignettes are set in a modern, westernized context, and our sample is from the United States. It is possible that cultures that exhibit an attenuated role for mental states in moral evaluation will show even less sensitivity when it comes to conventional violations. Investigating other cultures will be crucial in painting a more accurate and complete picture of human moral psychology.

Finally, how might our results from Experiments 3-4 translate to more "earthly" situations? We chose to situate Experiments 3-4 in an alien world because we feared that

participants' beliefs about how arbitrary – or not – a rule is on earth would be relatively hard to manipulate. However, to the extent we could convincingly state that our actors reasonably believed a real moral rule was set arbitrarily, we would expect to see a reduction in the influence of mental states on judgments. Precisely which mental states matter and why, however, is another question that merits further research.

In sum, our findings are consistent with prior work demonstrating the importance of mental states in moral judgment. However, our findings go beyond prior work in pointing to a differential role for knowledge and intent when it comes to evaluating moral versus conventional transgressions. We also suggest a basis for this difference: to the extent a rule is believed to reflect non-arbitrary considerations, the intention to violate the rule will be bound up with intentions to cause other consequences, such as harm. This makes a knowing transgression more severe, and ignorance more mitigating, for moral transgressions than for conventional transgressions.

Chapter 4

Mind Reading Strategies Predict Outcomes for Defendants with, and without, Mental Illness

Abstract

Jurors often have to infer what a defendant felt and believed: Did he *really* fear for his life? Did he *truly* believe he was in danger? Determining the mental states of others is referred to as mind reading, and it is supported by various strategies, including projection (using oneself as a base for estimation) and stereotyping (using a stereotype about the other's social group as a base). In three experiments involving online participants evaluating descriptive vignettes, we explored the impact of mind reading strategies on legal decisions about defendants with and without mental illness. Experiment 1 (N =480, 235 female, 240 male, 5 other; mean age = 36) found that participants used less projection and more stereotyping when evaluating defendants with a mental illness, translating into more guilty verdicts. In Experiment 2 (N =640, 352 female, 286 male, 2 other; mean age = 37), we found that a short prompt to put oneself in the other's position increased levels of projection and modestly attenuated punishment. In Experiment 3 (N =480, 280 female, 199 male, 1 other; mean age = 37), we found that providing insanity standards (either M'Naghten or MPC) led to more guilty verdicts than a condition in which no standard was given. These findings suggest an important role for mind reading strategies in the law, as well as a simple way to direct these strategies. They also suggest that defendants with a mental illness are evaluated differently, and treated more harshly, than defendants who commit comparable crimes.

keywords: mind reading; law; decision making; punishment

Introduction

Imagine the following scenario: two people get into a fight that begins with words, but quickly escalates into a physical altercation. There is bad behavior on both sides, but one person throws the first punch, and it is this person who is later arrested for the crime of battery. At trial, the defendant argues that he punched the other man because he believed the man posed a threat and could injure him.

Should the defendant be found guilty of battery? Should he be fined, or serve time in jail? If you are a juror, the way you evaluate the defendant is likely to depend crucially on how you reason about his mental states. Did he really *believe* that the other man posed a threat? Did he experience genuine *fear*? Under the circumstances, was it *reasonable* for him to have these beliefs and emotional reactions? Would *you* have thought and felt the same way?

Researchers have studied people's ability to determine what others think, intend, believe, and feel for years. This is often referred to as mind reading (Eyal & Epley, 2010), but it is not a power found in science fiction: it refers to the everyday activity of inferring and evaluating the mental states of those around us. The literature on this topic is vast (for instance, see Epley, 2014; Carrington & Bailey, 2008; Imuta, et al., 2016; Henry et al., 2013), but it has overwhelmingly focused on non-legal contexts. From a different tradition, research on juror decision making has revealed that a juror's mental representations are crucial to subsequent judgments, but this research has not investigated the role of mind reading per se. Pennington and Hastie (1986), for example, argued that jurors represent evidence as stories, and they elucidated some of the factors that impact story construction and verdicts – such as the coherence of the story (Pennington & Hastie, 1992). While Pennington and Hastie did not directly test the importance of mind reading, they did note that 45% of the coded interviews were inferences that the participants made from the material that was provided, and some of these inferences involved or followed from the posited mental states of the key actors (Pennington & Hastie, 1986). From the perspective of the story model, it is therefore highly plausible that whether and how a juror engages on mind reading will affect the story she constructs and the verdict she ultimately reaches, though this avenue of research has not been pursued.

In the current paper, we bring these two research traditions together by investigating the role of mind reading in a legal context. Specifically, we investigate whether and how different *strategies* for mind reading affect people's intuitive judgments about whether a defendant should be found guilty and how he should be punished. Given that this is among the first efforts to systematically manipulate mind reading in legally-relevant judgments, we begin by adapting the methods of social psychological research, using descriptive vignettes that allow us to exert considerable experimental control over individual participants. We also focus on two mind reading strategies that have been well documented in prior psychological work, projection and stereotyping (Ames, 2005), and we consider how they are deployed in reasoning about a defendant who does or does not have a mental illness.

Projection and stereotyping are common mind reading strategies that could plausibly describe how a juror reasons about a defendant. When engaged in projection, a person uses him or herself as a basis from which to infer the likely thoughts, feelings, and behaviors of another (Ames, 2004a & 2004b). For example, in trying to predict whether another person would find a particular situation frightening, someone might imagine whether she herself would find it frightening. By contrast, stereotyping is a mind reading strategy in which a person uses a social

group that the other belongs to as a basis for inference (Fiske, 1998, Fiske & Neuberg, 1990). For instance, in trying to understand why a particular woman stopped to help a stranger, an observer might appeal to his stereotypes about women (in general) as nurturing, and thus likely to have intentions and desires to help.

Studies have found evidence that both of these processes occur, with their relative prevalence affected by a number of factors. One key factor is the perceived similarity of the target – that is, how similar we think the other person is to ourselves. The more similar we believe the other person is to ourselves, the more likely we are to use projection (Ames, Weber, & Zhou 2012; Ames, 2004a). This is true even when we do not particularly like the other person (Davis, 2017). Intuitively, this makes sense: the more similar we believe another person is to ourselves, the more we should consider ourselves to be an appropriate model for the other's mind and behavior. Importantly, it is *perceived* similarity, not actual similarity, that matters. A person must only *believe* that another person is similar to themselves in order to increase the likelihood that he or she will use projection rather than stereotyping (Ames, Weber, & Zhou, 2012).

One reason mindreading strategy is important is because it has been associated with different forms of treatment. For instance, Batson and colleagues (2003) asked participants to assign two tasks – one pleasant and one boring – to themselves and to another participant they had not met, and they were given a coin to flip to help them make this decision. Prior to assigning tasks, however, participants were asked to either imagine themselves in the other's place, to imagine the other person's feelings, or given no prompt at all. The researchers found that those participants who imagined themselves in the other's place were more likely than participants in the other two groups to flip the coin and abide by the flip, while those who imagined the other's feelings were more likely than those in either other group to assign the other the more pleasant task outright (Batson et al., 2003). Other studies have found that taking another person's perspective leads to greater feelings of perceiver-target "overlap" (similarity), and thereby to more positive assessments and intentions to help them (Laurent & Myers, 2011; Manner et al., 2002). These findings suggest that engaging in projection could lead to more lenient treatment in a legal context, but this hypothesis has yet to be tested.

Prior work also suggests that differences in mental state attributions can affect moral and legal judgments, but this work has not measured or manipulated mind reading strategy. For instance, Giffin and Lombrozo (2016) found that the beliefs and intent attributed to a defendant had a significant impact on judgments concerning a variety of crimes, including battery and theft. Specifically, ignorance that some act constituted a crime (because, for example, the defendant erroneously and unreasonably believed he was acting in defense of another) was mitigating, at least for crimes that are not strict liability. More generally, there's evidence that mental state attributions affect judgments of blame, moral responsibility, and punishment (e.g., Cushman, 2008; Young & Saxe, 2011). Together, this prior research suggests that mind reading strategy could affect the mental states that one attributes to a defendant, and that mental state attributions could in turn affect legally-relevant judgments.

What specific predictions might we make about how mind reading strategies influence lay people's legal judgments? First, based on the prior literature just reviewed, we might expect projection to be associated with more lenient treatment, as reflected in fewer verdicts of guilty and more modest sentences. Second, we might expect similarity to the defendant to moderate the use of projection, such that dissimilar defendants are less likely to receive the benefits of

projection. It is reasonable to assume that most decision-makers consider themselves fairly dissimilar to someone who is standing trial for a crime. While jurors might perceive themselves to be dissimilar from any person accused of a crime, dissimilarity still comes in degrees. A defendant with a mental illness (that is not shared by the juror) could be regarded as particularly dissimilar, making the juror especially unlikely to engage in projection. In the legal system, this could mean that defendants receive harsher treatment based on something as legally irrelevant as *how* a decision-maker assesses their beliefs and intentions.

To test these ideas, our studies involve hypothetical defendants who do or do not have a mental illness. We anticipate that defendants with a mental illness will (on average) be regarded as less similar and therefore prompt less projection. We also anticipate that defendants with a mental illness will trigger more stereotyping, as stereotypes about mental illness are prevalent, if often inaccurate (Corrigan et al., 2012: Torrey, 2011; Angermeyer & Matschinger, 2004; Phelan, 1998; Monahan, 1992; Steadman, 1981). For example, Angermeyer and Matschinger (2004) found that 19% of participants agreed that "the number of violent crimes could be markedly reduced" if all people with schizophrenia were put in locked wards, and 36% believed that people with schizophrenia posed "a great threat to small children." While our intention is not to identify some mind reading strategies as better or more appropriate than others, widespread use of stereotyping – especially when those stereotypes are inaccurate – could have legal consequences that are important to recognize.

A final question is how – or whether – jury instructions on insanity standards have any impact on mind reading strategy, on laypeople's verdict choice, or on their judgments concerning appropriate levels of punishment. Prior research comparing different insanity standards has found little evidence that different standards produce different results (Olgoff, 1991; Finkel et al., 1985; James, 1959). While this remains the common wisdom of legal scholarship, few studies address the impact of insanity standards directly, and none have considered potential interactions with strategies for mind reading.[22] Introducing a legal standard could in fact have an unintended effect by focusing attention on the defendant's mental illness, thereby decreasing perceived similarity, reducing projection, and increasing stereotyping, potentially resulting in less favorable outcomes. On the other hand, providing a legal standard could also remind participants that in the legal system, mental illness information is presented as mitigation.

Across three studies, we test these ideas by presenting participants with a hypothetical case of battery involving a defendant with or without schizophrenia. Our studies are designed to test the following hypotheses. First, we hypothesize that participants will engage in less projection for a defendant who has a stated mental illness than for one who does not, even when their actions were identical, because they will perceive these defendants to be less similar to themselves (Hypothesis 1a). We also make the complementary prediction that stereotyping will be significantly higher when the defendant has a mental illness (Hypothesis 1b). Second, we predict that greater levels of projection will be associated with less punitive outcomes (Hypothesis 2a), while greater stereotyping will lead to more punitive outcomes (Hypothesis 2b). Combining these hypotheses leads to a third prediction, that a defendant with a mental illness will be treated more harshly than a defendant without one (Hypothesis 3).

---

[22] More of the studies address laypeople's feelings towards the insanity defense and their personal, quite detailed concept of insanity (i.e. Skeem & Golding, 2008; Finkel & Handel, 1989).

Our final hypotheses concern interventions that could potentially affect the treatment of defendants with a mental illness. First, we speculate that prompts to engage in projection could increase the extent to which participants use projection when reasoning about a defendant with a mental illness, thereby resulting in more lenient treatment (Hypothesis 4). While prior work has not – to our knowledge – investigated how interventions on mind reading strategies affect legal judgments, Weiner and Hurt (2000) found that women who were asked to use a "reasonable woman" as opposed to "reasonable person" standard were more likely to find that conduct met the level of sexual harassment, which suggests some malleability that a shift in mind reading strategy could potentially bring about. Finally, and most tentatively, we predict that providing a legal standard of insanity – regardless of which standard – will increase stereotyping, thus resulting in harsher treatment (Hypothesis 5). While certainly vignette studies such as those reported here lack important components of ecological validity, they are a first step towards understanding phenomena with both theoretical implications for psychology and practical implications for the law.

Experiment 1

The primary purpose of Experiment 1 is to test hypotheses 1-3. Will projection be lower, and stereotyping higher, when the defendant has a mental illness (Hypotheses 1a and b)? Will higher projection lead to more lenient treatment, and higher stereotyping to less lenient treatment (Hypotheses 2a and b)? Finally, will defendants with a mental illness be treated more harshly for committing identical crimes (Hypothesis 3)? If, as we predict, projection will be lower and punishment higher for defendants with a mental illness, we can assess whether perceived similarity explains these differences.

To answer these questions, we compare three hypothetical defendants who have all performed identical actions, and who all plead not guilty. Across conditions, however, we vary the explanation offered for their behavior. In the *Neurotypical* condition, the defendant argues that the other man started the altercation, and that he felt threatened. In the *Labeled* condition, the defendant similarly argues that he felt threatened, but participants additionally read that he has been diagnosed with schizophrenia and that he suffers from certain symptoms. Finally, in the *Unlabeled* condition, the defendant again argues that he felt threatened, and participants read that he suffers from the same symptoms as in the *Labeled* condition, but the word "schizophrenia" is never used.

We chose schizophrenia because it is a mental illness about which people have strong, though often inaccurate, stereotypes (Angermeyer &Matschinger, 2004). It is therefore a good candidate for investigating the role of stereotyping. We included both the *Labeled* and *Unlabeled* conditions because prior research suggests that labeling a category can affect how it is treated (Lebowitz & Ahn, 2014; Giffin, Wilkenfeld, & Lomborzo, 2017). By including conditions that differed only in the provision of a label, we can test whether the label itself – schizophrenia – is responsible for any observed differences, or whether possession of an unidentified mental illness with equivalent symptoms is sufficient to generate the differences that we predict.

Methods

Participants. Four-hundred-and-eighty adults (235 female, 240 male, 5 other/prefer not to specify, mean age = 36, *SD* = 13) participated in the study through Amazon Mechanical Turk

for monetary compensation. Sample size was determined based on a power analysis of a pilot study. An additional 54 participants were tested, but were excluded for failing attention checks (explained below). In all studies, participation was restricted to workers with an IP address within the United States and a prior approval rating of 95% or higher on previous tasks. In addition to restricting our workers' IP addresses, all participants indicated that they were United States' citizens. Thus, our participants met the basic requirements for jury service: United States' citizens over the age of 18.

      Materials & Procedures. The experimental stimuli consisted of three vignettes to which participants were randomly assigned, corresponding to the *Neurotypical*, *Labeled*, and *Unlabeled* conditions. All three vignettes described the same criminal action, an act of battery, but concluded by offering different explanations for the action. The initial portion of each vignette was thus identical across conditions, and read as follows:

> "John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

<div align="center">—Page Break—</div>

> One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over.
>
> John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

<div align="center">—Page Break—</div>

> As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.
>
> The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John sung at him."

      In all cases, participants were told that the police found a security tape that confirmed the facts as presented. This information was added to mitigate possible concerns that participants might believe a defendant with a mental illness was not accurately recalling the facts due to his mental illness. Also note that in all cases, the defendant had a false belief: that the other man's cellphone was a gun. This was to ensure that the defendant had a false belief in *all* conditions, not only in the case of mental illness. This false belief also provided an identical claim of self-defense for each defendant. In each condition, John believed, based on identical, confirmed facts that the man was going to injure him.

The final paragraph of each vignette offered an explanation for John's actions, where the explanation differed across the *Neurotypical*, *Labeled*, and *Unlabeled* conditions. The relevant text is excerpted below; the *Labeled* and *Unlabeled* conditions varied in the underlined clause, which was omitted in the *Unlabeled* condition (for full stimuli, see Appendix A).

> *Neurotypical*. "At John's trial, his lawyers plead not guilty. John's lawyer argued that John was trying to defend himself, saying that the altercation was the result of the other man's actions. John testified that at the time of the fight, he believed the other man posed a threat and could injure him."
>
> *Labeled / Unlabeled*. "At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John <u>had been diagnosed with schizophrenia, explaining that he</u> suffers from abnormal thoughts and feelings. John testified that at the time of the fight, he believed the other man had been sent by his enemies to track and injure him."

After reading their assigned vignette, participants first selected a verdict: either "guilty" or "not guilty." Participants were then asked "If a jury did find John Guilty, what kind of punishment should John receive?", and could select a monetary fine, jail time, counseling, and/or community service. Thus, even participants who selected "not guilty" selected a punishment option. For each kind of punishment that a participant selected, a follow-up question asked them to specify an amount:

> "How much [money/time] should John be required to [pay/spend in jail/spend in counseling/perform community service]?"

Answers were indicated on sliding bars that went from zero to twice the legally applicable amount. In the case of counseling, it went up to 24 months.[23]

After answering questions about verdicts and sentences, participants saw a new screen with the following three questions about blame and punishment presented in a randomized order:

> *Blame*. "How much blame do you think John deserves for his actions?" 1 (None at all) to 7 (A great deal).
>
> *Deterrence*: "Punishing John is important because it will deter him from committing this crime in the future." 1 (Strongly Disagree) to 7 (Strongly agree).
>
> *Incapacitation*: "Imprisoning or monitoring John is important because it protects the public from him while he is incarcerated." 1 (Strongly Disagree) to 7 (Strongly agree).
>
> *Retribution*: "Even if it won't prevent John from committing this crime in the future, he deserves to be punished for this action." 1 (Strongly Disagree) to 7 (Strongly agree).

On a separate screen, participants then answered questions designed to assess the strategies that they employed in reasoning about John. These questions were prefaced by a short

---

[23] Although jurors would not be asked to render sentences in an actual trial, allowing our participants to provide sentences gives us more insight into how they are reasoning about defendants and verdicts. For example, a guilty verdict associated with only a week of community service is importantly different from a guilty verdict with three months of jail in terms of how the participant is viewing the defendant and the implications of a guilt determination.

sentence: "If you were trying to predict what John would do in another situation, you would rely on:" Participants rated all the following on a scale from 1 (Not at all) to 7 (A great deal).

> *Individual Information*. "What I know about how John thinks."
> *Projection*. "What I would do if I were in that situation."
> *Stereotyping*. "What I know about people like John."
> *Reasonable Person*. "What another reasonable person would do in that situation."

Next participants answered the following question on its own screen:

> *Confidence*. "How confident are you that you could predict John's future behavior?" 1 (Not at all) to 7 (Very).

The last substantive questions participants answered assessed similarity. These questions were modeled on prior work (Ames, 2004a; 2004b). The second was reverse scored, and the two questions were averaged together into one composite measure of similarity:[24]

> *Similar to John*. "I am a similar person to John." 1 (Strongly Disagree) to 7 (Strongly agree).
> *Similar Friends*. "The people I identify with are a lot different from John." 1 (Strongly Disagree) to 7 (Strongly agree).

Finally, on a separate screen, participants answered two questions designed to assess whether they had been paying attention and reading questions carefully. All participants answered a true/false question regarding the content of the vignette they had just read and a reading comprehension question designed to test whether they carefully read instructions (Oppenheimer, Meyvis, & Davydenko, 2009). The former attention check was placed at the end of the task on the assumption that if participants remembered these details at the end, it is safe to assume they remembered them while answering substantive questions. Those who answered either question incorrectly were excluded from further analysis. Finally, participants answered demographic questions about age and gender.

Results

Across our measures, the *labelled* and *unlabeled* conditions never differed significantly. As a result, we collapse them in our analyses into a single category of *Mental Illness*, leaving us with two between-subjects conditions: *Neurotypical* and *Mental Illness*.

Verdict. In the *Neurotypical* condition, 28% of participants selected a guilty verdict. In the *Mental Illness* condition, 43% of participants did so. A 2 (defendant condition: *Neurotypical*, *Mental Illness*) by 2 (verdict: not guilty, guilty) chi-square test revealed that this difference was significant, X^2 (1, $N = 480$) = 10.52, $p < .001$.

Sentences. To analyze sentences, we treated the omission of some form of punishment as equivalent to a value of "0" (i.e., a fine of $0, zero jail time, zero community service, zero hours

---

[24] Analyses using just one or the other of these questions separately show the same patterns of results.

of counseling). Because the resulting distributions of punishment levels were not normal, we used (non-parametric) Mann-Whitney tests to evaluate the effect of defendant condition on levels of assigned punishment (see Table 1-4). Participants assigned marginally higher fines to the *Neurotypical* defendant, $U = 23022.50$, $p = .052$, but significantly more jail time, $U = 24339.00$, $p < .039$, and counseling, $U = 9369.00$, $p < .001$, to the defendant with a mental illness. Community service hours did not differ across conditions, $U = 25350.50$, $p = .860$.

Table 1-4: Means and standard deviations for punishment dependent variables for Experiments 1 and 2. Significant differences across defendant conditions are indicated in bold. Fines were measured in dollars with a maximum of $4,000, jail time and community service were measured in days with a maximum of 365, and counseling was measured in months with a maximum of 24.

| | Experiment 1 | | Experiment 2 | |
	Neurotypical	*Mental Illness*	*Neurotypical*	*Mental Illness*
Fines	406 (622)	376 (660)	**336 (652)**	**271 (569)**
Jail time	**1.89 (13.4)**	**3.08 (15.4)**	**4.18 (28.0)**	**6.55 (31.6)**
Community service	37.9 (47.8)	46.4 (63.6)	33.1 (45.6)	35.9 (51.1)
Counseling	**1.68 (3.69)**	**9.17 (8.03)**	**1.68 (3.65)**	**9.62 (7.76)**

Mind Reading Strategies. To assess whether participants adopted different strategies for reasoning about each defendant, we conducted *t*-tests comparing ratings for each judgment strategy question as a function of defendant condition (*Neurotypical*, *Mental Illness*).

For the neurotypical defendant, participants reported significantly higher levels of projection, $t(477) = 3.61$, $p < .001$, $d = .33$, reliance on a "reasonable person," $t(478) = 2.74$, $p < .006$, $d = .25$, and similarity, $t(478) = 6.28$, $p < .001$, $d = .57$. For the defendant with mental illness, participants reported significantly higher levels of stereotyping, $t(478) = 2.24$, $p < .026$, $d = .20$. Additionally, as predicted, similarity was significantly higher for the *Neurotypical* condition, $t(478) = 6.28$, $p < .001$, $d = .57$. No other significant differences were found (see Figure 1-4).

Figure 1-4: Mind reading variables by defendant condition. Error bars are one SEM in each direction. All differences are significant.
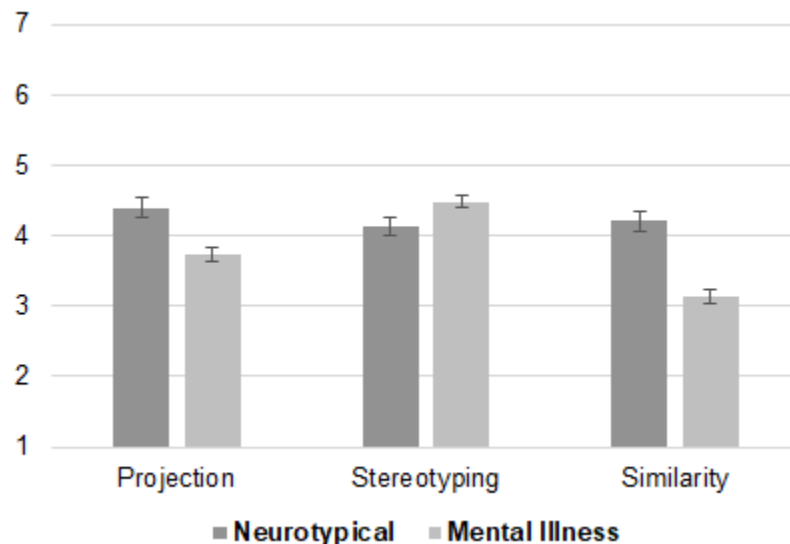
Previous research has proposed similarity as a key moderator of the use of projection, with higher levels of similarity leading to greater projection (Ames, 2004b). We conducted a Sobel test to determine whether similarity mediated levels of projection across our *Neurotypical* and *Mental Illness* conditions, and found a significant result, $t = -4.24, SE = 0.07, p < 0.001$, with partial mediation.[25]

Blame, Punishment Justifications, & Confidence. We analyzed each of our blame, justification for punishment, and confidence questions with *t*-tests comparing mean ratings as a function of defendant condition (*Neurotypical*, *Mental Illness*; see Table 2-4).

For the defendant with mental illness, participants reported that the defendant deserved significantly greater blame, $t(478) = 3.89, p < .001, d = .36$, and that punishment was more warranted on the grounds of incapacitation, $t(478) = 6.49, p < .001, d = .59$, retribution, $t(478) = 3.05, p < .002, d = .28$, and deterrence, $t(478) = 2.44, p < .015, d = .22$. For the neurotypical defendant, participants reported significantly greater confidence that they could predict John's actions, $t(478) = 3.70, p < .001, d = .34$.

Table 2-4: Blame, Punishment Rationale, and Confidence Means by Defendant Condition. Means and standard deviations for blame, punishment justification, and prediction confidence questions in Experiment 1. Significant differences across defendant conditions are indicated in bold.

	Neurotypical	Mental Illness
Blame	**3.48 (1.53)**	**4.05 (1.52)**
Incapacitation	**1.97 (1.43)**	**2.98 (1.69)**
Retribution	**3.24 (1.82)**	**3.78 (1.82)**
Deterrence	**3.05 (1.79)**	**3.46 (1.68)**
Confidence	**3.71 (1.54)**	**3.18 (1.42**

Does mind reading strategy predict verdict? The findings reported so far reveal that participants did judge defendants with a mental illness significantly differently from those without a mental illness, more often rendering verdicts of guilty. We also found evidence that participants used different mindreading strategies: they relied more heavily on stereotyping for the defendant with a mental illness, and more heavily on projection (partially mediated by similarity) for the neurotypical defendant. Our next step was to determine whether these finding were related by testing whether mindreading strategies predicted verdicts. To do so, we ran a binary logistic regression predicting verdict (1=not guilty, 2=guilty) as a function of projection and stereotyping. We found that projection was a significant, negative predictor of verdicts, $\beta = -.184, p < .001$, while stereotyping was a positive, but non-significant predictor, $\beta = .058, p = .341$. These results indicate that the more a participant used projection, the less likely she was to choose a guilty verdict. To test whether projection mediated the effect of defendant condition on

---

[25] A linear regression with defendant condition and similarity entered stepwise as predictors revealed that similarity partially, but not fully, mediated the effect of defendant condition on projection. In the first step, defendant condition was a significant predictor of projection, accounting for 3% of the variance, $F(1, 478) = 13.04, p < .001, R^2 = .027$. In the second step, the addition of similarity was a significant improvement, accounting for 7% of the variance, $F(1, 476) = 19.09, p < .001, R^2 = .074$; R^2 change $= .048, p < .001$.

verdict, we conducted a Sobel test, which revealed partial mediation, *t = 2.51, SE = 0.05, p = 0.012.*[26]

Discussion

Experiment 1 found that participants rendered different verdicts and assigned different levels of punishment for an act of battery when the defendant was described as having a mental illness, whether the mental illness was or was not labelled as schizophrenia. The findings confirmed most, but not all, of our predictions.

First, consistent with Hypotheses 1a and 1b, we found that projection was higher for the *Neurotypical* condition, while stereotyping was higher for the *Mental Illness* condition. Further, we found that similarity partially mediated this difference in projection. Second, consistent with Hypothesis 2a, we found that higher projection was associated with more verdicts of not guilty. The corresponding effect for stereotyping (Hypothesis 2b) was not significant. Finally, supporting Hypothesis 3, we found that participants judged the defendant with a mental illness more harshly: they were more likely to render a verdict of guilty, and assigned significantly more jail time. We also found that projection partially mediated the effect of defendant condition on verdicts.

One important limitation of Experiment 1 is that it focused on a single criminal act and a single instance of mental illness. It is therefore unclear whether our findings would generalize to a broad range of crimes and to other kinds of mental illness. Before proceeding to test our other hypotheses, we therefore conducted an additional experiment, reported in full in the online supplementary materials as Supplemental Experiment 1 (S1), to test the generality of our effects. Specifically, we tested similar materials but using hypothetical defendants who were diagnosed with Post-Traumatic Stress Disorder (PTSD) or Obsessive Compulsive Disorder (OCD), in addition to schizophrenia or no documented mental illness, as in Experiment 1. We chose to focus on PTSD and OCD because prior work investigating the stereotypes associated with different forms of mental illness has found that PTSD and OCD differ both from each other and from schizophrenia along dimensions of warmth and competence, which play important roles in social judgment (Fisk, Cuddy, & Glick, 2007; Sadler, Meagor, & Kaye, 2012; Sadler, Kaye, & Vaughn, 2015).

As expected, Experiment S1 confirmed that participants' impressions of the warmth and competence of defendants with different forms of mental illness did vary, as did their judgments concerning their unpredictability and the extent to which they could control their behavior. However, defendants with mental illness were treated more harshly than the neurotypical defendant (in some cases significantly, in some cases numerically), regardless of their particular illness, supporting the generality of Hypothesis 3. Further, as predicted by Hypotheses 1a and 1b, projection was significantly lower, and stereotyping significantly higher, for all defendants with a mental illness relative to the defendant without. Finally, supporting Hypotheses 2a and 2b, we also found that greater projection predicted verdicts of not guilty, while greater stereotyping was associated with fewer verdicts of not guilty. Having verified that support for Hypotheses 1-3 is not limited to the conditions tested in Experiment 1, we went on to consider Hypothesis 4.

---

[26] When only defendant condition was included in the regression the Cox and Snell R Squared is .021; but when projection is added in the second step the Cox and Snell increases to .040.

Experiment 2

Experiment 1 provided initial answers to some of the questions posed in the introduction. The mind reading strategies for defendants with and without mental illness do appear to be different: projection dominates in the latter case and stereotyping in the former. Different mind reading strategies also appear to have different implications for verdicts and punishment, with projection being the less punitive strategy.

In Experiment 2, we aimed to replicate the key findings of Experiment 1, and additionally to test Hypothesis 4, which predicts that prompts to engage in projection could increase the extent to which participants use projection when reasoning about a defendant with a mental illness, thereby resulting in more lenient treatment. To do so, we included an experimental manipulation of the instructions that preceded and followed the vignette, with one set of instructions designed to increase levels of projection, and a second set designed to serve as a matched control. These prompts were not designed to be instructions that the judge would give to jurors. Rather, we framed them as a request from John's attorney, language that could easily be inserted into opening and closing arguments. To gain a more nuanced impression of the punitive impulses of our sample, we also offered participants in the *Mental Illness* condition an additional verdict choice: Not Guilty by Reason of Insanity (NGRI).

Methods

Participants. Six-hundred-and-forty adults (352 female, 286 male, 2 other/prefer not to specify, mean age = 37, *SD* = 12) participated in the study through Amazon Mechanical Turk as in Experiment 1. Qualifications were the same as in Experiment 1. An additional 78 participants were tested, but were excluded for failing attention checks (explained below).

Materials & Procedures. The experimental stimuli consisted of two vignettes crossed with two prompts, resulting in four conditions to which participants were randomly assigned. The vignettes were identical to the *Neurotypical* and *Labeled* versions from Experiment 1, with one key exception. In Experiment 1, the *Labeled* version ended by saying that John believed the other man "had been sent by his enemies to track and injure him." This could potentially lead participants to think that John's impressions were still importantly different in the *Labeled* condition, despite our inclusion of the security camera footage to confirm identical actions. Thus, in Experiment 2, both vignettes end with John saying that he "believed the other man posed a threat and could injure him," making the actions *and claims* of the defendant identical in each case. The only aspect that varied was whether John was described as having schizophrenia or not. These defendant conditions will be referred to as *Neurotypical* and *Mental Illness*.

The prompt conditions were either *Project*, which encouraged participants to put themselves in the defendant's place, or *Remember*, which encouraged participants to carefully remember the information presented to them. After giving informed consent, all participants were told: "You are being asked to serve as a juror in John's trial. In his opening statement at the beginning of the trial, John's lawyer makes the following request." Participants then saw the initial prompt for either the *Project* or *Remember* condition.

Project. "As you get the information about John, we would like for you to imagine yourself in John's place. That is, while you are taking in this information, try to imagine you are having the same experiences in the same situation."

Remember. "As you get the information about John, we would like for you to try remember the facts of John's story. That is, while you are taking in this information, try to remember the facts as they are given to you about the experiences and situation."

Participants next read their assigned vignette, and then saw the following statement: "In his closing statement, John's lawyer made the following request. You will be given two minutes to complete this request, and will not be allowed to advance until those two minutes are up." Participants were then reminded of their assigned prompt with the following instructions:

Project. "We would like for you to imagine yourself in John's situation. That is, imagine you are walking down a crowded street on a warm night and someone runs into you. Imagine also how you will feel when that person then shoves you. Take about one minute for this imagination exercise, getting as clear a sense as possible of how you would feel if you were in John's place. Then, at the end of the minute, write down what you imagined. We have found that carefully following this procedure can ensure understanding."

Remember. "We would like for you to remember John's story. That is, remember the facts as they were reported to you about John walking down a crowded street on a warm night when someone pushed him. Remember also the facts surrounding the person shoving John. Take about one minute for this exercise, getting as clear a memory as possible of the story. Then, at the end of the minute, write down what you remembered. We have found that carefully following this procedure can ensure understanding."

Participants typed what they had either imagined or remembered into a text box. These responses were intended to reinforce the prompt manipulation, and were not themselves analyzed for this study. Participants were not allowed to advance before 2 minutes had passed, at which point the screen advanced automatically.

Next, participants answered a verdict question that was identical to that from Experiment 1, except that participants in the Mental Illness condition were offered the additional choice of "Not Guilty by Reason of Insanity (NGRI)." We provided no explanation for this verdict choice, which allowed us to evaluate whether participants tend to view this verdict as more punitive or protective. Next participants answered sentencing and blame questions identical to those seen in Experiment 1.

On the screen with the blame question, participants were asked a new projection question. Because the projection question used in Experiment 1 dealt with a future event, it remained possible that participants in the *Neurotypical* condition had not been using higher levels of projection when reasoning about the event in our vignette. Further, jurors in actual cases are asked to evaluate the alleged *past* actions of the defendant, not future actions as were the focus of our original projection assessment. Thus, in Experiment 2 we included a "backward" projection question as well. The question read as follows:

Backward Projection: "When I was deciding whether John was guilty and whether/how he should be punished, I thought about what I would have done in his situation." This was rated from 1 (Strongly disagree) to 7 (Strongly agree).

On the next screen, participants answered the same projection, stereotyping, individual information, and reasonable person questions from Experiment 1. On a separate screen participants were asked the two similarity questions from Experiment 1. Participants then answered content and a reading comprehension attention checks similar to those in Experiment 1. Finally, participants provided their age and gender.

Results

Verdict. Verdict choices are reported in Figure 2-4. Because participants in the *Neurotypical* condition were not offered the NGRI option, we analyzed verdict choices by comparing convictions with non-convictions – not guilty or NGRI. While NGRI verdicts are substantively different from either other category, they are not convictions. Further, our analysis of the sentencing variables, reviewed below, suggests that our participants treated NGRI verdicts more like not guilty verdicts in this study. In the *Neurotypical* condition, 69% of verdicts were not guilty and 31% were guilty. In the *Mental Illness condition*, 38% of verdicts were not guilty, 26% were guilty, and 36% were NGRI.

We ran a binary logistic regression with backward elimination. Defendant condition, prompt type, and an interaction term were entered as predictors of guilty versus other (not guilty + NGRI). This analysis revealed that the interaction term was a significant, positive predictor of verdicts, $\beta = .540$, $p < .014$.[27] We followed this up with two tests. For the neurotypical defendant, we found that prompt type was not a significant predictor of verdicts, $\beta = .236$, $p < .332$. For the defendant with a mental illness, however, prompt type was a significant predictor, with the prompt to project associated with fewer verdicts of guilty, $\beta = .559$, $p < .031$. Within the *Mental Illness* condition, 32% of the non-convictions were NGRI, and 34% were not guilty. These numbers did not differ from each other ($p = .559$).

---

[27] An alternative analysis comparing not guilty verdicts to a collapsed NGRI/guilty category revealed that only defendant condition was a significant, negative predictor of verdicts, $\beta = -1.30$, $p < .001$, although prompt type was marginal, $\beta = .321$, $p < .055$. This shows that participants in our *Mental Illness* condition were more likely to choose the guilty/NGRI collapsed category than participants in the *Neurotypical* condition, and that participants receiving the project prompt were marginally more likely to choose not guilty.

Figure 2-4: Verdict choice as a function of defendant condition and prompt.
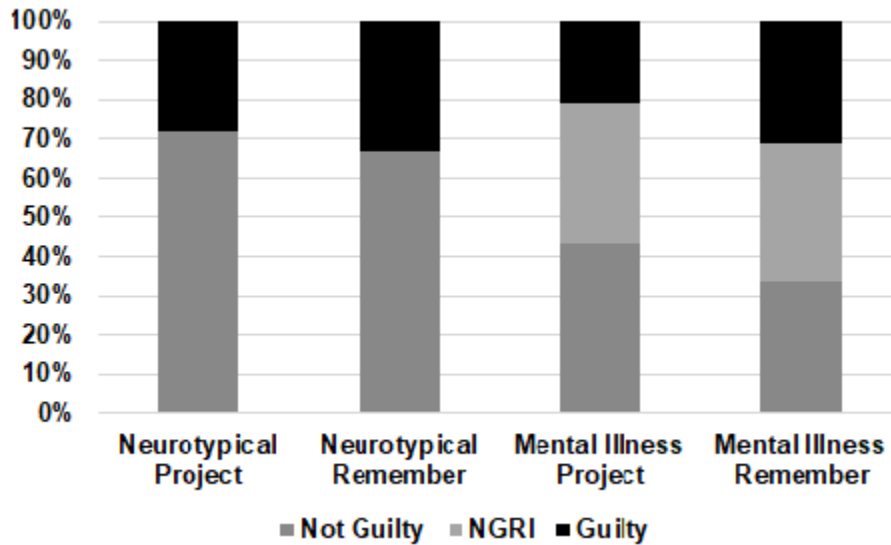
Sentences. We again coded those participants who did not choose a sentence as indicating a zero (see Table 1-4). Mann Whitney U tests on each sentence measure revealed no effects of prompt; we therefore performed Mann Whitney U tests for each measure as a function of defendant condition. Consistent with the results from Experiment 1, we found that defendants with a mental illness were assigned significantly more jail time, $U = 49292.50$, $p = .046$, and significantly more counseling, $U = 17593$, $p < .001$. Reinforcing the trend we saw in Experiment 1, fines were significantly higher for the neurotypical defendant, $U = 43845.00$, $p < .001$. There were no significant effects for service hours.

One additional question that we sought to address in Experiment 2 is whether participants employ an NGRI verdict somewhat protectively (to ensure that a defendant receives treatment) versus punitively (as a way to punish the defendant, and hence more akin to a "guilty" verdict). To do so, we compared sentences in the *Mental Illness* condition to see whether participants who chose the NGRI option patterned more closely with those who chose guilty or not guilty. Four Kruskal Wallis H tests, one for each sentence measure, revealed a significant difference between the verdict groups for fines, X^2 $(2, N = 320) = 23.34$, $p < .001$, jail time, X^2 $(2, N = 320) = 10.19$, $p < .006$, and counseling time, X^2 $(2, N = 320) = 35.34$, $p < .001$. We followed up these significant results with Mann Whitney U tests, which revealed that participants who chose guilty chose significantly higher fines ($M = 493.19$, $SD = 691.71$) than participants who selected either not guilty ($M = 165.02$, $SD = 399.13$), $U = 3559.0$, $p < .001$, or NGRI ($M = 224.25$, $SD = 586.60$), $U = 3417.5$, $p < .001$, who did not differ from each other. Jail sentences were also higher after a guilty verdict ($M = 16.08$, $SD = 53.80$) than either a not guilty ($M = 1.95$, $SD = 9.97$), $U = 4503.5$, $p < .003$, or an NGRI verdict ($M = 4.57$, $SD = 22.82$), $U = 4275.0$, $p < .028$, which did not differ from each other. Fittingly, the pattern was different for counseling. Participants who chose NGRI assigned significantly longer counseling terms ($M = 12.74$, $SD = 7.21$) than participants who assigned either guilty ($M = 8.37$, $SD = 8.13$), $U = 3139.0$, $p < .001$, or not guilty ($M = 7.57$, $SD = 7.11$), $U = 4013.0$, $p < .001$, and these latter conditions did not differ from each other. These results suggest that while the NGRI verdict may not be selected for exclusively therapeutic purposes, it is associated with levels of punishment akin to those given by participants who chose not guilty.

74

Mind Reading Strategy. We conducted a 2 (defendant condition: *Neurotypical*, *Mental Illness*) by 2 (prompt: *Remember*, *Project*) ANOVA for each of our six questions about mind reading strategy: backwards projection, forward projection, stereotyping, reasonable person, individual information, and similarity (see Figure 3-4).

Figure 3-4: Mind reading strategy variables by defendant condition. Error bars are one SEM in each direction. All differences are significant.
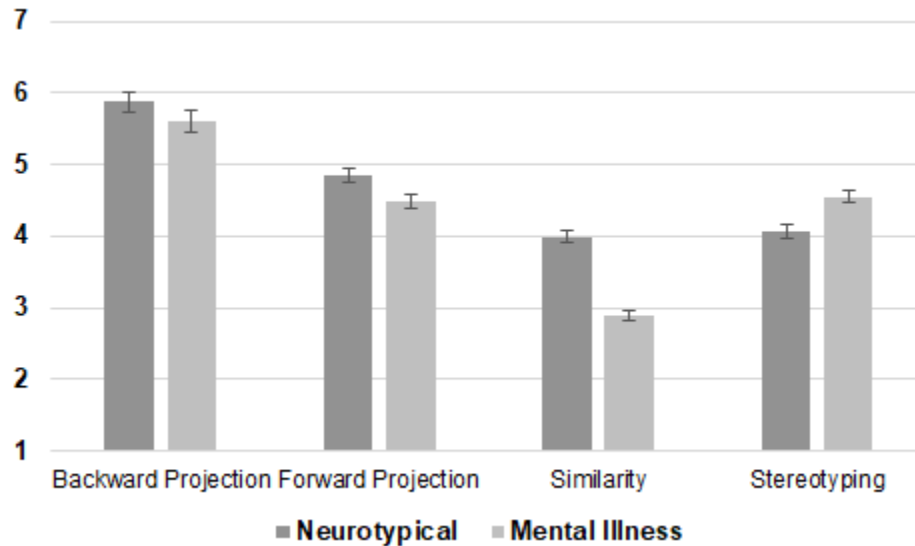
All six measures revealed main effects of defendant condition. Backward projection, $F(1, 636) = 7.16$, $p < .008$, $\eta^2 = .011$, forward projection, $F(1, 635) = 6.24$, $p < .013$, $\eta^2 = .010$, reliance on a reasonable person, $F(1, 636) = 22.55$, $p < .001$, $\eta^2 = .034$, individual information, $F(1, 636) = 24.50$, $p < .001$, $\eta^2 = .037$, and similarity, $F(1, 636) = 98.70$, $p < .001$, $\eta^2 = .134$, were significantly higher for *Neurotypical* than for *Mental Illness*, while stereotyping was significantly higher for *Mental Illness* than for *Neurotypical,* $F(1, 635) = 14.27$, $p < .001$, $\eta^2 = .022$. This mirrors the findings from Experiment 1, with the addition of a significant effect for individual information.

Three measures revealed significant effects of prompt condition. First, backward projection, $F(1, 636) = 106.36$, $p < .001$, $\eta^2 = .143$, and forward projection, $F(1, 635) = 31.67$, $p < .001$, $\eta^2 = .048$, were both were significantly higher in the *Project* condition ($M_{bp} = 6.28$, $SD_{bp} = .967$; $M_{fp} = 5.07$, $SD_{fp} = 1.73$) than in the *Remember* condition ($M_{bp} = 5.21$, $SD_{bp} = 1.604$; $M_{fp} = 4.27$, $SD_{fp} = 1.84$). This suggests that our manipulation of projection via instructional prompts was successful. Second, participants reported that they relied significantly more on a "reasonable person" in the *Project* condition ($M = 5.33$, $SD = 1.57$) than in the *Remember* condition ($M = 4.95$, $SD = 1.58$), $F(1, 636) = 9.70$, $p < .002$, $\eta^2 = .015$.

Finally, the analysis of backward projection revealed a significant interaction between defendant condition and prompt condition, $F(1, 636) = 4.06$, $p < .044$, $\eta^2 = .006$. With the remember prompt, backward projection was significantly higher for the *Neurotypical* defendant ($M = 5.45$, $SD = 1.42$) than for the defendant with a *Mental Illness* ($M = 4.96$, $SD = 1.74$), $F(1, 318) = 7.55$, $p < .006$, $\eta^2 = .023$. With the projection prompt, this difference was no longer significant ($M_n = 6.31$, $SD_n = 1.04$ vs. $M_{mi} = 6.24$, $SD_{mi} = .895$), $F(1, 318) = .404$, $p < .526$, $\eta^2 =$

.001. This interaction was limited to the backwards projection question, but it suggests that while default levels of projection may favor neurotypical defendants, an appropriate prompt can have a mitigating effect, eliminating this difference entirely.

As in Experiment 1, we also tested whether similarity mediated the effect of defendant condition on projection. A Sobel test revealed full mediation for both forward projection, $z = -6.62$, $p < .001$, and backward projection, $z = -4.64$, $p < .001$.[28]

Blame. A 2 (defendant condition: *Neurotypical*, *Mental Illness*) by 2 (prompt: *Remember*, *Project*) ANOVA with blame as the dependent variable revealed a significant effect of prompt: participants assigned higher levels of blame in the *Remember* condition ($M = 3.89$, $SD = 1.46$) than in the *Project* condition ($M = 3.58$, $SD = 1.53$), $F(1, 636) = 6.98$, $p < .008$, $\eta^2 = .011$. We did not replicate the effect of defendant condition found in Experiment 1, with ratings being just barely higher in the *Mental Illness* condition (M = 3.75, SD = 1.50) than in the *Neurotypical* condition (M = 3.72, SD = 1.50). The interaction was not significant.

Does mind reading strategy predict verdict? As in Experiment 1, we evaluated whether projection and stereotyping were associated with different verdicts. Given that we saw a significant effect of defendant condition on verdict choices when collapsing guilty with NGRI, we ran a binary logistic regression predicting verdict (1=not guilty, 2=guilty/NGRI) as a function of both kinds of projection and stereotyping. We found that forward projection was a significant, negative predictor of verdicts, $\beta = -.303$, $p < .001$, indicating that greater forward projection was associated with verdicts of not guilty. Backward projection was also a negative predictor, but did not reach significance. In contrast, stereotyping was a significant, positive predictor, $\beta = .218$, $p < .001$, indicating that greater stereotyping was associated with fewer verdicts of not guilty. We next conducted two Sobel tests to determine whether forward projection and/or stereotyping significantly mediated the effect of defendant condition on verdict choices. This analysis revealed that forward projection partially mediated the effect, $t = 2.24$, $SE = .04$, $p < .025$,[29] as did stereotyping, $t = 2.37$, $SE = .03$, $p < .018$.[30]

Discussion

The first aim of Experiment 2 was to replicate key findings from Experiment 1 supporting Hypotheses 1a, 1b, 2a, and 3. Consistent with Hypotheses 1a and 1b and replicating the results of Experiment 1, we found that projection was higher for the *Neurotypical* condition, while stereotyping was higher for the *Mental Illness* condition. Going beyond Experiment 1, we

---

[28] In the first step of the forward projection regression, defendant condition was a significant predictor of forward projection, accounting for 1% of the variance, $F(1, 637) = 5.99$, $p < .015$, $R^2 = .009$. In the second step, the addition of similarity was a significant improvement, accounting for 11% of the variance, $F(1, 636) = 39.60$, $p < .001$, $R^2 = .111$; R^2 change = .101, $p < .001$. In the first step of the backward projection regression, defendant condition was a significant predictor of backward projection, accounting for 1% of the variance, $F(1, 638) = 6.12$, $p < .014$, $R^2 = .010$. In the second step, the addition of similarity was a significant improvement, accounting for 4% of the variance, $F(1, 637) = 13.67$, $p < .001$, $R^2 = .041$; R^2 change = .032, $p < .001$.

[29] When only defendant condition was included in the regression the Cox and Snell R Squared is $r = .093$; but when forward projection is added in the second step the Cox and Snell increases to $r = .131$

[30] For stereotyping, we also find significant, partial mediation. Including only defendant condition in the regression results in a Cox and Snell R Squared of $r = .093$, but adding stereotyping in the second step increases the Cox and Snell to $r = .096$.

found that similarity *fully* mediated this difference in projection. Consistent with Hypothesis 2a and again replicating the results of Experiment 1, we found that higher projection was associated with more verdicts of "not guilty." Unlike Experiment 1, we also found support for Hypothesis 2b: greater stereotyping was associated with fewer verdicts of "not guilty." Finally, supporting Hypothesis 3 and reinforcing the findings from Experiment 1, we found that participants judged the defendant with a mental illness more harshly: assigned significantly more jail time. However, they assigned significantly higher fines to the neurotypical defendant. Also as in Experiment 1, we found that projection partially mediated the effect of defendant condition on verdicts.

The second aim of Experiment 2 was to determine whether mindreading strategies can be altered by a prompt to engage in projection, as predicted in Hypothesis 4. Our manipulation check suggests that the prompt was successful: participants reported engaging in significantly more projection when prompted to project versus remember, and this effect was especially pronounced when the defendant had a mental illness. However, the effect of greater projection on subsequent judgments was mixed. A prompt to project was associated with fewer verdicts of guilty for defendants with mental illness, and marginally associated with more verdicts of not guilty overall. The prompt to project also resulted in lower attributions of blame, but did not affect sentences. The effect on verdicts is especially noteworthy, however, as this is the outcome measure that real jurors are asked to provide.

Experiment 3

Experiment 3 tests our final prediction (Hypothesis 5): that including instructions about a legal insanity standard will decrease projection, increase stereotyping, and lead to greater punishment. This prediction stems from the observation that the function of providing an insanity standard is to draw attention to the defendant's mental illness. While legally this focus is meant to have a mitigating impact, the results from our previous experiments suggest that causing decision-makers to focus on a defendant's mental illness could have the opposite effect.

Methods

Participants. Four-hundred-and-eighty adults (280 female, 199 male, 1 other/prefer not to specify, mean age = 37, *SD* = 12) participated in the study through Amazon Mechanical Turk as in Experiments 1-2. An additional 35 participants were tested, but were excluded for failing attention checks (explained below).

Materials & Procedures. The experimental stimuli consisted of one vignette crossed with three standard conditions, resulting in three conditions to which participants were randomly assigned. The vignette was identical to the *Labeled* condition from Experiment 1 and the *Mental Illness* condition from Experiment 2. The standard conditions were either *Model Penal Code* (*MPC*), *M'Naghten*, or *None*. While insanity standards vary across states, the vast majority of states model their standards on either the MPC or M'Naghten standards.[31] Participants read the vignette, and were then randomly assigned to read either no further instructions, or one of the

---

[31] 20 states and the District of Columbia model their standards on the MPC while 25 states model their standards on the M'Naghten rule.

following sets of instructions, modeled after the jury instructions used in California and Michigan, respectively.[32]

> *M'Naghten*. "You must decide whether the defendant was legally insane when he committed the crime. The defendant must prove that it is more likely than not that he was legally insane when he committed the crime.
> The defendant was legally insane if:
> 1. When he committed the crime, he had a mental disease or defect;
> AND
> 2. Because of that disease or defect, he did not know or understand the nature and quality of his act; or did not know or understand that his act was morally or legally wrong.
> ***
> You may consider any evidence that the defendant had a mental disease or defect before the commission of the crime. If you are satisfied that he had a mental disease or defect before he committed the crime, you may conclude that he suffered from that same condition when he committed the crime. You must still decide whether that mental disease or defect constitutes legal insanity."
>
> *MPC*. "In order to establish that the defendant was legally insane, the defendant must prove two elements by a preponderance of the evidence. A preponderance of the evidence means that he must prove that it is more likely than not that each of the elements is true.
> (1)   First, the defendant must prove that he was mentally ill.
> (a)   "Mental illness" means a substantial disorder of thought or mood that significantly impairs judgment, behavior, capacity to recognize reality, or the ability to cope with the ordinary demands of life.
> (2)   Second, the defendant must prove that, as a result of his mental illness he either lacked substantial capacity to appreciate the nature and wrongfulness of his act, or lacked substantial capacity to conform his conduct to the requirements of the law.
> (3)   You should consider these elements separately. If you find that the defendant has proved both of these elements by a preponderance of the evidence, then you must find him not guilty by reason of insanity. If the defendant has failed to prove either or both elements, he was not legally insane."

Next, all participants answered a verdict question identical to that used in Experiment 2, including the option of NGRI. Then participants answered sentencing questions identical to those used in Experiments 1-2. On successive screens, participants answered the projection, stereotyping, and similarity questions used in Experiment 1. Next, participants answered true/false content and a reading comprehension attention checks similar to those in Experiments 1-2. Finally, participants provided their age and gender.

Results

**Verdicts**. In the *None* condition, participants assigned 23% not guilty, 23% guilty, and 54% NGRI verdicts. In the *M'Naghten* condition, participants assigned 26% not guilty, 35% guilty, and 38% NGRI verdicts. In the *MPC* condition, participants assigned 26% not guilty,

---

[32] It is worth noting that insanity standard instructions are given at different stages of trial in different states. For uniformity of comparison, we have given them at the same point in this experiment.

31% guilty, and 43% NGRI verdicts. We found no significant differences between the two standards, so we collapsed them into one *Standard* condition for further analysis (see Figure 4-4). A 2 (standards: *Standard*, *None*) by 3 (verdict: not guilty, guilty, NGRI) chi-square test revealed a significant effect, X^2 (1, N = 480) = 9.08, p < .011. As predicted, the standards had the effect of significantly increasing guilty and decreasing NGRI verdicts.

Figure 4-4: Verdicts as a function of standard condition in Experiment 3.
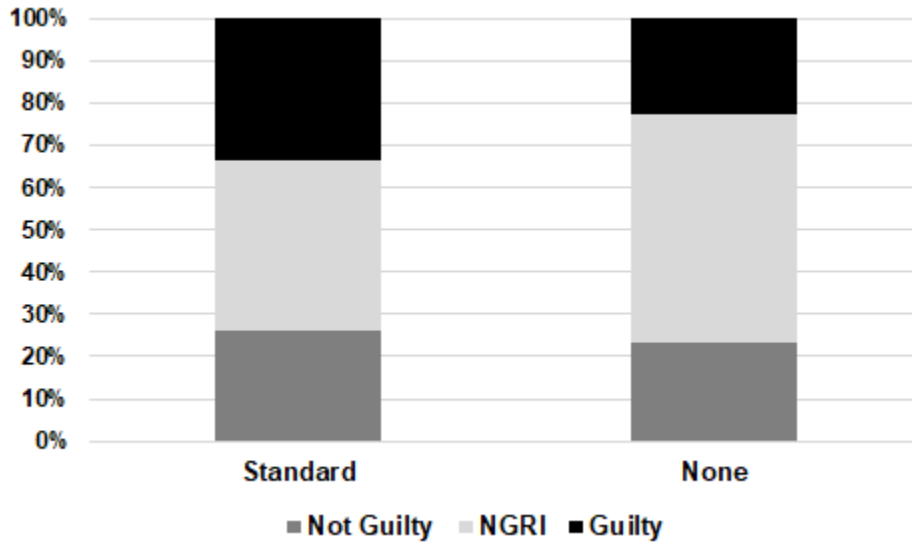
Sentences. As in Experiments 1 and 2, we coded those participants who did not choose a sentence as indicating a zero. We ran a Kruskal-Wallis test on each sentencing variable, but only counseling showed a significant difference, X^2 (2, N = 480) = 6.62, p < .037. A Mann Whitney U test found that participants assigned significantly more counseling in the *M'Naghten* condition than either *None*, U = 11206.00, p < .049, or *MPC*, U = 10862.00, p < .016.

Mind Reading Strategy. We conducted *t*-tests on projection, stereotyping, and similarity and found no significant differences between the conditions on any of these measures.

Does mind reading strategy predict verdict? We conducted a multinomial regression using guilty as the comparison category, and with projection and stereotyping as predictors. Projection was a significant, positive predictor for not guilty verdicts, β = .213, p < .001, indicating that the more projection participants reported, the more likely they were to render a not guilty verdict. Stereotyping was a significant, positive predictor for NGRI verdicts, β = .413, p < .001, with more stereotyping leading to a significantly greater likelihood of choosing the NGRI verdict.

Discussion

Hypothesis 5 predicted that participants given a standard would increase punishments by decreasing projection and potentially increasing stereotyping. Experiment 3 partially supports this prediction. When provided with a standard, participants generated more guilty verdicts, and fewer NGRI verdicts, than when no standard was provided. Participants in the *M'Naghten* condition also assigned significantly more counseling than either of the other conditions.

However, the effect of standards was not otherwise reflected in levels of punishment or in mindreading strategies.

The effect of instructions on verdicts is notable in light of the fact that previous research on insanity instructions has found no significant difference in outcomes based on which, or whether, instruction is provided. However, important differences in methodologies may help explain this apparent inconsistency. First, most of the previous studies dealt with the killing of at least one other person. Finkel et al.'s studies had participants consider defendants who had killed one person (Finkel & Handel, 1988; Finkel et al., 1985), and Ogloff's study concerned a defendant who had killed four college-aged people, including his own daughter (1991). Fearing that an extreme crime might obscure effects of more subtle moderators, we chose the crime of battery because it is not rated as a particularly severe crime (Bureau of Justice Statistics, 1985). The only study that used a less serious crime, breaking and entering, compared the M'Naghten and Durham standards to one another and asked only for a verdict (James, 1959). We did not test the Durham standard because it is now only used in one state, and we did not find a difference between standards in verdicts, only in sentencing. Our results are therefore consistent with James' findings, to the limited extent that the two studies are comparable.

Second, some of the prior work on standards compared standards only to each other, and not to a condition involving no instruction at all. Ogloff (1991) did include a no instruction condition, but used an extreme crime as discussed above. Finkel and colleagues compared results across two papers: Finkel and Handel (1988), which did provide standards, and Finkel et al. (1985), which did not. They found no significant difference in patterns of verdicts across the two studies. However, both studies still concerned killings. Finally, only the Finkel et al. (1985) study considered what they called consequences – such as our counseling measure – and they do not report whether these consequences varied by instruction given, so we cannot evaluate whether our sentencing result is inconsistent with their findings. Thus, while our results may initially appear inconsistent with previous research on insanity standards, this may reflect the limited range of crimes tested rather than a genuine inconsistency.

General Discussion

Across three studies, we explored three broad questions: 1) would mind reading strategies behave in the legal context as they have in previous, non-legal research, 2) would the presence of a mental illness alter mind reading strategies, and 3) would mind reading strategies have an impact on verdicts and punishment? Across three experiments, we find evidence that the answer to each of these broad questions is "yes." The evidence supporting, and partially refuting, each of our specific hypotheses provides more nuanced answers.

Our first set of hypotheses was that participants would engage in less projection (Hypothesis 1a) and more stereotyping (Hypothesis 1b) when a defendant was described as having a mental illness. Supporting these hypotheses, Experiments 1 and 2 found that participants in our *Mental Illness* conditions reported significantly lower projection and significantly greater stereotyping than participants in our *Neurotypical* condition. Further, linking these findings to prior research outside of the legal domain, we found that similarity mediated this difference in projection levels – partially in Experiment 1 and fully in Experiment 2. This is consistent with recent findings that people use less mentalizing language when interacting with or describing members of outgroups (McClung & Reicher, 2017). Outgroup membership may be one way people assess perceived similarity; this is an interesting direction for future research.

Our second set of hypotheses related mind reading strategies to punitive outcomes. We predicted that higher projection would be associated with outcomes that were less punitive (Hypothesis 2a) while higher stereotyping would be associated with outcomes that were more punitive (Hypothesis 2b). In Experiments 1 and 2, we found clear support for Hypothesis 2a: higher projection was a significant predictor of not guilty verdicts and partially mediated the effect of defendant condition on verdicts. Support for Hypothesis 2b was mixed. In Experiment 2, stereotyping was a significant predictor of guilty verdicts, but in Experiment 1 it was not.

Hypothesis 3 predicted that defendants with a mental illness would be treated more harshly than defendants without a mental illness, even for identical actions. Supporting this prediction, we find in Experiments 1 and 2 that participants in the *Mental Illness* condition were assigned significantly more jail time and were significantly less likely to receive not guilty verdicts than participants in the *Neurotypical* condition. However, participants in the *Neurotypical* condition assigned higher fines, marginally in Experiment 1 and significantly in Experiment 2. While fines are punishment, they are certainly a lesser punishment than jail time. Our findings are thus consistent with Hypothesis 3.

Our fourth hypothesis predicted that we could increase participants' level of projection, and thus decrease punishment, by prompting participants to put themselves in the defendants' place. While the prompt had no impact on sentences, it did impact verdicts. Participants in the *Project* condition giving significantly higher ratings for both of our measures of projection. Regressions also show that, overall, participants in the *Project* condition were marginally more likely to assign a verdict of not guilty, with a significant decrease in guilty verdicts for defendants with a mental illness. These findings suggest that a simple verbal prompt can increase levels of projection, leading to less punitive verdicts, particularly when projection may naturally be low, as when the defendant has a mental illness.

Our final hypothesis, that insanity standards would increase stereotyping and lead to more punitive outcomes, was partially supported. Participants did assign more guilty verdicts when given an insanity standard than when they were not provided with a standard, and participants in the *M'Naghten* condition gave longer counseling sentencings than participants in either the *MPC* or *None* condition. However, we found no difference in levels of projection, stereotyping, or similarity among the groups, leaving open what it was that drove the difference in verdicts and counseling.

Our findings largely support our hypotheses, and are also consistent with prior research insofar as the effects of mental illness were driven by differences in perceived similarity. However, this prompts an important further question: what led our participants to view defendants with a mental illness as less similar? Our supplemental experiment provides some hints. In Experiment S1, we found that defendants with a mental illness – any mental illness – were rated as less warm and significantly more unpredictable than a neurotypical defendant. Warmth was a significant predictor of not guilty verdicts, while unpredictability was a significant predictor of guilty verdicts. Unpredictability was also a significant, negative predictor of similarity. Thus, it could be that participants view mental illness as a proxy for these facets of warmth and unpredictability, and it is for this reason that defendants with a mental illness are rated as less similar and treated more harshly.

Implications for Law

The implications of our results for the legal system are simultaneously troubling and hopeful. The finding that providing a legal standard of insanity decreased NGRI verdicts is a

discouraging finding for defense attorneys, as an NGRI plea must be accompanied by instructions on the standard to be used. However, our results also suggest that a simple verbal prompt to imagine the event from the defendant's perspective could lead to fewer guilty verdicts, especially when the defendant has a mental illness. This kind of prompt could easily be incorporated into opening and closing statements, does not require any legislative changes, and could offset some of the negative impact of the legal standard.

Implications for Psychology

This research also enriches the psychological literature on mind reading. The current work shows that the basic mind reading processes described in other contexts are applicable in the legal context, even when asking a fundamentally different set of questions. Further, our study extends prior research by showing how mind reading interacts with mental illness, and that mind reading strategies do have downstream consequences on measures of punishment. Finally, our supplemental experiment begins to address what it is about mental illness – and about the perception of social groups more generally – that could affect perceived similarity and the subsequent choice to engage in projection. Specifically, we find additional evidence that warmth is an important dimensions of social evaluation, and more tentatively propose that it least in the legal context, predictability may be as well.

Limitations

The current study has important limitations. Juries make decisions after live testimony and in groups, not individually on the basis of written materials. Our materials were also brief, lacked the vividness of a real trial, and considered a limited range of crimes. In future research, we hope to address these limitations. However, it is worth noting that a translation to more realistic scenarios could well exaggerate, rather than attenuate, the effects that we find here. In particular, longer and more vivid materials could heighten the salience of a mental illness, increasing the gap in judgments for defendants with a mental illness and those without. Either way, we think the current work offers some first steps in bridging methods and theories from basic research on mind reading with real legal judgments; more work certainly needs to be done.

Future Research

The present research serves as the basis for much future research. It would be instructive to investigate a wider variety of crimes to see whether the differences we find with battery would be exaggerated, attenuated, or remain stable with more or less serious crimes. The impact of standards also deserves further research. We found, as predicted, that including a legal standard of insanity led to more punitive outcomes, but it did not have an effect on mind reading strategies. Thus it is not clear *how* the standards affected verdicts. It is also not clear what impact a prompt to project would have when legal standards are presented: would the prompt still be mitigating? Finally, we considered only two mindreading strategies, each assessed by a single measure. Mind reading is undoubtedly more complex, and additional strategies and measures should be considered. Further experiments are necessary to address these questions and more.

Conclusions

Our findings demonstrate, for the first time, the impact of mind reading strategy on legally-relevant outcome measures. We find that mind reading operates in the legal system in largely the same way it has in previous research, with the choice of mind reading strategy

depending importantly on perceptions of similarity. We extend these results not only by finding them in a legal context, but also by finding that the presence of a mental illness reduces perceived similarity and projection. Our results also show that higher levels of projection lead to less punitive outcomes for defendants, and that the use of projection can be increased with simple prompts. These findings suggest that mind reading strategy is an important consideration in the legal system, and one which lawyers have the power to address.

Chapter 5

Conclusion

As Chapter One described, all three studies simply take different avenues to explore the same key questions. At the broadest level, these three studies consider: what makes a good legal explanation? However, each study uses a slightly different approach and set of materials to address a more specific question: when are these explanations more or less mitigating? In Chapter 1, I raised several question about the role of mental state explanations in legal decision making, and Chapters 2 through 4 shed light on these questions, each in slightly different ways. While Chapter Three is an extension of Chapter Two's work on when ignorance is a mitigating explanation, it concerns the entirely different context of a school. Chapter Four is further removed, considering for the first time the strategies people use to evaluate explanations as well as defendants' with a mental illness. Thus, while the three preceding chapters might seem somewhat disconnected from each other, their results converge on unifying themes that begin to answer these questions.

The first undeniable theme from my research is that the mental state explanation offered by a defendant impacts legal decisions. More broadly, the mental state explanation offered by an actor impacts how that actor is judged. The overall trend discussed in chapter one shows that, typically, as an actor's mental state moves from ignorance to knowledge, punishment and blame also increase (i.e. Giffin & Lombrozo, 2016; Mueller, Solan, & Darley, 2012; Shen, Hoffman, Jones, Greene, & Marois, 2011; Mikhail, 2009). My research does not entirely contradict this trend. In Chapter Two, we found that for *mens rea* crimes punishment and censure increased significantly as an actor's mental state progressed from ignorance to knowledge. Chapter Three replicated this finding in a school setting, even when the wrong doers are children and the punishment is school service hours, the actor's ignorance was mitigating for most violations. That is, both Chapters Two and Three find that an actor who explains his or her actions by saying, "yes I committed that act, but I did not know I was breaking the law/rule," are generally afforded a measure of leniency compared to actors who knowingly violate the law or rule.

Chapter Four demonstrated the importance of mental state explanations in a different way. We found that how a defendant explained their mental state, whether or not they said, "yes I committed that act, but I have a mental illness," changed how their mental state was evaluated. This finding builds on prior research outside the legal context. However, in a novel finding, we also found that these evaluation strategies have different punitive consequences. The more projection a participant reported using, the more likely they were to find the defendant not guilty, and projection was significantly lower when the defendant explained her behavior by reference to a mental illness. While readers may not be surprised that different mental state explanations have different results, in this study what is interesting is *how* the explanation impacts the results. Different explanations lead to varying levels of perceived similarity, which elicited distinct evaluation strategies, and it is these strategies that are at least partly responsible for the different treatment defendants received. This leads to the prediction that other explanations that highlight difference might also lead to lower similarity, lower projection, and more guilty verdicts.

The second theme of these studies is related to the first: while mental state explanations impact legal decisions, they do not always do so in the manner laypeople, scholars, or the law would expect. While an explanation that appeals to a defendant's ignorance is typically

mitigating, the studies in Chapters Two and Three find specific, and surprising, instances where it is not. Participants found ignorance of vital facts a less mitigating explanation for violations of strict liability crimes and conventional rules. This finding was contrary to the weight of the legal scholarship as well as lay intuition that ignorance should always be a mitigating explanation. However, the judgments did coincide with already identified legal and psychological categories. Statistical tests as well as experimental manipulations across the two studies suggest that this difference in treatment lies in the way these kinds of laws and rules are set. Strict liability laws and conventional rules seem arbitrary in that the line of permissibility could have been drawn differently – the speed limit could have been 35 MPH instead of 25 MPH or the acceptable shirt color could have been red rather than blue. The act of driving 35 MPH or wearing a red shirt is not inherently bad, it is only bad because it has been prohibited. That the action itself is not seen as inherently wrong is because performing the act is not believed to have any inevitable consequences. Driving ten miles over the speed limit will not inevitably cause a traffic accident, it is not even clear how much more likely such an accident becomes. Thus, violating these laws or rules knowingly is not seen as *intending* any additional consequences, a knowing violation merely entails knowingly breaking the law or rule. This alone does not seem worthy of significantly greater punishment or censure. Chapters Two and Three use converging statistical and experimental evidence to support a theory of why, counterintuitively, ignorance is not always a mitigating explanation.

Chapter Four also exemplifies this second theme, that mental state explanations do not always have the expected impacts. In this study, we find, counter to the law's explicit expectation, that a defendant who may be unable to form the necessary *mens rea* due to a mental illness was not given more lenient treatment. In the first two experiments in this study, a defendant who is described as having a mental illness is treated worse than a defendant who is not, even though the actions are identical. In these experiments, though, participants were not offered any legal standard by which to judge the defendant's mental illness, leaving open a question of how the legal standards would impact judgments. The third experiment in this study provided some participants with insanity standards by which to judge the defendant. While the differences across conditions were subtle, those participants who read a legal standard were more likely than participants who read no standard at all to find the defendant guilty. It could be that the actions described in our vignette simply did not satisfy the standards given, that is, participants who were given standards realized that contrary to their lay intuitions, the actions described do not meet the legal standard. In some respects, that could be characterized as participants carefully applying the law they have been given. It is, however, potentially troubling that the law that governs insanity pleas make them *less* likely to succeed than participants' unguided intuition would allow. This finding also seems to contradict the essence of the insanity plea as a defense: a defendant who uses her mental illness as an explanation for her behavior may find that she fairs worse, not better, than had she not brought this evaluation to bear.

While mental state explanations may not always do what is predicted, what we have learned from the studies presented just in this dissertation may help us understand some of the most well-known findings in legal psychological research. Pennington and Hastie famously found that jurors represent trial evidence in the form of narrative stories (1986). One surprising, though often overlooked, finding was that a full forty-five percent of the narratives participants constructed had not been presented as evidence. That is, forty-five percent of these narratives were composed of inferences the participants made because the information fit their narrative

and filled an information gap. Pennington and Hastie did not address how participants went about making these inferences, but the results in this dissertation suggest possible answers. Characteristics of the defendant may have made him seem more or less similar to participants, activating more or less stereotype consistent information which could be used to fill in gaps. Further, the crime the defendant had been accused of, homicide, is certainly not arbitrary leading participants to focus on his mental state, a key distinction in this trial. This is just one example of how present results might relate to past literature. This example illustrates that while mental state explanations are not always *a priori* predictable, what we learn by studying them sheds light on important decision making processes.

The two themes of this dissertation, as well as the preceding application example, emphasize the need for more research in this area. Mental state explanations impact legal decisions, but laypeople, scholars, and legislators are not always able to predict what these impacts will be. In some instances, this research finds unexpected correspondence between the law and cognitive processes. Participants in our studies found the actor's knowledge less important when judging strict liability crimes than when judging *mens rea* crimes. This sorting mirrored a legal distinction that has long been criticized by scholars and attorneys. In our studies, this sorting appears to be partially based on the arbitrariness of the line of illegality. In this instance, it appears that the law reflects a real cognitive distinction, and this cognitive distinction may help to explain why crimes that on the surface may look quite different were placed in the same category. In other instances, mental state explanations have impacts that are in contrast with the law. We find that explaining behavior by reference to mental illness elicits a different mind reading strategy that leads to harsher outcomes for defendants. These issues are complicated and the studies in this dissertation have only begun to explore them, but it is safe to say that the law does not intend for a person with a mental illness to be penalized for that fact, unrelated to the action she has committed, which our findings suggest she may well be. More research is necessary to gain a better understanding of how mental state explanations impact a variety of legal decisions, including those I have begun to explore in this dissertation.

While the picture is more complicated than may be assumed by laypeople, legal scholars, psychologists, and legislators, it is not entirely abstract, and these studies and others continue to make it clearer. My own continuing research seeks to answer one broad question: what makes a good legal explanation? To answer this big question, my ongoing work addresses questions in three main categories: 1) how do mental state explanations impact legal decisions, what aspects of belief, knowledge, and intent are more or less persuasive; 2) what features of an explanation – source, structure, proper names – can change legal decisions, and 3) to what extent does the structure of our legal system reflect cognitive realities of categorization, priorities, and moral values.

Descriptive findings are not normative, and I do not mean to suggest that these results always suggest clear paths for change or advocacy. However, critique and reform attempts will be ill-formed, unhelpful, and unrealistic if they are based on inaccurate beliefs about how mental state explanations impact outcomes for defendants. Continuing research on this topic will provide insights that can be used to guide those who wish to change the system or merely wish to operate more efficiently within its current bounds. This research will also continue to enrich our knowledge of general human psychology. As the studies in this dissertation show, the answers to specifically legal questions can be applied to non-legal questions, and vice versus. This should not be surprising. The legal system is a concrete manifestation of categories, priorities, and

values. Studying this complex system cannot help but teach us about the even more complex minds that made it.

References

13, C.P.C., §§451, 452, 484(a).

Alicke, M. (2008). Blaming Badly. *Journal of Cognition & Culture, 8*, 179-186. doi:10.1163/156770908X289279

Ames, D. (2005). "Everyday Solutions to the Problem of Other Minds: Which Tools Are Used When?" In B. F. Malle and S. D. Hodges (Eds.), *Other Minds* (pp. 158-173). New York: Guilford.

Ames, D. (2004a). Inside the Mind Reader's Tool Kit: Projection and Stereotyping in Mental State Inference. *Journal of Personality and Social Psychology, 87* (3), 340-353.

Ames, D. (2004b). Strategies for Social Inference: A Similarity Contingency Model of Projection and Stereotyping in Attribute Prevalence Estimates. *Journal of Personality and Social Psychology, 87* (5), 573-585.

Ames, D. & Fiske, S. (2015). Perceived intent motivates people to magnify observed harms. *Proceedings of the National Academy of Sciences, 112*(12), 3599-3605. DOI: 10.1073/pnas.1501592112

Ames, D., Weber, E., & Zou, X. (2012). Mind-reading in strategic interaction: The impact of perceived similarity on projection and stereotyping. *Organizational Behavior and Human Decision Process, 117* (1), 96-110.

American Law Institute. (1985). Model penal code : official draft and explanatory notes : complete text of model penal code as adopted at the 1962 annual meeting of the American Law Institute at Washington, D.C., May 24, 1962. Philadelphia, Pa.

Angermeyer, M., & Matschinger, H. (2004). The Stereotype of Schizophrenia and Its Impact on Discrimination Against People With Schizophrenia: Results From a Representative Survey in Germany. *Schizophrenia Bulletin, 30*(4), 1049-1061.

Appiah, K.A. (2006). *Cosmopolitanism: Ethics in a World of Strangers* (New York: W. W. Norton).

Baldus, D.C., Woodworth, G., Zuckerman, D., Weiner, N.A., & Broffitt, B. (1998). Racial discrimination and the death penalty in the post-Furman era: An empirical and legal overview, with recent findings from Philadelphia. *Cornell Law Review, 83*, 1638–1770.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., ... & Scelza, B. A. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, 201522070.

Batson, C., Lishner, D., Carpenter, A., Dulin, L., Harjusola-Webb, S., Stocks, E., Gale, S., Hassan, O., & Sampat, B. (2003). "... As you Would have Them Do Unto You": Does Imagining Yourself in the Other's Place Stimulate Moral Action?. *Personality and Social Psychology Bulletin, 29* (9), 1190-1201.

Blackstone, W. (1769). *Commentaries on the Laws of England: Public Wrongs*. Oxford, UK.: Clarendon Press.

Bureau of Justice Statistics. (1985). The National Survey of Crime Severity.

Carrington, S. & Bailey, A. (2008). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping, 30* (8), 2313-2335. doi: 10.1002/hbm.20671

Chakroff, A., Dungan, J., Koster-Hale, J., Brown, A., Saxe, R., & Young, L. (2015). When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Social Cognitive and Affective Neuroscience*, 1-9. doi:10.1093/scan/nsv131

Corrigan, P.W., Morris, S.B., Michaels, P.J., Rafacz, J.D., Rüsch, N. (2012). Challenging the public stigma of mental illness: A meta-analysis of outcome studies. *Psychiatric Services 63*, 963–973.

Craswell, R., & Calfee, J. (1986). Deterrence and Uncertain Legal Standards. *J. Law, Economics, & Organization, 2(2)*, 279-314.

Cushman, F. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, 6, 97-103. DOI: 10.1016/j.copsyc.2015.06.003

Cushman, F. (2008). Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*(2), 353-380. DOI:10.1016/j.cognition.2008.03.006

Cushman, F., Inbar, Y., & Pizarro, D. (2012). Benefiting from Misfortune: When Harmless Actions are Judged to be Morally Blameworthy. *Personality & Social Psychology Bulletin, 38*, 52-62. doi: 10.1177/0146167211430232

Dar-Nimrod, I., & Heine, S. (2011). Genetic Essentialism: On the Deceptive Determinism of DNA. *Psychological Bulletin, 137*(5), 800-818.

Davis, M. (2017). Social projection to liked and disliked targets: The role of perceived similarity. *Journal of Experimental Social Psychology, 70*, 286–293. http://dx.doi.org/10.1016/j.jesp.2016.11.012

Davidson, P., Turiel, E., & Black, A. (1983). The Effect of Stimulus Familiarity on the Use of Criteria and Justifications in Children's Social Reasoning. *British Journal of Developmental Psychology, 1*, 49-65. doi: 10.1111/j.2044-835X.1983.tb00543.x

Diamond, S. & Rose, M. (2005). Real Juries. *Annual Review of Law and Social Science, 1*, 255-284.

Downs, J. S., Holbrook, M. B., Sheng, S., & Cranor, L. F. (2010, April). Are your participants gaming the system?: screening mechanical turk workers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2399–2402). ACM

Duff, R.A. (2005). "Strict Liability, Legal Presumptions, and the Presumption of Innocence." *In* Simester, A. (ed.), *Appraising Strict Liability*. Oxford, UK: Oxford University Press, 125-149.

Eberhardt, J., Davies, P., Purdi-Vaughns, V., & Johnson, S. (2006). Looking Deathworthy; Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes. *Psychological Science, 17*(5), 383-386.

Epley, N. (2014). *Mindwise: How we understand what others think, believe, feel, and want*. New York: Knopf.

Eyal, T., & Epley, N. (2010). How to seem telepathic: Enabling mind reading by matching self-construal. *Psychological Science, 21*, 700-705.

Federal Bureau of Investigation. (2012). Crime in the United States. Uniform Crime Reports.

Finkel, N. & Handel, S. (1989). How Jurors Construe "Insanity." *Law and Human Behavior, 13*(1), 41-59

Finkel, N. J., & Handel, S. F. (1988). Jurors and insanity: Do test instructions instruct? *Forensic Reports, 1*(1), 65-79.

Finkel, N., Shaw, R., Bercaw, S. & Koch, J. (1985). Insanity Defenses: From the Jurors' Perspective. *Law and Psychology Review, 9*, 77-92.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.). *The handbook of social psychology* (pp. 357–411). Boston: McGraw-Hill.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social perception: warmth and competence. *Trends in Cognitive Science, 11*,77-83.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–74). San Diego, CA: Academic Press.

Gardner, M. R. (1993). The Mens Rea Enigma: Observations on the Role of Motive in the Criminal Law, Past and Present. *Utah Law Review*, 635.

Garnett v. United States. 632 A.2d 797 (1993).

Giffin, C. & Lombrozo, T. (2016). Wrong or Merely Prohibited: Special Treatment of Strict Liability in Intuitive Moral Judgment. *Law and Human Behavior, 40*(6), 707-720. doi.org/10.1037/lhb0000212

Giffin, C. & Lombrozo, T. (2017). An actor's knowledge and intent are more important in evaluating moral transgressions than conventional transgressions. *Cognitive Science, pending*. doi.org/10.1111/cogs.12504

Giffin, C., Wilkenfeld, D., & Lombrozo, T. (2017). The explanatory effect of a label: Explanations with named categories are more satisfying. *Cognition, 168*, 357-369. doi.org/10.1016/j.cognition.2017.07.011

Gilbert, E., Tenney, E., Holland, C., & Spellman, B. (2014). Counterfactuals, Control, and Causation: Why Knowledgeable People Get Blamed More, *under review*.

Goodwin, G. P., & Darley, J. (2012). Why are some moral beliefs seen as more objective than others? *Journal of Experimental Social Psychology, 48,* 250-256. doi: 10.1016/j.jesp.2011.08.006

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *J Pers and Soc Psych*,*101*(2), 366. doi: 10.1037/a0021847

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology, 96*, 505–520. doi:10.1037/a0013748

Gray, K., &Wegner, D. M. (2012). Morality takes two: Dyadic morality and mind perception. In M. Mikulincer, & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil*. (pp. 109–127).Washington, DC: American Psychological Association. doi:10.1037/13091-006

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry, 23*, 206–215. doi:10.1080/1047840X.2012.686247

Green, S. (2005). "Six Sense of Strict Liability: A Plea for Formalism." *In* Simester, A. (ed.), *Appraising Strict Liability*. Oxford, UK: Oxford University Press, 1-20.

Hardecker, S., Schmidt, M., Roden, M., & Tomasello, M. (2016). Young children's behavioral and emotional responses to different social norm violations. *Journal of Experimental Child Psychology, 150*, 364-379. doi:10.1016/j.jecp.2016.06.012

Hart, H.L.A. & Honore, A.M., (1962). *Causation in the Law*. 1st ed. Oxford, UK.: Oxford University Press.

Hauser, M. D., Young, L., & Cushman, F. A. (2008). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral Psychology and Biology*. New York: Oxford University Press.

Hawley-Dolan, A., & Young, L. (2013). Whose Mind Matters More—The Agent or the Artist? An Investigation of Ethical and Aesthetic Evaluations. PLoS ONE 8(9): e70759. doi:10.1371/journal.pone.0070759.

Henry, J.D., Phillips, L.H., Ruffman, T., & Bailey, P.E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*(3), 826-839. http://dx.doi.org/10.1037/a0030677

Holmes, O. W. (1881). *The Common Law*. Boston, Massachusetts: Little, Brown, and Company.

House Report No. 113–46. *Mens Rea:* The Need for a Meaningful Intent Requirement in Federal Criminal Law (2013).

Imuta, K., Henry, J.D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology, 52*(8), 1192-1205. http://dx.doi.org/10.1037/dev0000140

Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin, 38*, 52–62. DOI: 10.1177/0146167211430232

James, R. (1959). Jurors' Assessment of Criminal Responsibility. *Social Problems, 7*(1), 58-69.

Josephs, M. & Rakoczy, H. (2016). Young children think you can opt out of social-conventional but not moral practices. *Cognitive Development, 39*, 197-204. doi:/10.1016/j.cogdev.2016.07.002

Josephs, M., Kushnir, T., Grafenhain, M. & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology, 141*, 247-255. doi: /10.1016/j.jecp.2015.08.002

Kant, I. (2013). Metaphysics of Morals. Cambridge, UK: Cambridge University Press.

Kappes, H. & Morewedge, C. (2016). Mental Simulation as Substitute for Experience. *Social & Personality Psychology Compass, 10*(7), 405-420. DOI: 10.1111/spc3.12257

Keating, G. (2006). Strict Liability and the Mitigation of Moral Luck. *Journal of Ethics & Social Philosophy, 2*, 1-33.

Knobe, J. (2005). Theory of Mind and Moral Cognition: Exploring the Connections. *Trends in Cognitive Science, 9*(8), 357-359. doi:10.1016/j.tics.2005.06.011

Laurent, S. & Myers, M. (2011). I know you're me, but who am I? Perspective taking and seeing the other in the self. *Journal of Experimental Social Psychology, 47*, 1316-1319.

Lebowitz, M., Rosenthal, J., & Ahn, W. (2013). Effects of Biological Versus Psychosocial Explanations on Stigmatization of Children with ADHD. *Journal of Attention Disorders, 20*(3), 240-250. DOI: 10.1177/1087054712469255

Lebowitz, M., & Ahn, W. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceedings of the National Academy of Sciences, 111* (50), 17786-90. doi: 10.1073/pnas.1414058111

Levenson, L. (1992). Good Faith Defenses: Reshaping Strict Liability Crimes. *Cornell Law Review, 78*, 401-469.

Lombrozo, T. (2012). Explanation and Abductive Inference. In Holyoak, K.J. & Morrison, R.G. (Eds.), *Oxford Handbook of Thinking and Reasoning* (pp. 260-276), Oxford, UK: Oxford University Press.

Malle, B.F., Guglielmo, S., & Monroe, A. E. (2014) A Theory of Blame. *Psychological Inquiry, 25*(2), 147-186. DOI: 10.1080/1047840X.2014.877340 31

Malle, B.F., & Knobe, J. (1997). The Folk Concept of Intentionality. *J Exp Soc Psych, 33*, 101 - 121. doi:10.1006/jesp.1996.1314

Malle, B.F., & Nelson, S. (2003). Judging Mens Rea: The Tension Between Folk Concepts and Legal Concepts of Intentionality. *Behavioral Sciences and the Law, 21*, 563-580. DOI: 10.1002/bsl.554

Manchester, C. (1977). The Origins of Strict Criminal Liability. *Anglo-American Law Review, 6*, 277-283.

Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin, 28*, 1601–1610.

McClung, S., & Reicher, S. (2017). Representing other minds: Mental state reference is moderated by group membership. *Journal of Experimental Social Psychology*. DOI: 10.1016/j.jesp.2017.11.013

Mikhail, J. (2009). Moral Grammar and Intuitive Jurisprudence: A Formal Model of Unconscious Moral and Legal Knowledge. In B.H. Ross (Ed.), *Psychology of Learning and Motivation* (pp. 27-100). Cambridge, MA: Academic Press. DOI: 10.1016/S0079-7421(08)00402-7

Monahan, J. (1992). Mental disorder and violent behavior: Perceptions and evidence. *American Psychologist, 47*(4), 511-521.

Mueller, P, Solan, J., & Darley, J. (2012). When Does Knowledge Become Intent? Perceiving the Minds of Wrongdoers. *Journal of Empirical Legal Studies, 9* (4), 859-892.

Murray, D. & Lombrozo, T. (2017). Effects of Manipulation on Attributions of Causation, Free Will, and Moral Responsibility. Cognitive Science, 41(2): 447-481. DOI: 10.1111/cogs.12338

Nichols, S. (2008). Sentimentalism Naturalized. In W. Sinnott-Armstrong (ed.) *The Psychology and Biology of Morality*. Cambridge, MA: MIT Press.

Nichols, S. (2002). Norms with Feeling: Towards a Psychological Account of Moral Judgment, *Cognition*, *84*, 221-236. doi:10.1016/S0010-0277(02)00048-3

Nucci, L. P, & Herman, S. (1982). Behavioral disordered children's conceptions of moral, conventional, and personal issues. *Journal of Abnormal Child Psychology, 10*, 411-426. doi: 10.1007/BF00912330

Nucci, L. & Turiel, E. (2009). Capturing the Complexity of Moral Development and Education. *Mind, Brain, & Education, 3(3),* 151-159. doi: 10.1111/j.1751-228X.2009.01065.x

Oppenheimer, D. M., Meyvis, T., & Davydenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867-872. doi:10.1016/j.jesp.2009.03.009

Ogloff, J. (1991). A Comparison of Insanity Defense Standards on Juror Decision Making. *Law & Human Behavior, 15*(5), 509-531.

Pennington, N. & Hastie, R. (1992). Explaining the Evidence: Tests of the Story Model for Juror Decision Making. *Attitudes & Social Cognition, 62*, 189-206.

Pennington, N. & Hastie, R. (1986). Evidence Evaluation in Complex Decision Making: A Story Model of Evidence Evaluation. *Journal of Personality & Social Psychology, 51*, 242-258.

Pennington, R. & Hastie, N. (1981). Juror Decision-Making Models: The Generalization Gap. *Psychological Bulletin, 89*(2), 246-287.

Perkins, R. (1939). A Rationale of Mens Rea. *Harvard Law Review, 52*(6), 905-928.

Phelan, J. C., & Link, B. G. (1998). The growing belief that people with mental illnesses are violent: The role of the dangerousness criterion for civil commitment. *Social Psychiatry & Psychiatric Epidemiology, 33*, S7-S12.

Phillips, J. & Shaw, A. (2014). Manipulating Morality: Third-Party Intentions Alter Moral Judgments by Changing Causal Reasoning, *in preparation*.

Rai, T., and Fiske, A. (2011). Moral Psychology Is Relationship Regulation: Moral Motives for Unity, Hierarchy, Equality, and Proportionality. *Psych Rev, 118*, 57-75. doi:10.1037/a0021867

Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, *11*(2), 233. doi:10.1037/a0022598.

Sadler, M., Kaye, K., & Vaughn, A. (2015). Competence and warmth stereotypes prompt mental illness stigma through emotions. *Journal of Applied Social Psychology, 45*, 602-612. doi: 10.1111/jasp.12323

Sadler, M., Meagor, E., & Kaye, K. (2012). Stereotypes of mental disorders differ in competence and warmth. *Social Science & Medicine, 74*(6), 915-922. doi:10.1016/j.socscimed.2011.12.019

Saltzman, A. (1978). Strict Criminal Liability and the United States Constitution: Substantive Criminal Law Due Process. *Wayne Law Review, 24*, 1571-1640.

Sanders, J., Kugler, M., Solan, L., & Darley, J. (2014). Must Torts Be Wrongs? An Empirical Perspective. *Wake Forest Law Review, 49*, 1-45.

Sarkissian, H. Parks, J., Tien, D., Wright, J.C., & Knobe. J. (2011) Folk moral relativism. *Mind & Language, 26* (4), 482-505. doi: 10.1111/j.1468-0017.2011.01428.x

Sayre, F. (1932). Mens Rea. *Harvard Law Review, 45*(6), 974-1026.

Schad v. Arizona, 501 U.S. 624 (1991).

Schein, C., & Gray, K. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, 22(1) 32-70. DOI: 10.1177/1088868317698288

Shafto, P., Goodman, N., & Frank, M. (2012). Learning From Others: The Consequences of Psychological Reasoning for Human Learning. *Perspectives on Psychological Science, 7*(4), 341-351. DOI: 10.1177/1745691612448481

Shariff, A. et al. (2014). Free Will and Punishment: A Mechanistic View of Human Nature Reduces Retribution. *Psychological Science*, 1-8. doi:10.1177/0956797614534693

Shen, F., Hoffman, M., Jones, O., Greene, J., & Marois, R. (2011). Sorting Guilty Minds. *New York University Law Review, 86*(5), 1306-1360.

Skeem, J., & Golding, S. (2001). Describing Jurors' Personal Conceptions of Insanity and their Relationship to Case Judgments. *Psychology Public Policy, and Law, 7*(3), 561-621.

Simester, A. (2004). "Is Strict Liability Always Wrong?" *In* Simester, A. (ed.), *Appraising Strict Liability*. Oxford, UK: Oxford University Press, 21-50.

Simons, K. (1997). When Is Strict Criminal Liability Just? *J. of Criminal Law & Criminology, 87(4)*, 1075-1137. doi: 0091-4169/97/8704-1075

Singer, R. (1989). The Resurgence of Mens Rea: The Rise and Fall of Strict Criminal Liability. *Boston College Law Review, 30(2)*, 337-408.

Smetana, J. G. (1995). Morality in context: Abstractions, ambiguities, and applications. In R. Vasta (ed.), *Annals of child development* (Vol 10, pp. 83-130). London: Jessica Kingsley.

Smetana, J. & Braeges, J. (1990). The Development of Toddlers ' Moral and Conventional Judgments. *Merrill-Palmer Quarterly, 36* (3), 329-346.

Stahlkopf, C., Males, M., & Macallair, D. (2010). Testing Incapacitation Theory: Youth Crime and Incarceration in California. *Crime & Delinquency, 56(2)*, 253-268. doi:10.1177/0011128707307227

Stanton-Ife, J. (2007). Strict Liability: Stigma and Regret. *Oxford J. of Legal Studies, 27(1)*, 151-173. doi: 10.1093/ojls/gql033

Staples v. United States, 511 US 600 (1994).

Steadman, H. J. (1981). Critically reassessing the accuracy of public perceptions of the dangerousness of the mentally ill. *Journal of Health & Social Behavior, 22*(3), 310-316.

Tisak, M. & Turiel, E. (1988). Variation in Seriousness of Transgressions and Children's Moral and Conventional Concepts. *Developmental Psychology, 24*(3), 325-357. doi: 10.1037/0012-1649.24.3.352

Torrey, E. F. (2011). Stigma and violence: Isn't it time to connect the dots? *Schizophrenia Bulletin, 37*, 892-896.

Turiel, E. (2008a). The Development of Children's Orientations toward Moral, Social, and Personal Orders: More than a Sequence in Development. *Human Development, 51*(1), 21-39. doi: 10.1159/000113154

Turiel, E. (2008b). Thought About Actions in Social Domains: Morality, Social Conventions, and Social Interactions. (2008). *Cognitive Development, 23*, 136-154. doi: 10.1016/j.cogdev.2007.04.001

Turiel, E. (1994). The development of social-conventional and moral concepts. In *Fundamental Research in Moral Development*. Puka, B. (Ed.). New York, NY: Routledge.

United States v. Morissette, 342 U.S. 246 (1952).

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

Weston, D. & Turiel, E. (1980). Act-Rule Relations: Children's Concepts of Social Rules. *Developmental Psychology, 16(5)*, 417-424. doi: 10.1037/0012-1649.16.5.417

Young, L., Cushman, F., Hauser, M., Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *PNAS*, *104* (20), 8235-8240. doi:10.1073/pnas.0701408104

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition, 120*, 202–214. DOI: 10.1016/j.cognition.2011.04.005

Young, L. & Tsoi, L. (2013). When Mental States Matter, When They Don't, and What That Means for Morality. *Social & Personality Psychology Compass, 7*, 585-604.

Chapter 2 Appendices and Supplemental Material.

Appendix A: Difference scores and mental state correlations by crime.

Crime	Censure Difference Score	Fines Difference Score	Jail Difference Score	Mental State & Moral Censure Correlation	Mental State & Fines Correlation	Mental State & Jail Correlation
Speeding	1.65	273.06	0.000	0.403	0.051	-0.019
Statutory Rape	1.60	4645.07	0.945	0.031	0.298	N/A
Migratory Bird	1.45	4580.59	0.255	0.481*	0.415	0.092
Reckless Driving	1.30	-1600.95	0.130	0.067	0.104	0.091
DUI	0.85	3330.82	0.360	0.187	0.025	0.13
Dumping	0.85	10094.39	0.285	0.666**	0.192	0.532*
Seducing a Minor	0.65	13246.42	1.395	0.083	0.031	-0.195
Minor in Possession	0.60	116.40	0.000	0.112	-0.384	N/A
1,500 Feet	0.55	6143.78	0.596	-0.086	-0.047	-0.084
Drugs to Minors	0.55	6264.11	0.435	0.364	-0.148	0.296
Sexual Battery	0.55	3560.32	0.965	0.402	0.356	0.233
Public Drunkenness	0.55	54.49	0.005	0.301	0.100	0.515*
Disturbing the Peace	0.50	683.41	0.070	0.484*	0.141	0.449*
Unregistered Firearm	0.45	6421.05	0.410	0.137	0.420	0.304
Incest	0.30	307.06	0.090	0.111	0.168	0.128
Battery	0.30	1526.40	0.185	0.054	-0.164	0.099
Burglary	0.25	859.47	0.300	0.549*	-0.014	0.132
Theft	0.20	614.03	0.120	0.560*	0.207	0.39
First Cousins	0.20	2326.00	0.265	0.211	-0.064	-0.099
Drug Distribution	0.15	716.17	0.220	0.698**	.493*	.524*

*Indicates significance at p< .05.
**Indicates significance at p< .01.

Appendix B: Arbitrariness Ratings by Crime

Arbitrariness ratings from Experiment 3. Mean ratings are followed in parentheses by standard deviations. Strict liability crimes are italicized.

Crime	Arbitrariness Rating
Theft	2.37 (2.06)
Burglary	2.38 (2.15)
Sexual Battery	2.54 (2.20)
Battery	2.61 (2.12)
Dumping	2.63 (2.07)
Reckless Driving	2.74 (2.06)
DUI	2.82 (2.07)
Seducing a Minor	2.84 (2.15)
Drug Distribution	2.87 (2.07)
Unregistered Firearm	2.87 (2.14)
Drugs to Minors	3.03 (2.15)
Speeding	3.03 (1.90)
Migratory Birds	3.05 (1.83)
Incest	3.09 (2.21)
Statutory Rape	3.17 (2.14)
1500 Feet	3.19 (2.23)
Disturbing the Peace	3.20 (1.84)
Public Drunkenness	3.43 (1.87)
Minor in Possession	3.50 (2.08)
First Cousin Marriage	3.70 (2.08)

Online Supplement: Brief Report, Do Participants Know Which Crimes are Strict Liability

Introduction

      A potential concern is that rather than tapping into intuitive moral and legal judgment, what we actually tested was American citizens' knowledge of their own legal system. While we never told participants which crimes were strict liability or when they should take mental states into account, they may have been familiar enough with our crimes to know – for example – that burglary requires the knowledge that we varied while speeding does not.

Methods

      To address this concern, we ran an additional experiment with the same population. Participants (N = 50) were asked to indicate whether they thought knowledge was required for conviction of each crime. For example, for speeding, participants indicated "true" or "false" for the following: "Speeding: Defendant must know how fast they were driving in order to be convicted of this crime. In other words, the defendant can't be convicted (or punishment will decrease) if it turns out that the defendant did not know the speed they were driving." The question corresponding to each crime referred to the relevant knowledge, specified in Tables 1 and 3.

Results

      We found no significant difference (p = .232) in the rate at which participants responded true versus false across strict liability crimes versus mens rea crimes. This finding indicates that pre-existing knowledge of the legal categories of our crimes did not drive our results, and further suggests that our findings speak to an aspect of cognition that is not merely confined to those with exposure to our legal system.

Some of the *deceived* vignettes below have a paragraph denoted with **. These paragraphs were updated in Experiment 2 to ensure that the information provided did not license further bad character evaluations of the main actor.

The statute changes in the *unspecified* stimuli were shown to participants only in Experiment 3.

Strict Liability

Statutory Rape – Knowing. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura talked about how much fun she had with her friends the previous weekend celebrating her 15th birthday. Alan, a college sophomore, had thought Laura was in college. Regardless, the two had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer asked for identification and discovered that Laura is 15 and Alan is 20.

In Alan and Laura's state, it is illegal for any person to have sex with another person under the age of 16. Alan's action constitutes a crime: statutory rape. The officer arrested Alan.

The following questions refer to Alan's actions.

Statutory Rape – Unspecified. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer asked for identification and discovered that Laura is 15 and Alan is 20.

In Alan and Laura's state, it is illegal for any person to have sex with another person under the age of 16. Alan's action constitutes a crime: statutory rape. The officer arrested Alan.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. The age of consent had been lowered to 15. Alan's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Statutory Rape – Unknowing. Laura and Alan met at a coffee shop. Alan asked the coffee shop owner about Laura, and the owner told Alan that Laura was a student at the same college Alan attended. After that first meeting, Alan and Laura ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer asked for identification and discovered that Laura is 15 and Alan is 20.

The coffee shop owner had heard Laura talking about going to an event on campus and honestly believed she was a student there. In fact, Laura had attended the event, but was never enrolled at the school.

In Alan and Laura's state, it is illegal for any person to have sex with another person under the age of 16. Alan's action constitutes a crime: statutory rape. The officer arrested Alan.

The following questions refer to Alan's actions.

Statutory Rape – Deceived. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura talked about how much fun she had with her friends the previous weekend celebrating her 18th birthday. Alan, a college sophomore, was pleased to hear that they were about the same age. Laura and Alan found that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer asked for identification and discovered that Laura is 15 and Alan is 20. Laura had lied about her age earlier in the evening to impress Alan.

** Laura had deliberately lied about her age earlier in the evening because she was angry at her parents and thought spending the evening with Alan would be a great way to get back at them. She thought if Alan knew how old she really was, he'd just take her back to her parents' house.

In Alan and Laura's state, it is illegal for any person to have sex with another person under the age of 16. Alan's action constitutes a crime: statutory rape. The officer arrested Alan.

The following questions refer to Alan's actions.

Incest – Knowing. Laura and Alan ran into each other at a coffee shop. Laura and Alan had the same father, but their father's divorce from Laura's mother had been difficult so Laura and Alan knew each other but had never been allowed to spend any time together. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found out

that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer took Laura and Alan to the police station to book them for indecent exposure. During booking, the officer pulled up Laura's and Alan's files and discovered that Laura and Alan were half-siblings. Although they were related, Alan and Laura did not regret the encounter.

In Alan and Laura's state, it is illegal for half-siblings to have sexual intercourse. Alan and Laura's action constitutes a crime: incest. The officer arrested Alan and Laura.

The following questions refer to Laura's actions.

Incest – Unspecified. Laura and Alan ran into each other at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found out that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer took Laura and Alan to the police station to book them for indecent exposure. During booking, the officer pulled up Laura's and Alan's files and discovered that Laura and Alan were half-siblings. Although they were related, Alan and Laura did not regret the encounter.

In Alan and Laura's state, it is illegal for half-siblings to have sexual intercourse. Alan and Laura's action constitutes a crime: incest. The officer arrested Alan and Laura.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with incest unless they had sex with a full sibling. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Incest – Unknowing. Laura and Alan ran into each other at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan talked about growing up as only children, and Laura talked about being raised by a single mother. Laura and Alan found out that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the

officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer took Laura and Alan to the police station to book then for indecent exposure. During booking, the officer pulled up Laura and Alan's files and discovered that Laura and Alan were half-siblings. Neither Laura nor Alan knew they had the same father.

Laura's mother had never told the father about Laura and had lost track of him after she got pregnant so did not realize he later had a son. Thus, both Alan and Laura had thought they were only children. Even after this discovery, Alan and Laura did not regret the encounter.

In Alan and Laura's state, it is illegal for half-siblings to have sexual intercourse. Alan and Laura's action constitutes a crime: incest. The officer arrested Alan and Laura.

The following questions refer to Laura's actions.

Incest – Deceived. Laura and Alan ran into each other at a coffee shop. Alan recognized Laura from photographs his father kept. Laura was Alan's father's daughter from his first marriage (i.e., a half-sibling), but the divorce had been difficult and their father had never been a part of Laura's life. Alan did not tell Laura what he knew. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura talked about being an only child raised by a single mother. Alan did not reveal that this was untrue, as he was Laura's half-sibling. Laura and Alan found out that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove her to a secluded section of the local park where the two eventually began to have consensual sex.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours before so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan and Laura having sex. The officer took Laura and Alan to the police station to book them for indecent exposure. During booking, the officer pulled up Laura's and Alan's files and discovered that Laura and Alan were half-siblings. Laura had never known she had a brother; her mother had told her she had no siblings. Even after this discovery, Alan and Laura did not regret the encounter.

In Alan and Laura's state, it is illegal for half-siblings to have sexual intercourse. Alan and Laura's action constitutes a crime: incest. The officer arrested Alan and Laura.

The following questions refer to Laura's actions.

Speeding – Knowing. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window. Alan knew the speed limit was 55 miles per hour and that he was driving over the speed limit, but Alan didn't slow down as the speedometer crept up to 70 miles per hour.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan passed by, the officer looked at the gun and noticed that Alan was going 70 miles per hour. The officer pulled Alan over.

In Alan's state, it is illegal to drive a vehicle upon a two-lane, undivided highway at a speed greater than 55 miles per hour unless the highway has been posted for a higher speed. The highway Alan was driving on was not posted for a higher speed. Alan's action constitutes a crime: speeding. The officer gave Alan a ticket.

The following questions refer to Alan's actions.

Speeding – Unspecified. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan passed by, the officer looked at the gun and noticed that Alan was going 70 miles per hour. The officer pulled Alan over.

In Alan's state, it is illegal to drive a vehicle upon a two-lane, undivided highway at a speed greater than 55 miles per hour unless the highway has been posted for a higher speed. The highway Alan was driving on was not posted for a higher speed. Alan's action constitutes a crime: speeding. The officer gave Alan a ticket.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed, due to a technicality. Speeding, even in a school zone, is no longer enough to show reckless driving. Alan's actions were therefore no longer technically reckless driving.

Please answer the following questions based on the change in statute.

Speeding – Unknowing. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window. Alan knew the speed limit was 55 miles per hour and was careful to check the speedometer periodically to make sure the car wasn't going faster than the speed limit.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan passed by, the officer looked at the gun and noticed that Alan was going 70 miles per hour. The officer pulled Alan over. Alan was surprised to hear this and told the officer the speedometer had said 55 miles per hour the entire time. Alan had never had any reason to suspect the speedometer was broken but later found out that it was.

In Alan's state, it is illegal to drive a vehicle upon a two-lane, undivided highway at a speed greater than 55 miles per hour unless the highway has been posted for a higher speed. The highway Alan was driving on was not posted for a higher speed. Alan's action constitutes a crime: speeding. The officer gave Alan a ticket.

The following questions refer to Alan's actions.

Speeding – Deceived. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window. Alan knew the speed limit was 55 miles per hour and was careful to check the speedometer periodically to make sure the car wasn't going faster than the speed limit.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan passed by, the officer looked at the gun and noticed that Alan was going 70 miles per hour. The officer pulled Alan over. Alan was surprised to hear this and told the officer the speedometer had said 55 miles per hour the entire time. Alan had never had any reason to suspect the speedometer was broken but later found out that his mechanic had intentionally tampered with it after Alan disputed a bill.

** A police officer was parked alongside the road with a speed gun pointed towards oncoming traffic. As Alan passed by, the officer looked at the gun and noticed that Alan was going 50 miles per hour. The officer pulled Alan over. Alan was surprised to hear he was speeding and told the officer the speedometer had said 35 miles per hour the entire time. Alan had never had any reason to suspect the speedometer was broken but later found out that his mechanic had intentionally tampered with it because the mechanic was angry at his boss and wanted to make the garage look bad by doing shoddy work.

In Alan's state, it is illegal to drive a vehicle upon a two-lane, undivided highway at a speed greater than 55 miles per hour unless the highway has been posted for a higher speed. The highway Alan was driving on was not posted for a higher speed. Alan's action constitutes a crime: speeding. The officer gave Alan a ticket.

The following questions refer to Alan's actions.

Drugs to Minors – Knowing. Laura had been packaging and selling illegal drugs for several years, but, recently, business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot near the local high school. The area got a lot of foot traffic, and she knew that most of the high school kids passed by there at the end of the school day. Laura thought high school kids could be good customers. The second day in the new spot, Laura was approached by Alan who was wearing a high school shirt and carrying an algebra book. Alan asked for twenty ecstasy pills. Laura could tell Alan was in high school, but sold him the drugs anyway.

The transaction was seen by an off duty FBI agent who was at the school to pick up the agent's child. The agent promptly approached Laura and Alan, took the drugs, and asked for identification. The agent discovered that Laura was 23 and Alan was 16.

In Laura's state, it is illegal for any person eighteen years or older to distribute, dispense, or possess with intent to distribute or dispense a controlled substance to a person under twenty-one years of age. Laura's action constitutes a crime: distributing drugs to minors. The agent arrested Laura.

The following questions refer to Laura's actions.

Drugs to Minors – Unspecified. Laura had been packaging and selling illegal drugs for several years, but, recently, business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. She knew the area got a lot of foot traffic. The second day in the new spot, Laura was approached by Alan. Alan asked for twenty ecstasy pills. Laura sold Alan the drugs.

The transaction was seen by an off duty FBI agent who was in the area running errands. The agent promptly approached Laura and Alan, took the drugs, and asked for identification. The agent discovered that Laura was 23 and Alan was 16 [for Experiment 3, the age was raised to 18].

In Laura's state, it is illegal for any person eighteen years or older to distribute, dispense, or possess with intent to distribute or dispense a controlled substance to a person under twenty-one years of age. Laura's action constitutes a crime: distributing drugs to minors. The agent arrested Laura.

The following questions refer to Laura's actions.

<u>Statute change</u>: After Laura's arrest, the officer was informed that the statute had been changed. A minor was now defined as anyone younger than 18. Laura's actions were therefore no longer technically eligible for a sentence increase.

Please answer the following questions based on the change in statute.

      Drugs to Minors – Unknowing. Laura had been packaging and selling illegal drugs for several years, but, recently, business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. She knew the area got a lot of foot traffic. The second day in the new spot, Laura was approached by Alan and one of Laura's regular customers. The regular customer told Laura he had recently met Alan at a bar. Alan asked for twenty ecstasy pills. Laura sold Alan the drugs.

The transaction was seen by an off duty FBI agent who was in the area running errands. The agent promptly approached Laura and Alan, took the drugs, and asked for identification. The agent discovered that Laura was 23 and Alan was 16.

Laura's regular customer had met Alan at the bar and honestly believed Alan was 21. In fact, Alan was at the bar for to see a band he liked and had a special wristband that indicated he was under 21. Laura's customer just hadn't noticed it.

In Laura's state, it is illegal for any person eighteen years or older to distribute, dispense, or possess with intent to distribute or dispense a controlled substance to a person under twenty-one years of age. Laura's action constitutes a crime: distributing drugs to minors. The agent arrested Laura.

The following questions refer to Laura's actions.

      Drugs to Minors – Deceived. Laura had been packaging and selling illegal drugs for several years, but, recently, business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. The area got a lot of foot traffic, and while she knew that most of the high school kids passed by there at the end of the school day, Laura had no plans to sell to the kids. The second day in the new spot, Laura was approached by Alan who was wearing a sports jacket and carrying a briefcase. Laura wanted to make absolutely sure Alan wasn't a high school student so she asked him how old he was. Alan said he was 22 and worked at a business nearby. Alan asked for twenty ecstasy pills. Laura sold Alan the drugs.

The transaction was seen by an off duty FBI agent who had been at the high school to pick up the agent's child. The agent promptly approached Laura and Alan, took the drugs, and asked for identification. The agent discovered that Laura was 23 and Alan was 16.

In Laura's state, it is illegal for any person eighteen years or older to distribute, dispense, or possess with intent to distribute or dispense a controlled substance to a person under twenty-one years of age. Laura's action constitutes a crime: distributing drugs to minors. The agent arrested Laura.

The following questions refer to Laura's actions.

      1,500 Feet – Knowing. Laura had been packaging and selling ecstasy for several years, but recently business had been slipping. In an effort to get new customers, Laura decided to start

selling in a new spot. She wanted a spot that got a lot of foot traffic. One of Laura's regular customers suggested a location near a train stop. Laura knew there was a school around the corner, about 800 feet away, but decided to sell at that spot anyway because the foot traffic was good.

One day a police officer on routine patrol saw Laura on the street. The officer thought Laura might be selling drugs and waited out of sight until she sold some pills to a customer. The officer arrested Laura for selling drugs and informed her she would face additional time for selling so close to a school.

In Laura's state, it is illegal for any person to sell or possess with intent to sell to another person any controlled substance within fifteen hundred feet of any public or private school. Laura's action constitutes a crime: distributing drugs with fifteen hundred feet of a school.

The following questions refer to Laura's actions.

1,500 Feet – Unspecified. Laura had been packaging and selling ecstasy for several years, but recently business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. She wanted a spot that got a lot of foot traffic. One of Laura's regular customers suggested a location near a train stop.

One day a police officer on routine patrol saw Laura on the street. The officer thought Laura might be selling drugs and waited out of sight until she sold some pills to a customer. The officer arrested Laura for selling drugs and informed her she would face additional time for selling so close to a school.: a school was around the corner, about 800 feet away.

In Laura's state, it is illegal for any person to sell or possess with intent to sell to another person any controlled substance within fifteen hundred feet of any public or private school. Laura's action constitutes a crime: distributing drugs with fifteen hundred feet of a school.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person now needed to be selling drugs within 400 feet of a school to face additional time. Laura's actions were therefore no longer technically eligible for a sentence increase.

Please answer the following questions based on the change in statute.

1,500 Feet – Unknowing. Laura had been packaging and selling ecstasy for several years, but recently business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. She wanted a spot that got a lot of foot traffic. One of Laura's regular customers suggested a location near a train stop. Laura thought that spot was around the corner from a private high school, about 800 feet away, but her customer assured her the building she was talking about was actually a charity foundation, not a school.

One day a police officer on routine patrol saw Laura on the street. The officer thought Laura might be selling drugs and waited out of sight until she sold some pills to a customer. The officer arrested Laura for selling drugs and informed her she would face additional time for selling so close to a school.

Laura's regular customer – the one who had recommended the location – really believed that the school was a charity. The customer had heard a woman say that a lot of good work went on in

106

the building, but the woman was referring to the number of low-income students the school accepted, not charity work.

In Laura's state, it is illegal for any person to sell or possess with intent to sell to another person any controlled substance within fifteen hundred feet of any public or private school. Laura's action constitutes a crime: distributing drugs with fifteen hundred feet of a school.

The following questions refer to Laura's actions.

1,500 Feet – Deceived. Laura had been packaging and selling ecstasy for several years, but recently business had been slipping. In an effort to get new customers, Laura decided to start selling in a new spot. She wanted a spot that got a lot of foot traffic. One of Laura's regular customers suggested a location near a train stop. Laura thought that spot was around the corner from a private high school, about 800 feet away, but her customer assured her the building she was talking about was actually a charity foundation, not a school.

One day a police officer on routine patrol saw Laura on the street. The officer thought Laura might be selling drugs, and knew that a small, charter school was just around the corner, about 800 feet away. The officer waited out of sight until Laura sold some pills to a customer. The officer arrested Laura for selling drugs and informed her she would face additional time for selling so close to a school.

Laura's regular customer – the one who had recommended the location - had deliberately lied to Laura about the school being a charity because the spot was really close to the customer's apartment and it would be really convenient to have Laura nearby.

In Laura's state, it is illegal for any person to sell or possess with intent to sell to another person any controlled substance within fifteen hundred feet of any public or private school. Laura's action constitutes a crime: distributing drugs with fifteen hundred feet of a school.

The following questions refer to Laura's actions.

DUI – Knowing. Laura had just gotten a promotion at work, and she and her friends decided to go out to celebrate. They went to their favorite restaurant and stayed for several hours. Laura knew she had to drive home and that she should watch what she drank. She figured out how many drinks she could have – based on her weight and height - and still be able to drive home at the end of the evening, but she was having such a good time that she stopped worrying about it. At the end of the night Laura said goodbye to her friends, got in her car, and drove home, knowing that she'd had more to drink than she should have.

A police officer was out doing a routine patrol through town. The officer saw Laura's car swerve a little and pulled her over. When the officer got to Laura's car, the officer noticed the faint smell of alcohol on Laura's breath. The officer asked Laura to take a Breathalyzer test. Laura reluctantly agreed. The test indicated that Laura's blood alcohol was above the legal limit.

In Laura's state, it is illegal for a person who has 0.08 percent or more, by weight, of alcohol in his or her blood to drive a vehicle. Laura's test showed she had 0.09. Laura's action constitutes a crime: driving while intoxicated. The officer arrested Laura.

The following questions refer to Laura's actions.

DUI – Unspecified. Laura had just gotten a promotion at work, and she and her friends decided to go out to celebrate. They went to their favorite restaurant and stayed for several hours. Laura and her friends said goodbye, and Laura started her drive home.

A police officer was out doing a routine patrol through town. The officer saw Laura's car swerve a little and pulled her over. When the officer got to Laura's car, the officer noticed the faint smell of alcohol on Laura's breath. The officer asked Laura to take a Breathalyzer test. Laura agreed. The test indicated that Laura's blood alcohol was above the legal limit.

In Laura's state, it is illegal for a person who has 0.08 percent or more, by weight, of alcohol in his or her blood to drive a vehicle. Laura's test showed she had 0.09. Laura's action constitutes a crime: driving while intoxicated. The officer arrested Laura.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed due to a measurement technicality. A person could no longer be charged with DUI unless their blood alcohol was measured at 0.10. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

DUI – Unknowing. Laura had just gotten a promotion at work, and she and her friends decided to go out to celebrate. They went to their favorite restaurant and stayed for several hours. Laura knew she had to drive home so she was very careful about how much she drank. When she ordered her first cocktail, she specifically asked the waitress how much alcohol was in it. The waitress said it contained one shot. Based on this information and her weight and height, Laura figured out how many drinks she could have and still be able to drive home at the end of the evening. By the time Laura left, her calculations told her she was fine to drive. Laura and her friends said goodbye, and Laura started her drive home.

A police officer was out doing a routine patrol through town. The officer saw Laura's car swerve a little and pulled her over. When the officer got to Laura's car, the officer noticed the faint smell of alcohol on Laura's breath. The officer asked Laura to take a Breathalyzer test. Laura agreed. The test indicated that Laura's blood alcohol was above the legal limit.

The waitress had made a mistake about how much alcohol was in Laura's cocktail. Each drink actually had two and a half shots in it, not one.

In Laura's state, it is illegal for a person who has 0.08 percent or more, by weight, of alcohol in his or her blood to drive a vehicle. Laura's test showed she had 0.09. Laura's action constitutes a crime: driving while intoxicated. The officer arrested Laura.

The following questions refer to Laura's actions.

DUI – Deceived. Laura had just gotten a promotion at work, and she and her friends decided to go out to celebrate. They went to their favorite restaurant and stayed for several hours. Laura knew she had to drive home so she was very careful about how much she drank. When she ordered her first cocktail, she specifically asked the waitress how much alcohol was in it. The waitress said it had one shot in it. Based on this information and her weight and height, Laura figured out how many drinks she could have and still be able to drive home at the end of

the evening. By the time Laura left, her calculations told her she was fine to drive. Laura and her friends said goodbye, and Laura started her drive home.

A police officer was out doing a routine patrol through town. The officer saw Laura's car swerve a little and pulled her over. When the officer got to Laura's car, the officer noticed the faint smell of alcohol on Laura's breath. The officer asked Laura to take a Breathalyzer test. Laura agreed. The test indicated that Laura's blood alcohol was above the legal limit.

The waitress had deliberately lied about how much alcohol was in Laura's cocktail because she thought that Laura might not buy as many if she knew how much alcohol was really in each one. Each drink actually had two and a half shots in it, not one.

In Laura's state, it is illegal for a person who has 0.08 percent or more, by weight, of alcohol in his or her blood to drive a vehicle. Laura's test showed she had 0.09. Laura's action constitutes a crime: driving while intoxicated. The officer arrested Laura.

The following questions refer to Laura's actions.

Unregistered Firearm. Laura had grown up in a family that kept guns in the house. Some of her family and friends even competed in target shooting competitions, so she was used to being around guns. Even though she'd been around guns her whole life, Laura had never participated in any of the competitions or paid too much attention to the guns one way or another. However, when Laura moved into her own place, she decided she wanted to keep a gun for protection and bought a gun from one of her friends.

A few days later, Laura threw a housewarming party and invited all her friends to come see her new place. Laura was giving everyone a tour when one of her friends, a police officer, noticed the gun on the bottom shelf of her nightstand. The officer was startled and asked Laura if the gun was registered. In fact, the gun was not registered.

In Laura's state, it is illegal for anyone to receive or possess a firearm which is not registered in the National Firearms Registration and Transfer Record. Laura's actions constitute a crime: possessing an unregistered firearm. The officer arrested Laura.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the firearm registration statute had been changed due to a technicality. Because of the state it happened to be manufactured in, the gun Laura had in her possession was old enough that it was not required to be registered under the new statute. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Illegal Dumping. Alan had owned and operated a small fabrication business for many years. He made custom ordered cabinets, metal pieces, and some glass pieces. He usually sold his pieces to architects building custom houses. Alan didn't have room to keep all his scraps and trash in his shop between garbage pickups so he usually just put everything in the backyard until the garbage truck came.

One day, when the garbage people arrived to pick up Alan's trash, they saw Alan hauling it from the backyard. The garbage people knew that Alan used some chemicals in his work and that Alan's house was pretty close to a river. The garbage people reported Alan to the Fish and

Game warden who confirmed that some of the substances Alan had been keeping in his backyard were on the list of prohibited substances and that these substances could have gotten into the nearby river. These substances had been placed on the list because they could potentially harm fish and wildlife.

In Alan's state, it is illegal for any person to deposit in, permit to pass into, or place where it can pass into state waters any of the prohibited substances. Alan's actions constitute a crime: illegal dumping. The warden arrested Alan.

The following questions refer to Alan's actions.

Statute Change: After Alan's arrest, the warden was informed that the illegal dumping statute had been changed. Due to a technicality, (not to a change in the anticipated risk of harming fish and wildlife), the substances in Alan's backyard were no longer on the list of prohibited substances. Alan's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Migratory Bird. Alan had been an avid hunter his entire life. His family had gone on hunting trips regularly when he was younger, and he had continued hunting with family, friends, and solo as he became an adult. One day Alan was hunting in a spot that he had had good luck at before and quickly shot a large bird.

A fish and game warden doing a routine patrol of the area spotted Alan heading back to his car. The warden noticed that Alan was carrying a bird and stopped him to make sure that Alan had a hunting permit. Alan did have a hunting permit, but when the warden got closer, she realized that the bird Alan shot was from a protected species.

In Alan's state, it is illegal for anyone to pursue, hunt, take, capture, or kill any migratory bird, or any part, nest, or egg of any such bird. Alan's actions constitute a crime: killing a protected bird. The warden arrested Alan.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. Due to a technicality (not to a change in the bird species' status), the bird Alan shot was no longer on the list of protected birds. Alan's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Mens Rea Crimes

Seducing a Minor – Knowing. Alan and Laura met in an online chat room. After that first meeting, they had three or four more conversations before the topics began to get more personal. During one conversation, Laura talked about how much fun she had with friends the previous weekend celebrating her 15th birthday. Alan, a college sophomore, had thought Laura was in college. Regardless, the two had a lot in common and had a long conversation that evening. Towards the end of that conversation, Alan sent Laura some pornographic photos hoping to make their conversation more sexual. Laura responded positively.

The next day, Laura's father was using her computer for work and found the photographs and chat log. He immediately notified the police, who traced the photographs back to Alan.

In Laura and Alan's state, it is illegal for any person to knowingly send harmful matter to a minor with the intent of arousing, appealing to, or gratifying the lust or passions or sexual desires of that person or of a minor, and with the intent, or for the purpose of seducing a minor. A minor is anyone under 16 years of age. Alan's actions constitute a crime: seducing a minor. A police officer arrested Alan.

The following questions refer to Alan's actions.

Seducing a Minor – Unspecified. Alan and Laura met in an online chat room. After that first meeting, they had three or four more conversations before the topics began to get more personal. The fourth or fifth time they talked, the two found that they had a lot in common and had a long conversation that evening. Towards the end of that conversation, Alan sent Laura some pornographic photos hoping to make their conversation more sexual. Laura responded positively.

The next day, Laura's father was using her computer for work and found the photographs and chat log. He immediately notified the police that someone was sending these photographs to his 15-year-old daughter. The police traced the photographs back to Alan.

In Laura and Alan's state, it is illegal for any person to knowingly send harmful matter to a minor with the intent of arousing, appealing to, or gratifying the lust or passions or sexual desires of that person or of a minor, and with the intent, or for the purpose of seducing a minor. A minor is anyone under 16 years of age. Alan's actions constitute a crime: seducing a minor. A police officer arrested Alan.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with seducing a minor unless the victim was under 15 years of age. Alan's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Seducing a Minor – Unknowing. Alan and Laura met in an online chat room. Alan asked a friend of his who also used that chat room if he knew anything about Laura. Alan's friend said that Laura was a student at the same college Alan attended. After that first meeting, Alan and Laura had three or four more conversations before the topics began to get more personal. The fourth or fifth time they talked, the two found that they had a lot in common and had a long conversation that evening. Towards the end of that conversation, Alan sent Laura some pornographic photos hoping to make their conversation more sexual. Laura responded positively.

The next day, Laura's father was using her computer for work and found the photographs and chat log. He immediately notified the police that someone was sending these photographs to his 15-year-old daughter. The police traced the photographs back to Alan.

Alan's friend had talked to Laura about an event she attended at Alan's college. Alan's friend had honestly believed, based on that conversation, that Laura was a college student. In fact, Laura had attended the event, but was never enrolled at the school.

In Laura and Alan's state, it is illegal for any person to knowingly send harmful matter to a minor with the intent of arousing, appealing to, or gratifying the lust or passions or sexual desires of that person or of a minor, and with the intent, or for the purpose of seducing a minor. A minor is anyone under 16 years of age. Alan's actions constitute a crime: seducing a minor. A police officer arrested Alan.

The following questions refer to Alan's actions.

Seducing a Minor – Deceived. Alan and Laura met in an online chat room. After that first meeting, they had three or four more conversations before the topics began to get more personal. During one conversation, Laura talked about how much fun she had with friends the previous weekend celebrating her 18th birthday. Alan, a college sophomore, was pleased to hear that they were about the same age. Laura and Alan had a lot in common and had a long conversation that evening. Towards the end of that conversation, Alan sent Laura some pornographic photographs hoping to make their conversation more sexual. Laura responded positively.

The next day, Laura's father was using her computer for work and found the photographs and chat log. He immediately notified the police that someone was sending these photographs to his 15-year-old daughter. Laura had lied to Alan about her age. The police traced the photographs back to Alan.

In Laura and Alan's state, it is illegal for any person to knowingly send harmful matter to a minor with the intent of arousing, appealing to, or gratifying the lust or passions or sexual desires of that person or of a minor, and with the intent, or for the purpose of seducing a minor. A minor is anyone under 16 years of age. Alan's actions constitute a crime: seducing a minor. A police officer arrested Alan.

The following questions refer to Alan's actions.

First Cousins – Knowing. Laura and Alan have known each other their whole lives, but have only been dating for the past few months. They haven't told anyone about their relationship because they are first cousins and know that their friends and family won't approve. Laura and Alan know that as first cousins, they cannot marry in their state but decide to do so anyway when Laura gets pregnant. They go to a justice of the peace and are married.

The county clerk who processes Laura and Alan's marriage license knows their family and knows that Laura and Alan are first cousins and that Laura is pregnant. The clerk reports the issue to the police. The police verify these facts.

In Laura and Alan's state, it is illegal for first cousins under the age of 65 to marry unless they have presented proof to a judge that one of them is unable to reproduce. Laura and Alan's marriage is annulled. Despite this turn of events, Laura and Alan continue their relationship.

The following questions refer to Laura's actions.

First Cousins – Unspecified. Laura and Alan have known each other their whole lives, but have only been dating for the past few months. They haven't told anyone about their relationship because they know that their friends and family won't approve. Laura and Alan decide to get married when Laura gets pregnant. They go to a justice of the peace and are married.

The county clerk who processes Laura and Alan's marriage license knows their family and knows that Laura and Alan are first cousins and that Laura is pregnant. The clerk reports the issue to the police. The police verify these facts.

In Laura and Alan's state, it is illegal for first cousins under the age of 65 to marry unless they have presented proof to a judge that one of them is unable to reproduce. Laura and Alan's marriage is annulled. Despite this turn of events, Laura and Alan continue their relationship.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. It is no longer a crime for first cousins to marry. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

First Cousins – Unknowing. Laura and Alan have known each other their whole lives, but have only been dating for the past few months. They haven't told their families about their relationship because their families have never gotten along, and Laura and Alan know they won't approve of the relationship. Laura and Alan decide to get married when Laura gets pregnant. They go to a justice of the peace and are married.

The county clerk who processes Laura and Alan's marriage license has known Laura and Alan's families for years and knows that Laura and Alan's mothers are sisters. The clerk also knows that Laura is pregnant. The clerk reports the issue to the police. The police verify these facts.

Laura and Alan's mothers had the same father but different mothers. It was not a subject that the two women liked to discuss. They had never told Laura and Alan because it was embarrassing for them to talk about, and they had no idea Laura and Alan were seeing each other.

In Laura and Alan's state, it is illegal for first cousins under the age of 65 to marry unless they have presented proof to a judge that one of them is unable to reproduce. Laura and Alan's marriage is annulled.

The following questions refer to Laura's actions.

First Cousins – Deceived. Laura and Alan have known each other their whole lives, but have only been dating for the past few months. They haven't told their families about their relationship because their families have never gotten along, and Laura and Alan know they won't approve of the relationship. Alan has never told Laura that the reason their families don't get along is because their mothers are sisters and had a falling out years ago. Laura has no idea she and Alan are first cousins. Laura and Alan decide to get married when Laura gets pregnant. They go to a justice of the peace and are married.

The county clerk who processes Laura and Alan's marriage license has known Laura and Alan's families for years and knows that Laura and Alan's mothers are sisters. The clerk also knows that Laura is pregnant. The clerk reports the issue to the police. The police verify these facts.

In Laura and Alan's state, it is illegal for first cousins under the age of 65 to marry unless they have presented proof to a judge that one of them is unable to reproduce. Laura's mother had always said she was an only child, so Laura had no idea that she and Alan were first cousins.

Laura and Alan's marriage is annulled. Despite this turn of events, Laura and Alan continue their relationship.

The following questions refer to Laura's actions.

Reckless Driving – Knowing. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window. Alan knew the speed limit was 55 miles per hour, but Alan didn't care as the speedometer crept up to 70 miles per hour. Alan also saw the signs that said he was driving in a school zone. As the car started moving faster and faster, Alan enjoyed the speed more and more.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan flew by, the officer looked at the gun and noticed that Alan was going over the speed limit in a school zone. In fact, Alan was driving so fast that the officer was concerned about the safety of the children who were just getting out of school. The officer pulled Alan over.

In Alan's state, it is illegal to drive a vehicle upon a highway in willful or wanton disregard for the safety of persons or property. Due to Alan's proximity to the school zone, Alan's conduct constitutes a crime: reckless driving. The police officer gave Alan a ticket.

The following questions refer to Alan's actions.

Reckless Driving – Unspecified. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan flew by, the officer looked at the gun and noticed that Alan was going 70 miles per hour in a school zone. In fact, Alan was driving so fast that the officer was concerned about the safety of the children who were just getting out of school nearby. The officer pulled Alan over.

In Alan's state, it is illegal to drive a vehicle upon a highway in willful or wanton disregard for the safety of persons or property. Due to Alan's proximity to the school zone, Alan's conduct constitutes a crime: reckless driving. The police officer gave Alan a ticket.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed, due to a technicality. Speeding, even in a school zone, is no longer enough to show reckless driving. Alan's actions were therefore no longer technically reckless driving.

Please answer the following questions based on the change in statute.

Reckless Driving – Unknowing. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan flew by, the officer looked at the gun and noticed that Alan was going 70 miles per hour in a school zone. In fact, Alan was driving so fast that the officer was concerned about the safety of the children who were just getting out of school nearby. The officer pulled Alan over.

Alan did not know he was driving in a school zone because the signs had been blown down by a storm a few days before. He would not have driven that fast had he known.

In Alan's state, it is illegal to drive a vehicle upon a highway in willful or wanton disregard for the safety of persons or property. Due to Alan's proximity to the school zone, Alan's conduct constitutes a crime: reckless driving. The police officer gave Alan a ticket.

The following questions refer to Alan's actions.

Reckless Driving – Deceived. Alan was driving along a two lane highway. It was a nice day, and he was enjoying watching the scenery pass by outside the window. Alan knew the speed limit was 55 miles per hour and was careful to make sure that he wasn't driving in a way that might put people in danger.

A police officer was parked alongside the highway with a speed gun pointed towards oncoming traffic. As Alan drove by, the officer looked at the gun and noticed that Alan was going 70 miles per hour in a school zone. In fact, Alan was driving so fast that the officer was concerned about the safety of the children who were just getting out of school. The officer pulled Alan over. Alan had no idea he was in a school zone because an expelled student had stolen the signs a few days before. The student was angry about being expelled and hoped removing the sign would cause an accident. Alan would not have driven that fast had he known he was in a school zone.

In Alan's state, it is illegal to drive a vehicle upon a highway in willful or wanton disregard for the safety of persons or property. Due to Alan's proximity to the school zone, Alan's conduct constitutes a crime: reckless driving. The police officer gave Alan a ticket.

The following questions refer to Alan's actions.

Drugs Distribution – Knowing. Laura and Alan had been friends for several years. Laura had always been impressed that Alan managed to make a living working from home. Recently, though, Laura learned that Alan sold ecstasy which she knew was an illegal drug in her state. When Alan asked Laura if Laura would be willing to help with work, despite what Laura had recently learned about Alan's business, Laura readily agreed. Alan asked Laura to deliver twenty ecstasy pills to a customer. Alan gave Laura the name and address of the customer and instructed Laura not to leave the package with anyone other than the customer.

Laura went to the address Alan provided to deliver the package to the customer. The delivery was witnessed by FBI agents who had the customer's apartment under surveillance. The agents promptly approached Laura and the customer, took the drugs, and asked for identification.

It is illegal for any person to knowingly and intentionally distribute, dispense, or possess with the intent to distribute or dispense a controlled substance. The agents held Laura for questioning.

The following questions refer to Laura's actions.

Drugs Distribution – Unspecified. Laura and Alan had been friends for several years. Laura had always been impressed that Alan managed to make a living working from home. One day, Alan asked Laura if Laura would be willing to help with work. Laura readily agreed. Alan asked Laura to deliver a package to a customer. Alan gave Laura the name and address of the customer and instructed Laura not to leave the package with anyone other than the customer.

Laura went to the address Alan provided to deliver the package to the customer. The delivery was witnessed by FBI agents who had the customer's apartment under surveillance. The agents promptly approached Laura and the customer, took the package which contained twenty ecstasy pills, and asked for identification.

It is illegal for any person to knowingly and intentionally distribute, dispense, or possess with the intent to distribute or dispense a controlled substance. The agents held Laura for questioning.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person cannot be charged with drug distribution unless the customer has accepted the drugs and exchanged money for them. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Drugs Distribution – Unknowing. Laura and Alan had been friends for several years. Laura had always been impressed that Alan managed to make a living working from home, though she'd never been clear about what he did. One day, Alan's girlfriend, Diane, asked Laura if Laura would be willing to help with Alan's work while he was out of town. Laura readily agreed. Diane explained that Alan sold homeopathic medicine which needed to be delivered to a customer. She gave Laura the name and address of the customer and instructed Laura not to leave the package with anyone other than the customer.

Laura went to the address Alan provided to deliver the package to the customer. The delivery was witnessed by FBI agents who had the customer's apartment under surveillance. The agents promptly approached Laura and the customer, took the package which contained twenty ecstasy pills, and asked for identification.

Diane had no idea that Alan was selling ecstasy; she honestly believed that Alan was selling homeopathic medicine. She had met Alan at a homeopathic medicine seminar and when she saw the pills at his house, she had assumed that's what they were, though she'd never asked Alan directly. Alan had attended the seminar because he was interested in homeopathic medicine, but he had never realized that Diane thought that's what he sold. He had always assumed Diane was aware he sold ecstasy.

It is illegal for any person to knowingly and intentionally distribute, dispense, or possess with the intent to distribute or dispense a controlled substance. The agents held Laura for questioning.

The following questions refer to Laura's actions.

Drug Distribution – Deceived. Laura and Alan had been friends for several years. Laura had always been impressed that Alan managed to make a living working from home. Recently, Alan had told Laura he sold homeopathic medicine. One day, Alan asked Laura if she would be willing to help him with his work. Laura readily agreed. Alan asked Laura to deliver an order to a customer. Alan gave Laura the name and address of the customer and instructed Laura not to leave the order with anyone other than the customer.

Laura went to the address Alan provided to deliver the package to the customer. The delivery was witnessed by FBI agents who had the customer's apartment under surveillance. The agents promptly approached Laura and the customer, took the package which contained twenty ecstasy

pills, and asked for identification. It wasn't until that moment that Laura found out that she had been delivering ecstasy. Alan had intentionally lied to her so she would help him.

It is illegal for any person to knowingly and intentionally distribute, dispense, or possess with the intent to distribute or dispense a controlled substance. The agents held Laura for questioning.

The following questions refer to Laura's actions.

Sexual Battery – Knowing. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park. Alan made a move and Laura said no. Alan ignored Laura and started fondling her over her clothes.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours earlier so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan running his hands over the front of Laura's shirt. Laura told the police officer that Alan groped her over her clothes.

In Alan and Laura's state, it is illegal for any person to touch an intimate part of another person over their clothes, if the touching is against the will of the person touched, and is for the specific purpose of sexual arousal, sexual gratification, or sexual abuse. Alan's action constitutes a crime: sexual battery. The officer arrested Alan.

The following questions refer to Alan's actions.

Sexual Battery – Unspecified. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park and parked the car for a while.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours earlier so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan running his hands over the front of Laura's shirt. Laura told the police officer that Alan groped her over her clothes.

In Alan and Laura's state, it is illegal for any person to touch an intimate part of another person over their clothes, if the touching is against the will of the person touched, and is for the specific purpose of sexual arousal, sexual gratification, or sexual abuse. Alan's action constitutes a crime: sexual battery. The officer arrested Alan.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with sexual battery unless he or she touched the skin of the victim. Alan's actions were therefore no longer technically a sexual battery.

Please answer the following questions based on the change in statute.

Sexual Battery – Unknowing. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. At dinner, a friend they knew from the coffee shop, Tim, came over to say hello. When Laura excused herself to go to the bathroom, Tim whispered to Alan that he overheard Laura saying she hoped Alan would make a move that night to initiate a romantic relationship. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park. Alan made a move on Laura, and she resisted. Remembering what Tim had told him earlier, Alan ignored Laura and started fondling her over her clothes.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours earlier so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan running his hands over the front of Laura's shirt. Laura told the police officer that Alan groped her over her clothes.

Tim, the friend from the coffee shop, had honestly believed he was giving Alan good advice. Tim had overheard Laura talking on the phone about a relationship, but she had been talking about characters in a television series, not Alan and herself.

In Alan and Laura's state, it is illegal for any person to touch an intimate part of another person over their clothes, if the touching is against the will of the person touched, and is for the specific purpose of sexual arousal, sexual gratification, or sexual abuse. Alan's action constitutes a crime: sexual battery. The officer arrested Alan.

The following questions refer to Alan's actions.

Sexual Battery – Deceived. Laura and Alan met at a coffee shop. After that first meeting, the two ran in to each other at the coffee shop three or four more times before deciding to have dinner together at a restaurant nearby. During dinner, Laura and Alan found that they had a lot in common, and dinner lasted for several hours. At dinner, a friend they knew from the coffee shop, Tim, came over to say hello. When Laura excused herself to go to the bathroom, Tim whispered to Alan that he overheard Laura saying she hoped Alan would make a move that night to initiate a romantic relationship. After dinner, Alan offered to take Laura for a ride and drove to a secluded section of the local park. Alan made a move on Laura, and she resisted. Remembering what Tim had told him earlier, Alan ignored Laura and started fondling her over her clothes.

A local police officer was doing a routine patrol of the park and noticed Alan's car. The park had closed several hours earlier so the officer went to the car to tell the driver to leave. When the officer arrived at Alan's car, the officer found Alan running his hands over the front of Laura's shirt. Laura told the police officer that Alan groped her over her clothes.

Tim had lied to Alan because he had always liked Laura and was jealous that she went on a date with Alan instead of him; he thought that Laura would reject Alan if Alan pursued her too aggressively.

In Alan and Laura's state, it is illegal for any person to touch an intimate part of another person over their clothes, if the touching is against the will of the person touched, and is for the specific

purpose of sexual arousal, sexual gratification, or sexual abuse. Alan's action constitutes a crime: sexual battery. The officer arrested Alan.

The following questions refer to Alan's actions.

Burglary – Knowing. Laura worked at a nursing home for retired veterans. Laura had always been impressed by the way the home was decorated. She especially admired a particular vase in one of the common rooms. The vase was a family heirloom of a resident who donated it to the home. One day, Laura noticed the vase was missing. She asked a co-worker what happened to it. The co-worker told her it had been moved while the room was being redecorated and would be put back when it was safe. The next day, Laura drove to work, went to the back storage room where the vase was being kept, and took it home.

When the home's director noticed the vase was missing, she reviewed the security tape and saw Laura take it from the storage room. She called the police.

In Laura's state, it is illegal for any person to enter a house or other building with intent to commit grand or petit larceny or any felony. Laura's actions constitute a crime: burglary. The police arrest Laura.

The following questions concern Laura's actions.

Burglary – Unspecified. Laura worked at a nursing home for retired veterans. Laura had always been impressed by the way the home was decorated. She especially admired a particular vase in one of the common rooms. The vase was a family heirloom of a resident who donated it to the home. One day, Laura noticed the vase had been moved to a back storage room where the home kept things, including everything they planned on getting rid of. The next day, Laura drove to work, went to the back storage room where the vase was being kept, and took it home.

When the home's director noticed the vase was missing, she reviewed the security tape and saw Laura take it from the storage room. She called the police.

In Laura's state, it is illegal for any person to enter a house or other building with intent to commit grand or petit larceny or any felony. Laura's actions constitute a crime: burglary. The police arrest Laura.

The following questions concern Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with burglary if they had not damaged the property and were able to return it in its original condition. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Burglary – Unknowing. Laura worked at a nursing home for retired veterans. Laura had always been impressed by the way the home was decorated. She especially admired a particular vase in one of the common rooms. The vase was a family heirloom of a resident who donated it to the home. One day, Laura noticed the vase was missing. She asked a co-worker what happened to it. The co-worker told her they were redecorating and had decided to get rid of it. It was in a back storage room until the next garbage pickup. The next day, Laura drove to work, went to the back storage room where the vase was being kept, and took it home.

119

When the home's director noticed the vase was missing, she reviewed the security tape and saw Laura take it from the storage room. She called the police.

Laura's co-worker honestly thought the director meant to get rid of the vase. The co-worker had overheard the home's director saying the vase would not fit the style of the redecorated room. The director had indeed said this, but intended to move the vase to a different room. It was in the storage room temporarily for safekeeping.

In Laura's state, it is illegal for any person to enter a house or other building with intent to commit grand or petit larceny or any felony. Laura's actions constitute a crime: burglary. The police arrest Laura.

The following questions concern Laura's actions.

Burglary – Deceived. Laura worked at a nursing home for retired veterans. Laura had always been impressed by the way the home was decorated. She especially admired a particular vase in one of the common rooms. The vase was a family heirloom of a resident who donated it to the home. One day, Laura noticed the vase was missing. She asked a co-worker what happened to it. The co-worker told her they were redecorating and had decided to get rid of it. It was in a back storage room until the next garbage pickup. The next day, Laura drove to work, went to the back storage room where the vase was being kept, and took it home.

In fact, the vase had only been moved to keep it safe during the redecorating. When the home's director noticed the vase was missing, she reviewed the security tape and saw Laura take it from the storage room. She called the police.

Laura's co-worker deliberately lied to Laura because she was angry at the home's director. The co-worker knew Laura liked the vase and hoped she would take it and get the director in trouble for losing it.

In Laura's state, it is illegal for any person to enter a house or other building with intent to commit grand or petit larceny or any felony. Laura's actions constitute a crime: burglary. The police arrest Laura.

The following questions concern Laura's actions.

Theft – Knowing. Alan had always admired his neighbor Tim's electric saw. He was doing a renovation project and could really use a better saw than the one he had. Tim's was an older model, but from a good manufacturer. One day, Alan saw Tim working with the saw in his driveway. Alan waited until Tim went inside and picked it up off the sidewalk and took it to his own garage.

A few minutes later, Tim came back outside and noticed his saw was gone. He looked all over his garage for it, but couldn't find it anywhere and had to abandon the project he was working on. The next day, Tim saw it in Alan's garage and called the police.

In Alan's state it is illegal to feloniously steal, take, carry, lead, or drive away the personal property of another. Alan's actions constitute a crime: theft. The police arrested Alan.

Theft – Unspecified. Alan had always admired his neighbor Tim's electric saw. He was doing a renovation project and could really use a better saw than the one he had. Tim's was an older model, but from a good manufacturer. One day as Alan was rolling his trash out for

collection, he noticed the saw laying on the sidewalk in front of Tim's house, a few feet away from Tim's trash cans. Alan picked it up and took it to his own garage.

A few minutes later, Tim came back outside and noticed his saw was gone. He looked all over his garage for it, but couldn't find it anywhere and had to abandon the project he was working on. The next day, Tim saw it in Alan's garage and called the police.

In Alan's state it is illegal to feloniously steal, take, carry, lead, or drive away the personal property of another. Alan's actions constitute a crime: theft. The police arrested Alan.

The following questions concern Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with theft if the property was lying on the sidewalk. Alan's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Theft – Unknowing. Alan had always admired his neighbor Tim's electric saw. He was doing a renovation project and could really use a better saw than the one he had. Tim's was an older model, but from a good manufacturer. One day, Alan noticed the saw laying on the sidewalk in front of Tim's house. Tim's son was in the yard and Alan asked him why the saw was lying there. Tim's son said his father was buying a new one and had put the saw out for the garbage men. Alan picked it up and took it to his own garage.

A few minutes later, Tim came back outside and noticed his saw was gone. He looked all over his garage for it, but couldn't find it anywhere and had to abandon the project he was working on. The next day, Tim saw it in Alan's garage and called the police. Tim had never had any intention of getting rid of his saw, he had just set it down while he went into his house to have lunch. Tim's son had heard his father complaining about the saw and talking about getting a new one, so he had assumed his father had left it out because he was actually getting rid of it.

In Alan's state it is illegal to feloniously steal, take, carry, lead, or drive away the personal property of another. Alan's actions constitute a crime: theft. The police arrested Alan.

The following questions concern Alan's actions.

Theft – Deceived. Alan had always admired his neighbor Tim's electric saw. He was doing a renovation project and could really use a better saw than the one he had. Tim's was an older model, but from a good manufacturer. One day, Alan noticed the saw laying on the sidewalk in front of Tim's house. Tim's son was in the yard and Alan asked him why the saw was lying there. Tim's son said his father was buying a new one and had put the saw out for the garbage men. Alan picked it up and took it to his own garage.

A few minutes later, Tim came back outside and noticed his saw was gone. He looked all over his garage for it, but couldn't find it anywhere and had to abandon the project he was working on. The next day, Tim saw it in Alan's garage and called the police. Tim had never had any intention of getting rid of his saw, he had just set it down while he went into his house to have lunch. Tim's son deliberately lied because he was angry at his father for grounding him the previous week.

In Alan's state it is illegal to feloniously steal, take, carry, lead, or drive away the personal property of another. Alan's actions constitute a crime: theft. The police arrested Alan.

The following questions concern Alan's actions.

Battery – Knowing. Laura's friend Tina was always talking about the yelling matches she had with her boyfriend. Laura encouraged Tina to just breakup with him, but Tina never did. One day, Laura went to Tina's apartment and could hear loud voices through the door. She burst in and could tell that it was just the same old yelling they always did. She knew Tina wasn't in any danger, but Laura had just had enough and lunged at Tina's boyfriend, hitting him in the face.

One of the neighbors heard the commotion and called the police.

In Laura's state it is illegal for any person to willfully use force or violence upon the person of another unless they believe it is necessary to defend themselves or others. Laura's actions constitute a crime: battery. The police arrested Laura.

The following questions concern Laura's actions.

Battery – Unspecified. Laura's friend Tina was always talking about the yelling matches she had with her boyfriend. Laura encouraged Tina to just breakup with him, but Tina never did. One day, Laura went to Tina's apartment and could hear loud voices through the door. She walked in and lunged at Tina's boyfriend, hitting him in the face.

One of the neighbors heard the commotion and called the police.

In Laura's state it is illegal for any person to willfully use force or violence upon the person of another unless they believe it is necessary to defend themselves or others. Laura's actions constitute a crime: battery. The police arrested Laura.

The following questions concern Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with battery unless they had caused serious injury. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Battery – Unknowing. Laura's friend Tina was always talking about the yelling matches she had with her boyfriend. Laura encouraged Tina to just breakup with him, but Tina never did. One day, Laura went to Tina's apartment and could hear loud voices through the door. Laura walked in and saw a large red mark on Tina's face and thought that Tina's boyfriend had hit her. Laura believed Tina was in danger and lunged at the boyfriend, hitting him in the face.

One of the neighbors heard the commotion and called the police. It turned out that the boyfriend never hit Tina. She had been hit in a kickboxing class shortly before the argument.

In Laura's state it is illegal for any person to willfully use force or violence upon the person of another unless they believe it is necessary to defend themselves or others. Laura's actions constitute a crime: battery. The police arrested Laura.

The following questions concern Laura's actions.

Battery – Deceived. Laura's friend Tina was always talking about the yelling matches she had with her boyfriend. Laura encouraged Tina to just breakup with him, but Tina never did. One day, Laura went to Tina's apartment and could hear loud voices through the door. Laura walked in and saw a large red mark on Tina's face. Tina looked at Laura and said "he hit me." Laura believed Tina was in danger and lunged at Tina's boyfriend, hitting him in the face.

One of the neighbors heard the commotion and called the police. It turned out that the boyfriend never hit Tina. She had been hit in a kickboxing class shortly before the argument. Tina had deliberately lied because she was so angry at her boyfriend and hoped Laura would hit him.

In Laura's state it is illegal for any person to willfully use force or violence upon the person of another unless they believe it is necessary to defend themselves or others. Laura's actions constitute a crime: battery. The police arrested Laura.

The following questions concern Laura's actions.

Disturbing the Peace. Laura and her friends were at one of their favorite band's concerts. The concert was packed and it was difficult to see the stage. Laura was trying to get closer to the stage when she bumped into another girl and knocked her down. The other girl wasn't hurt, but got up in Laura's face yelling. Laura was angry that the other girl was yelling when the whole thing had been an accident. Laura was so angry that she asked the other girl if she wanted to fight. The other girl did not respond to the question but continued to yell at Laura for knocking her down. Laura continued to ask the girl if she wanted to fight.

A security guard noticed confrontation from across the room and came to break it up. When the security guard asked the girls to step away from each other, the other girl complied but Laura ignored the security guard. The guard asked again, and Laura continued yelling her question.

In Laura's state, it is illegal for any person to challenge another person in a public place to fight. Laura's actions constitute a crime: disturbing the peace. The security guard detained Laura and called the police. The officer arrested Laura.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with disturbing the peace for challenging someone to fight unless they also physically initiated a fight. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

Public Drunkenness. Alan and his friends were at a sports bar watching their local baseball team. It was an important game, and it wasn't going well. Alan kept ordering more drinks as the game got worse and worse. Eventually the bartender told Alan he wasn't going to serve him any more drinks because he was too drunk. Alan said he would climb over the bar and punch the bartender if he didn't serve him anymore drinks. The bartender thought Alan really might and called the police.

In Alan's state, it is illegal for any person to be found in any public place under the influence of intoxicating liquor, drugs, or controlled substances if that person is in a condition that he or she

is unable to exercise care for his or her own safety or the safety of others. Alan's actions constitute a crime: public drunkenness. The police arrested Alan.

The following questions refer to Alan's actions.

Statute change: After Alan's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with public drunkenness unless they had made some physical movement to support their threat. Alan's actions were therefore no longer technically public drunkenness.

Please answer the following questions based on the change in statute.

Minor in Possession. Laura was having a barbeque with her older sister and some of her sister's friends in the local park. It was a beautiful day and just before dinnertime so the park was packed with people, but Laura and her group had claimed one of the public grills. One of Laura's sister's friends opened a cooler and offered everyone something to drink. Laura reached in and took a can.

A local police officer was doing a routine patrol around the park when he noticed Laura and her group. He could see that several of them had beers in their hands and thought they all looked pretty young. Laura hadn't opened the beer can in her hand yet when the officer approached the group and asked for their IDs. Everyone except Laura was over 21. The officer arrested Laura.

In Laura's state, it is illegal for any person under the age of 21 years to have any alcoholic beverage in his or her possession on any street or highway or in any public place or in any place open to the public. Laura's actions constitute a crime: minor in possession of alcohol.

The following questions refer to Laura's actions.

Statute change: After Laura's arrest, the officer was informed that the statute had been changed. A person could no longer be charged with minor in possession of alcohol unless the minor had taken a drink of the beverage. Laura's actions were therefore no longer technically a crime.

Please answer the following questions based on the change in statute.

<div align="center">**Online Supplement: Experiment 2 Knowledge Manipulation**</div>

Strict Liability

   Dumping – Bad Info. Alan had owned and operated a small fabrication business for many years. He made custom ordered cabinets, metal pieces, and some glass pieces. He usually sold his pieces to architects building custom houses. Alan didn't have room to keep all his scraps and trash in his shop between garbage pickups so he usually just put everything in the backyard until the garbage truck came.

One day, when the garbage people arrived to pick up Alan's trash, they saw Alan hauling it from the backyard. The garbage people knew that Alan used some chemicals in his work and that Alan's house was pretty close to a river. The garbage people reported Alan to the Fish and Game warden who confirmed that some of the substances Alan had been keeping in his backyard were on the list of prohibited substances and that these substances could have gotten into the nearby river. These substances had been placed on the list because they could potentially harm fish and wildlife.

Alan did not know that the chemicals he was keeping out back were on the list of prohibited substances. Alan knew that there was a list and had asked a friend who also ran a fabrication business where he could find it. The friend had given Alan a copy of the list. The friend did not realize that the list he gave Alan was an older, out-of-date list that did not include all the chemicals on the current list, including some that Alan used.

In Alan's state, it is illegal for any person to deposit in, permit to pass into, or place where it can pass into state waters any of the prohibited substances. Alan's actions constitute a crime: illegal dumping. The warden arrested Alan.

The following questions refer to Alan's actions.

   Unregistered Firearm – Bad Info. Laura had grown up in a family that kept guns in the house. Some of her family and friends even competed in target shooting competitions, so she was used to being around guns. Even though she'd been around guns her whole life, Laura had never participated in any of the competitions or paid too much attention to the guns one way or another. However, when Laura moved into her own place, she decided she wanted to keep a gun for protection and bought a gun from one of her friends.

A few days later, Laura threw a housewarming party and invited all her friends to come see her new place. Laura was giving everyone a tour when one of her friends, a police officer, noticed the gun on the bottom shelf of her nightstand. The officer was startled and asked Laura if the gun was registered. In fact, the gun was not registered.

Laura did not realize that she needed to register her gun. When she purchased the gun, she asked her friend if it was necessary that she get it registered. The friend told her that due to the gun's size and model, it was not necessary for her to register it. The friend believed he was telling Laura the truth, but he was actually mistaken.

In Laura's state, it is illegal for anyone to receive or possess a firearm which is not registered in the National Firearms Registration and Transfer Record. Laura's actions constitute a crime: possessing an unregistered firearm. The officer arrested Laura.

The following questions refer to Laura's actions.

Birds – Bad Info. Alan had been an avid hunter his entire life. His family had gone on hunting trips regularly when he was younger, and he had continued hunting with family, friends, and solo as he became an adult. One day Alan was hunting in a spot that he had had good luck at before and quickly shot a large bird.

A fish and game warden doing a routine patrol of the area spotted Alan heading back to his car. The warden noticed that Alan was carrying a bird and stopped him to make sure that Alan had a hunting permit. Alan did have a hunting permit, but when the warden got closer, she realized that the bird Alan shot was from a protected species.

Alan had not realized that the bird was from a protected species. He knew that some birds in the area were and had consulted a list of them before his hunting trip. Alan did not realize that the list he consulted was out-of-date and did not contain all the species currently protected, including the one he shot.

In Alan's state, it is illegal for anyone to pursue, hunt, take, capture, or kill any migratory bird, or any part, nest, or egg of any such bird. Alan's actions constitute a crime: killing a protected bird. The warden arrested Alan.

The following questions refer to Alan's actions.

Mens Rea

Disturbing the Peace – Bad Info. Laura and her friends were at one of their favorite band's concerts. The concert was packed and it was difficult to see the stage. Laura was trying to get closer to the stage when she bumped into another girl and knocked her down. The other girl wasn't hurt, but got up in Laura's face yelling. Laura was angry that the other girl was yelling when the whole thing had been an accident. Laura was so angry that she asked the other girl if she wanted to fight. The other girl did not respond to the question but continued to yell at Laura for knocking her down. Laura continued to ask the girl if she wanted to fight.

A security guard noticed the confrontation from across the room and came to break it up. When the security guard asked the girls to step away from each other, the other girl complied but Laura ignored the security guard. The guard asked again, and Laura continued yelling her question.

Laura did not realize that her actions would disturb the peace. She had a friend who had been arrested for it, and had asked her friend what actions were considered disturbing the peace. Her friend had told her that you can't be charged with disturbing the peace unless you are being loud

and disruptive, and Laura had been careful not to be too loud when asking the girl whether she wanted to fight. The friend had believed this was true, but she was mistaken.

In Laura's state, it is illegal for any person to challenge another person in a public place to a fight. Laura's actions constitute a crime: disturbing the peace. The security guard detained Laura and called the police. The officer arrested Laura.

The following questions refer to Laura's actions.

Drunk in Public-Bad Info. Alan and his friends were at a sports bar watching their local baseball team. It was an important game, and it wasn't going well. Alan kept ordering more drinks as the game got worse and worse. Eventually the bartender told Alan he wasn't going to serve him any more drinks because he was too drunk. Alan said he would climb over the bar and punch the bartender if he didn't serve him anymore drinks. The bartender thought Alan really might and called the police.

Alan didn't realize his actions would be considered public drunkenness. Alan had a friend who had been arrested for public drunkenness and had ask his friend what actions were considered public drunkenness. His friend had said that you could only be charged with public drunkenness in a public space like a street or a public park, not in a bar. The friend thought he was telling the truth, but he was actually mistaken.

In Alan's state, it is illegal for any person to be found in any public place under the influence of intoxicating liquor, drugs, or controlled substances if that person is in a condition that he or she is unable to exercise care for his or her own safety or the safety of others. Alan's actions constitute a crime: public drunkenness. The police arrested Alan.

The following questions refer to Alan's actions.

Minor in Possession. Laura was having a barbeque with her older sister and some of her sister's friends in the local park. It was a beautiful day and just before dinnertime so the park was packed with people, but Laura and her group had claimed one of the public grills. One of Laura's sister's friends opened a cooler and offered everyone something to drink. Laura reached in and took a can.

A local police officer was doing a routine patrol around the park when he noticed Laura and her group. He could see that several of them had beers in their hands and thought they all looked pretty young. Laura hadn't opened the beer can in her hand yet when the officer approached the group and asked for their IDs. Everyone except Laura was over 21. The officer arrested Laura.

Laura didn't realize her actions would violate the law. She had only taken the can to fit in and didn't really plan to drink it. Her older sister had told her that as long as she wasn't drinking it, she wasn't breaking the law. Laura's sister had been told that by a bartender who was kidding, but the sister didn't realize the bartender was joking and thought it was true.

In Laura's state, it is illegal for any person under the age of 21 years to have any alcoholic beverage in his or her possession on any street or highway or in any public place or in any place open to the public. Laura's actions constitute a crime: minor in possession of alcohol.

The following questions refer to Laura's actions.

Strict Liability: Knowledge Manipulation

DUI. Suppose that Laura had actually realized, when she got in the car, that her actions broke the law – that is, that her blood alcohol was over the legal limit. And suppose that she decided to drive anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Speeding. Suppose that Alan had actually realized, while he was driving, that his actions broke the law – that is, that he was driving over the speed limit. And suppose that he decided to drive that speed anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Statutory Rape. Suppose that Alan had actually realized, when he had sex with Laura, that his actions broke the law – that is, that Laura was 15. And suppose that he decided to have sex with her anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

1500 Feet. Suppose that Laura had actually realized, when she began selling on that corner, that her actions broke the law – that is, that she was within 1500 of a school. And suppose that she decided to sell there anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Incest. Suppose that Laura had actually realized, when she had sex with Alan, that her actions broke the law – that is, that she was having sex with her half-brother. And suppose that she decided to have sex with him anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Drugs to Minors. Suppose that Laura had actually realized, when she sold Alan the drugs, that her actions broke the law – that is, that she was selling drugs to a minor. And suppose that she decided to sell him the drugs anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Dumping. Suppose that Alan had actually realized, when put the chemicals in his backyard, that his actions broke the law – that is, that he was placing prohibited chemicals where they could enter the river. And suppose that he decided to put the chemicals there anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Unregistered Firearm. Suppose that Laura had actually realized, before her friend pointed it out, that her actions broke the law – that is, that she should have registered the gun. And suppose that she decided to keep it unregistered anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Migratory Birds. Suppose that Alan had actually realized, when he shot the bird, that his actions broke the law – that is, that he was shooting a protected bird. And suppose that he decided to shoot the bird anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Mens Rea: Knowledge Manipulation

Theft. Suppose that Alan had actually realized, when he took the saw, that his actions broke the law – that is, that he was taking the property of another. And suppose that he decided to take it anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Battery. Suppose that Laura had actually realized, when she hit the boyfriend, that her actions broke the law – that is, that she was not defending herself or Tina. And suppose that she decided to hit him anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Burglary. Suppose that Laura had actually realized, when she took the vase, that her actions broke the law – that is, that she was entering the home with the intent to take the vase. And suppose that she decided to take it anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Sexual Battery. Suppose that Alan had actually realized, when he began touching Laura, that his actions broke the law – that is, that she did not consent to his touching. And suppose that he decided to touch her anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Seducing a Minor. Suppose that Alan had actually realized, when he sent the photographs Laura, that his actions broke the law – that is, that Laura was 15. And suppose that he decided to send the photographs anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

First Cousins. Suppose that Laura had actually realized, when she married Alan, that her actions broke the law – that is, that she was marrying her cousin. And suppose that she decided to marry him anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Reckless Driving. Suppose that Alan had actually realized, when he was driving in the school zone, that his actions broke the law – that is, that he was disregarding the safety of others. And suppose that he decided to drive that way anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Drug Distribution. Suppose that Laura had actually realized, when she delivered the package, that her actions broke the law – that is, that she was delivering ecstasy, a controlled substance. And suppose that she decided to delivery it anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Disturbing the Peace. Suppose that Laura had actually realized, when she was challenging the girl to fight, that her actions broke the law – that is, that she her challenge was disturbing the peace. And suppose that she decided to keep asking anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Public Drunkenness. Suppose that Alan had actually realized, when he threatened the bartender, that his actions broke the law – that is, that his threat constituted public drunkenness. And suppose that he decided to make his challenge anyway. In this case, where Alan knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Minor in Possession. Suppose that Laura had actually realized, when she took the beer, that her actions broke the law – that is, that she was violating the law even though she didn't plan to drink the beer. And suppose that she decided to take it anyway. In this case, where Laura knowingly violated the law, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Supplementary Material 1: Experiment 1 Full Stimuli
Conventional

Baseball Uniforms: Unknowing

Jack is a boy who likes to play games, especially baseball. Jack plays baseball for his school's team, the Blue Jays. Jack's teacher is the coach of the Blue Jays. There's a rule that all the students on the Blue Jays' team wear blue shirts with the school logo on the back to baseball practice. The school takes this rule and respect for the school and its logo very seriously.

One day, Jack was getting ready for a baseball practice. He was in a hurry to get to the bus on time, so he dressed quickly and left. Jack didn't realize he had grabbed the wrong blue shirt. So Jack went to the practice wearing a blue shirt that did not have the school logo on the back, in violation of the long-standing school policy.

Because Jack was wearing the wrong shirt, he was allowed to practice that day, but was sent to the principal's office after practice.

Baseball Uniforms: Knowing

Suppose that Jack had actually realized, while he was dressing, that the shirt he was about to put on for practice violated the rule – that is, that it didn't have the logo on the back. And suppose that he decided to wear it anyway. In this case, where Jack knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Baseball: No Rule

Finally, suppose that Jack's school had no rule prohibiting wearing a shirt without the school logo to practice, and Jack knowingly wore a shirt without the school logo to practice. In this case, with no rule about how to dress for practice in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Teacher's Title: Unknowing

Jerry's school has a new teacher named Taylor Madison. But the students at Jerry's school haven't learned the new teacher's name yet. At Jerry's school, there's a rule that children must call their teachers by their last names as a sign of respect. The rule has been in place for a long time, and it's almost never violated. All the teachers, though, always call the students by their first names.

One day, Jerry heard another teacher talking to Miss Madison. The other teacher called her "Taylor." Jerry thought that Taylor was her last name, and the next time he saw her, he said, "Hi, Miss Taylor."

Because students are not allowed to call teachers by their first names, and Jerry called Miss Madison by her first name, he was sent to the principal's office.

Teacher's Title: Knowing

Suppose that Jerry actually knew Miss Madison's full name and that his informal greeting violated the rule – that is, that he would be calling Miss Madison by her first name. And suppose he called her "Miss Taylor" anyway. In this case, where Jerry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Teacher's Title: No Rule

Finally, suppose that Jerry's school had no rule prohibiting calling a teacher by her first name, and that Jerry knowingly called a teacher by her first name. In this case, with no rule about how to address teachers, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Greetings: Unknowing

Sally is a student enrolled in her school's internship program. For her internship, she works as a secretary for a local district court judge. At the courthouse, there's a rule that judges should always be referred to by their full names and titles when speaking to lawyers or other judges. So when Sally sees people waiting for her boss Rebecca Rivers, she is supposed to say "the Honorable Rebecca Rivers will be with you shortly."

One day, Sally saw a woman waiting in the waiting area. The woman looked like the photo of her boss's daughter from her boss's desk, so Sally assumed that's who she was and said, "Becky will be with you shortly." In fact, the woman was a lawyer in one of the cases the judge was hearing.

Because Sally failed to use her boss's accurate title when speaking to a lawyer, her boss reported her to the school for violating the terms of the internship program, and she was called to the principal's office.

Greetings: Knowing

Suppose that Sally had actually realized that she was speaking to a lawyer and that her introduction violated the rule – that is, that she was supposed to be using the judge's proper title. And suppose that she decided to introduce the judge by her first name anyway. In this case, where Sally knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Greetings: No Rule

Finally, suppose that there was no rule against referring to the judge by her first name, and that Sally referred to the judge by her first name to someone she knew was a lawyer. In this case, with no rule about how to refer to the judges, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Dress Code - Unknowing

At Gregory's school, a dress code is strictly enforced: the rule is that students need to follow the dress code every day but "casual dress day," which occurs once a month and on which students are allowed to wear more casual clothes than usual. The dress code has been in place for

decades, and contributes to the unique character of the school. One day, Gregory puts on a new shirt and pants that he finds in his closet and heads to school.

Unbeknownst to Gregory, the new clothing he put on would only be acceptable on casual day. Gregory knew that it wasn't casual day, but he didn't know that what he was wearing would be considered casual.

Because it wasn't casual day that day at school, and Gregory was wearing casual clothes, he was sent to the principal's office.

Dress Code - Knowing

Suppose that Gregory had actually realized, while he was dressing, that the clothes he was about to put on for school violated the rule – that is, that they would be considered casual. And suppose that he decided to wear them anyway. In this case, where Gregory knowingly violated the rule, how you would respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Dress Code: No Rule

Finally, suppose that Gregory's school had no rule against dressing casually every day, and that he knowingly dressed in casual clothes. In this case, with no rule about dress code in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Lunch Table: Unknowing

At Georgia's school, students are assigned to lunch tables at the beginning of each week, and there's a rule that students must sit at their assigned tables. The teachers have found that this reduces the hurt feelings and fights that can arise when students are allowed to choose their own lunch tables. It also increases interaction between students who otherwise might not see much of each other. The policy has been in place for decades and reflects the school's commitment to social and emotional development alongside academics.

On Monday, Georgia didn't realize the lunch table assignments had changed, so she sat at a table other than the one to which she had been assigned. Georgia knew that she had to sit at her assigned table, but she did not realize she was at the wrong table.

Because Georgia was sitting at the wrong lunch table, she was sent to the principal's office.

Lunch Table: Knowing

Suppose that Georgia had actually known the correct table and that the table she chose to sit at violated the rule – that is, she wasn't sitting at her assigned lunch table. And suppose that she decided to sit there anyway. In this case, where Georgia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Lunch Table: No Rule

Finally, suppose that Georgia's school had no rule about where students have to sit for lunch, and that Georgia sat at a particular table. In this case, with no rule about where to sit for lunch, how

would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Tardy: Unknowing

Jane is taking chemistry, which is one of the largest classes at her school, taught in a large lecture hall. As there are so many students in the class, the teacher has established rules to try to make the class run as smoothly as possible. One rule is that students must arrive on time for class on test days so that the tests can be handed out promptly and in an orderly manner. Usually the school has a 5-minute grace period for arriving late before requiring students to obtain a tardy slip from their previous teacher, but on test days arriving even one minute after the bell rings is considered tardy.

On an exam day, Jane started walking to chemistry class, but stopped halfway there to talk to a friend. Jane mistakenly thought the grace period applied, and with the delay, Jane showed up for her test one minute late. However, with the teacher's rule in place, this was actually considered "tardy."

Because Jane was late for class on an exam day, she was sent to the principal's office after the test.

Tardy: Knowing

Suppose that Jane had actually realized, while she was talking to her friend, that there wasn't a grace period for being tardy on exam days. And suppose that she decided to stop and talk for the same amount of time anyway. In this case, where Jane knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Tardy: No Rule

Finally, suppose that Jane's school had no rule modifying the tardiness policy on exam days, and that she knowingly arrived to class one minute after the bell on an exam day. In this case, with no special rule about tardiness on test days, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Moral

The Dollar: Unknowing

One day, Ed helped his mother clean up the house. His mother gave him five dollars because he did such a good job. The next day Ed took the money with him to school so he could buy something for himself after school.

At lunchtime, Joe was bragging about the $5 he had found on the floor, and took out the bill to prove it. Not wanting to be outdone, Ed decided to show *his* $5 to Joe. Ed took it out of his pocket and put it down on the table. Just then, a teacher began to make an announcement to the lunchroom. Ed turned to listen and left the bill on the table. While Ed was turned, Joe picked up the bill and put it in his pocket, thinking it was his own. When Ed turned back, he couldn't find his dollar anywhere, and he accused Joe of taking it. At Ed and Joe's school, there is a rule against taking the property of others.

Because Joe was accused of taking Ed's money, he was sent to the principal's office.

The Dollar: Knowing

Suppose that Joe had actually realized, while he was picking up the money, that his action violated the rule – that is, that the five dollar bill was not his. And suppose that he decided to pick it up anyway. In this case, where Joe knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

The Dollar: No Rule

Finally, suppose that Joe's school had no rule prohibiting taking the property of others, and that Joe took the $5 bill knowing that it belonged to Ed. In this case, with no rule against taking another student's $5 bill, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Pushing: Unknowing

Sam was standing with a group of girls during recess. John walked up to the girls and started talking to them. Sam doesn't like John and wanted him to leave. Sam turned his back on John, trying to ignore him. When John still didn't leave, Sam swung his backpack onto his back, and walked away. Since his back was turned, Sam didn't realize how close John was or that he had swung his backpack right into John, hitting him and almost knocking him down. When John recovered his balance, he looked upset. At Sam and John's school, there is a rule against hitting or pushing other students.

Because Sam hit John with his backpack, he was sent to the principal's office.

Pushing: Knowing

Suppose that Sam had actually realized, while he swung his backpack onto his back, that his actions violated the rule – that is, that he was going to hit John with his backpack. And suppose that he decided to swing his backpack that way anyway. In this case, where Sam knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Pushing: No Rule

Finally, suppose that Sam's school had no rule against hitting or pushing other students, and that Sam swung his backpack, knowing that it would hit John. In this case, with no rule about how to treat other students, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

The Embezzler: Unknowing

Gloria was the bookkeeper for a club at her school. She was in charge of all the club's money, and Gloria was the only person who regularly managed the club's online accounts. The club had a rule against members mingling private funds with club funds.

One day, Gloria was managing her private accounts after school, and also had the club's accounts open. Gloria was switching back and forth between the two accounts and transferred

$5 of the club's money into the account that ended 8624. She thought that it was a club account, but it was actually one of her private accounts.

Because Gloria had transferred some of the club's funds into her private accounts, she was called to the principal's office.

The Embezzler: Knowing

Suppose that Gloria had actually realized, while she was working on the accounts, that her transfer violated the rule – that is, that she would be depositing club money into her account. And suppose that she decided to transfer it anyway. In this case, where Gloria knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

The Embezzler: No Rule

Finally, suppose that Gloria's school had no rule prohibiting the comingling of funds, and that Gloria knowingly deposited $5 of the club's money into her personal account. In this case, with no rule about how to manage the club's money, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Mascot: Unknowing

Gregory plays on his school's soccer team. His school's main rival is Bear Academy, which has a big furry bear as its mascot. Because the two schools have had confrontations in the past, there's a rule that students aren't allowed to hit each other at joint school events.

One day, before a game, Gregory saw the Bear Academy's mascot costume laying on a bench. He thought it would be funny to photograph himself hitting the mascot, so he had a friend snap a photo as he slapped the bear costume in the belly. Unbeknownst to Gregory, however, there was a student in the costume (the student had been taking a nap on the bench), so Gregory unknowingly hit the student.

Because Gregory hit another student at a school event, he was sent to the principal's office.

Mascot: Knowing

Suppose that Gregory had actually realized, while he was hitting the mascot, that his actions violated the rule – that is, that there was a student inside the suit whom he would be hitting. And suppose that he decided to slap the costume anyway. In this case, where Gregory knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Mascot: No Rule

Finally, suppose that Gregory's school had no rule prohibiting hitting other students, and that Gregory slapped the mascot costume knowing that a student was inside. In this case, with no rule about how to behave at school events, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Throwing: Unknowing

At Olivia's school, to avoid injuries, there's a rule that students are allowed to play ball games, but they aren't allowed to throw balls directly *at* each other. One day, Olivia throws a ball on the lawn during lunch, aiming for the center of a group of picnic tables. Unbeknownst to Olivia, another student is sitting there, hidden by the tables.

Olivia knew that she wasn't allowed to throw balls at other students, but she didn't know that there was another student exactly where she aimed the ball.

Because Olivia threw something, and it was aimed at another student, she was sent to the principal's office.

Throwing: Knowing

Suppose that Olivia had actually realized, while she was aiming, that there was a student near the picnic tables at which she was throwing, in the ball's path. And suppose that she decided to throw the ball anyway. In this case, where Olivia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Throwing: No Rule

Finally, suppose that Olivia's school had no rule prohibiting throwing a ball at another student, and that she knowingly threw the ball at the student near the picnic tables. In this case, with no rule about how to play with balls, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Bullying: Unknowing

Brittany's school has been having trouble with bullying and insensitive teasing. To try to improve this trend, the school has had informational assemblies, class discussions, and enacted some new rules. One of the new rules is that students may not read out loud the private papers or writing of another student.

Brittany, a junior, works as a classroom aid for the freshman English class. This means that Brittany helps collect and read assignments and suggests which assignments are good enough that they ought to be shared with the entire class. While reading the assignments one day, Brittany came across one that was especially well-written and moving, discussing a girl who had a stressful home life.

Brittany recommended to the teacher that it be shared with the class as an example of exceptional writing. Brittany read it to the class the next day. Neither Brittany nor the teacher realized that the paper was the private writing of one of the students, which Brittany had picked up by mistake – not the actual class assignment.

Because Brittany read out loud the private writing of another student, she was sent to the principal's office.

Bullying: Knowing

Suppose that Brittany had actually realized, when reading the student's writing, that it was private writing and not the assignment, such that reading it out loud would violate the rule. And suppose that she decided to read it out loud anyway. In this case, where Brittany knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Bullying: No Rule

Finally, suppose that Brittany's school had no rule prohibiting reading private writing out loud, and that Brittany read the students writing out loud knowing it to be private. In this case, with no rule about how to treat the writing of other students, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Supplementary Material 2: Replication of Knowledge Effects Within Subjects

Methods

     Participants. One-hundred-and-sixty adults (96 female, 64 male, mean age = 36, *SD* = 12) participated in the study through Amazon Mechanical Turk in exchange for monetary compensation. An additional 69 participants were tested but excluded for failing catch questions (55) or to ensure even numbers in all conditions (14). Participation was restricted to workers with IP addresses in the United States and a prior HIT approval rating of 95% or higher.

     Materials & Procedure. The experimental stimuli consisted of 12 distinct vignettes, 6 of which involved conventional violations and 6 of which involved moral violations. There were two versions of each story: one involved an agent who committed the violation knowingly, and the other involved an agent who knew the rule, but violated it unknowingly.

     Six of the vignettes (Teacher's Title, Greeting, Baseball, Dollar, Physician, and Embezzler) were based on vignettes originally presented to children by Davidson, Turiel, and Black (1983). These vignettes were modified to generate matched *knowing* and *unknowing* conditions. The Physician and Embezzler vignettes were additionally modified to take place in a school setting.

     Individual participants were randomly assigned to one of four conditions, the result of crossing violation domain (2: conventional, moral) with knowledge status (2: *knowing*, *unknowing*). Each participant received the corresponding six vignettes in a random order. Sample knowing and unknowing vignettes for one story, Baseball, are excerpted below (see Supplementary Material 3 for all vignettes). In this story, the rule was that students had to wear a blue shirt with the school logo on the back to baseball practice:

> *Knowing*: "One day, Jack was getting ready for a baseball practice. He was tired of always wearing his blue practice shirt; he thought it would be fun to wear another shirt for a change. So Jack went to the practice wearing a blue shirt that did not have the school logo on the back."
>
> *Unknowing:* "One day, Jack was getting ready for a baseball practice. He was in a hurry to get to the bus on time, so he dressed quickly and left. Jack didn't realize he had grabbed the wrong blue shirt. So Jack went to the practice wearing a blue shirt that did not have the school logo on the back."

In both cases, they then read: "Because Jack was wearing the wrong shirt, he was not allowed to practice that day and was sent home." The presentation of each vignette was then followed by moral censure and detention questions, presented on one screen in random order:

> *Censure*: "How wrong was [Actor's actions]?" Participants indicated their answer on a scale from 0 (not at all wrong) to 6 (very wrong).
>
> *Detention*: "How many hours of detention should [Actor] get?" Participants indicated their answer on a scale from 0 to 6 hours.

     After answering these two questions, participants were presented with another screen and asked to indicate the moral censure and detention ratings the actor would deserve if the school had never had a rule prohibiting the action. The wording of these questions (again presented in

random order) was identical to that above, but preceded by the following: "What if [Actor's] school had no rule prohibiting what [s]he did? Please answer the following questions based on this rule change."

Next, on a separate screen, participants were presented with an easy true/false comprehension question relating to the vignette they had just read. These questions were used to assess whether participants had read the vignette carefully; those who answered any comprehension questions incorrectly were excluded from further analyses.

After reading all six vignettes and answering their associated questions, participants answered one additional catch question designed to ensure that they were reading instructions carefully, modeled after Oppenheimer, Meyvis, and Davidenko (2009). Finally, participants answered demographic questions about their age and gender.

Results

Moral censure and detention with a rule. To test the prediction that judgments regarding conventional violations are less "knowledge dependent" than those regarding moral violations, we performed a 2 (knowledge status: *knowing*, *unknowing*) x 2 (violation domain: conventional, moral) between-subjects ANOVA on initial moral censure ratings, and another on initial detention ratings (see Figure S2). We expected to find an interaction between knowledge status and violation domain, with a larger effect of knowledge status for moral violations than for conventional violations.

This prediction was confirmed: we found significant interactions between knowledge status and domain for both censure, $F(1,156) = 19.02$, $p<.000$, $\eta_p^2 =.11$, and detention, $F(1,156) = 24.00$, $p<.000$, $\eta_p^2 =.13$. Independent samples t-tests showed that moral censure ratings were significantly higher for the *knowing* condition than for the *unknowing* condition for both conventional violations, $t(78) = 2.6$, $p<.012$, $r = .28$, and for moral violations, $t(78) = 8.62$, $p<.001$, $r = .70$, but the difference was greater for moral violations (2.28 points versus 1.34 points). Detention ratings showed the same pattern: responses were higher for the *knowing* condition than for the *unknowing* condition for both conventional violations, $t(78) = 1.99$, $p<.050$, $r = .22$, and moral violations, $t(78) = 6.06$, $p<.001$, $r = .57$, but the difference was greater for moral violations (1.67 points versus 0.22 points).

These analyses also revealed two significant pairs of main effects. First, there were main effects of violation domain: moral violations received significantly higher censure, $F(1,156) = 87.72$, $p<.000$, $\eta_p^2 =.36$, and detention ratings, $F(1,156) = 169.25$, $p<.000$, $\eta_p^2 =.52$. Second, there were main effects of knowledge status, with censure, $F(1,156) = 63.13$, $p<.000$, $\eta_p^2 =.29$, and detention, $F(1,156) = 40.5$, $p<.000$, $\eta_p^2 =.21$, ratings significantly higher for the *knowing* condition than the *unknowing* condition.
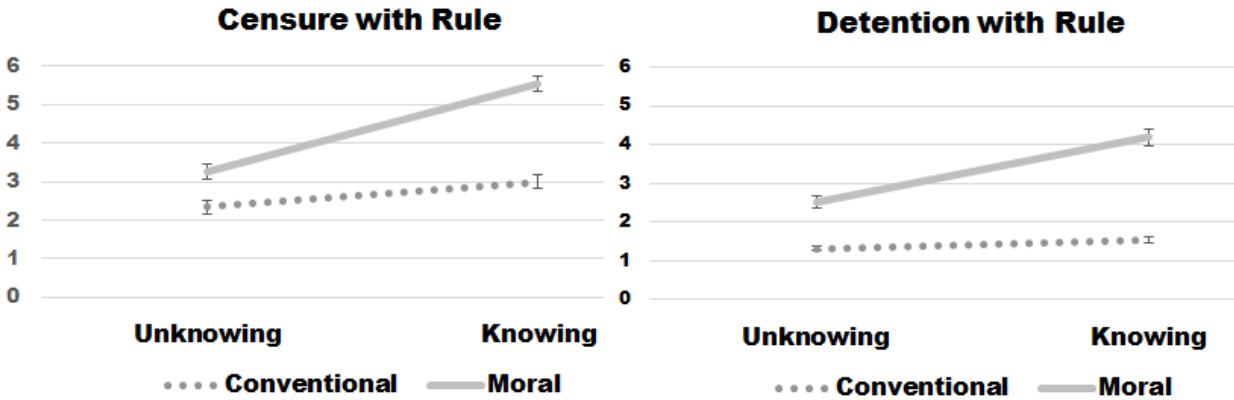
Figure S2: Interaction between knowledge condition and violation type for moral censure and detention ratings. Error bars correspond to one SEM in each direction.

Moral censure and detention without a rule. To what extent are conventional and moral violations condemned simply because they violate a rule – that is, to what extent are judgments of moral censure and punishment "rule dependent" in each domain? To address these questions, moral censure and detention difference scores were created by subtracting participants' moral censure and detention ratings after the rule change from their corresponding initial scores for each vignette. An average of these six difference scores from each participant for both moral censure and detention were then created. We performed 2 (knowledge status: *knowing*, *unknowing*) x 2 (violation domain: conventional, moral) ANOVAs on each average of difference scores, and predicted higher scores (i.e., a larger effect of the rule change) for conventional violations relative to moral violations.

This prediction was confirmed for censure: we found a significant main effect of violation domain, $F(1,156) = 33.76$, $p<.000$, $\eta_p^2 =.18$, with a larger drop in the perceived wrongness of conventional violations following the rule change relative to moral violations (1.11 points versus 0.41 points; see Table S2). The ratings for detention, however, were unexpected: we found a significant main effect of domain, but in the opposite direction, $F(1,156) = 5.37$, $p<.022$, $\eta_p^2 =.03$. This may be in part because conventional violations were assigned very low levels of detention initially; the response distribution was skewed, with little room for a drop in ratings after the rule change.

We also found a significant main effect of knowledge for moral censure ratings, with a larger drop in perceived wrongness for the *knowing* condition (M=.91, .STD=94) relative to the *unknowing* condition (M=.61, STD=.72), $F(1,156) = 6.45$, $p<.012$, $\eta_p^2 =.04$. There were no additional significant effects.[33]

---

[33] We ran further analyses to investigate the relationship between knowledge dependence and rule dependence. To obtain a measure of knowledge dependence for each vignette, the average wrongness rating from participants in the *unknowing* condition for that vignette was subtracted from the average wrongness rating from participants in the *knowing* condition for that vignette. To obtain a measure of rule dependence for each vignette, we averaged the wrongness difference scores (initial wrongness ratings minus wrongness ratings after rule change) for the knowing and unknowing versions of that vignette. This resulted in twelve pairs of numbers – one pair for each story – that revealed a significant negative correlation, r = -.74, p < .006: those vignettes for which wrongness judgments were more knowledge dependent tended to involve judgments that were less rule dependent. However, this correlation was not uniform over violation domain. Considering just conventional violations, the correlations between rule dependence and knowledge dependence was positive, though not significantly, r = .34, p = .52. For moral violations, the correlation remained negative and significant, r = -.86, p < .027. Because this analysis was performed across

			Moral	**Conventional**
Censure	*Knowing*	Rule in place	5.54 (SD =1.16)	3.00 (SD =1.21)
		No Rule	5.08 (SD =1.11)	1.63 (SD =0.73)
		Difference	*0.46*	*1.37*
	Unknowing	Rule in place	3.26 (SD =1.20)	2.34 (SD =1.11)
		No Rule	2.90 (SD =1.15)	1.48 (SD =0.59)
		Difference	*0.36*	*0.86*
Detention	*Knowing*	Rule in place	4.19 (SD =1.37)	1.53 (SD =0.52)
		No Rule	3.54 (SD =1.43)	1.10 (SD =0.19)
		Difference	*0.65*	*0.43*
	Unknowing	Rule in place	2.52 (SD =1.07)	1.32 (SD =0.45)
		No Rule	2.03 (SD =0.98)	1.09 (SD =0.22)
		Difference	*0.49*	*0.23*

Table S2: Average moral censure and detention ratings when a rule was and was not in place. The corresponding difference scores reflect the extent to which judgments were "rule dependent."

Discussion

Experiment S2 found a novel relationship between the kind of rule violated (moral versus conventional) and the extent to which knowledge influence judgments of wrongdoing and punishment. Specifically, we found that participants judged actors more harshly (in terms of both the wrongness of their action and hours of detention deserved) when they violated a rule knowingly as opposed to unknowingly. However, the effect of acting knowingly (versus unknowingly) was significantly greater for moral violations relative to conventional violations.

Consistent with prior work by Turiel and colleagues, Experiment S2 also found that when the operating rule was removed, judgments of moral censure were reduced to a greater extent for conventional violations than for moral violations. Unexpectedly, the same pattern did not emerge for judgments concerning punishment (i.e., hours of detention). This may be because detention ratings were generally low, especially for conventional violations; participants may not have felt that detention was an appropriate punishment given the minimal severity of the transgressions, and there was little room for downward movement in ratings after the rule change.

One concern about our results is that moral violations differed from conventional violations not only in their greater "knowledge dependence," as predicted, but also in overall severity. On average, the moral violations were judged more harshly than the conventional violations. While we might expect this to be the case when a violation was committed knowingly (since the intentions associated with violating a moral rule will tend to be worse than those associated with violating a conventional rule), it would be reassuring to find the same pattern of results when the unknowing versions of transgressions are comparable. This would help eliminate an alternative explanation for our results: that the lower knowledge-dependence observed for conventional violations stemmed not from the domain of the violation, but from the lower severity of the transgression. In fact, the observed pattern remained even when we

vignettes – and therefore limited by a sample size of 12, with only 6 per domain – the results are merely suggestive. These correlations make sense in light of the relationships suggested by Table 1.

considered only the three most harshly rated conventional violations and the three least harshly rated moral violations, which did not differ significantly in severity. Moreover, we observed a greater drop in moral censure ratings after the rule change for conventional violations than for moral violations, which indicates that these ratings were not displaying a floor effect. Nonetheless, in Experiment 1 reported in the main body, we used modified versions of these vignettes to better equate the perceived wrongness of violations across domains.

Supplementary Material 3: S2 Stimuli

Conventional

Baseball Uniforms: Unknowing

Jack is a boy who likes to play games, especially baseball. Jack plays baseball for his school's team, the Blue Jays. Jack's teacher is the coach of the Blue Jays. There's a rule that all the students on the Blue Jays' team wear blue shirts with the school logo on the back to baseball practice. The school takes this rule and respect for the school and its logo very seriously.

One day, Jack was getting ready for a baseball practice. He was in a hurry to get to the bus on time, so he dressed quickly and left. Jack didn't realize he had grabbed the wrong blue shirt. So Jack went to the practice wearing a blue shirt that did not have the school logo on the back, in violation of the long-standing school policy.

Because Jack was wearing the wrong shirt, he was allowed to practice that day, but was sent to the principal's office after practice.

Baseball Uniforms: Knowing

Suppose that Jack had actually realized, while he was dressing, that the shirt he was about to put on for practice violated the rule – that is, that it didn't have the logo on the back. And suppose that he decided to wear it anyway. In this case, where Jack knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Baseball: No Rule

Finally, suppose that Jack's school had no rule prohibiting wearing a shirt without the school logo to practice, and Jack knowingly wore a shirt without the school logo to practice. In this case, with no rule about how to dress for practice in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Teacher's Title: Unknowing

Jerry's school has a new teacher named Taylor Madison. But the students at Jerry's school haven't learned the new teacher's name yet. At Jerry's school, there's a rule that children must call their teachers by their last names as a sign of respect. The rule has been in place for a long time, and it's almost never violated. All the teachers, though, always call the students by their first names.

One day, Jerry heard another teacher talking to Miss Madison. The other teacher called her "Taylor." Jerry thought that Taylor was her last name, and the next time he saw her, he said, "Hi, Miss Taylor."

Because students are not allowed to call teachers by their first names, and Jerry called Miss Madison by her first name, he was sent to the principal's office.

Teacher's Title: Knowing

Suppose that Jerry actually knew Miss Madison's full name and that his informal greeting violated the rule – that is, that he would be calling Miss Madison by her first name. And suppose he called her "Miss Taylor" anyway. In this case, where Jerry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Teacher's Title: No Rule

Finally, suppose that Jerry's school had no rule prohibiting calling a teacher by her first name, and that Jerry knowingly called a teacher by her first name. In this case, with no rule about how to address teachers, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Greetings: Unknowing

Sally is a student enrolled in her school's internship program. For her internship, she works as a secretary for a local district court judge. At the courthouse, there's a rule that judges should always be referred to by their full names and titles when speaking to lawyers or other judges. So when Sally sees people waiting for her boss Rebecca Rivers, she is supposed to say "the Honorable Rebecca Rivers will be with you shortly."

One day, Sally saw a woman waiting in the waiting area. The woman looked like the photo of her boss's daughter from her boss's desk, so Sally assumed that's who she was and said, "Becky will be with you shortly." In fact, the woman was a lawyer in one of the cases the judge was hearing.

Because Sally failed to use her boss's accurate title when speaking to a lawyer, her boss reported her to the school for violating the terms of the internship program, and she was called to the principal's office.

Greetings: Knowing

Suppose that Sally had actually realized that she was speaking to a lawyer and that her introduction violated the rule – that is, that she was supposed to be using the judge's proper title. And suppose that she decided to introduce the judge by her first name anyway. In this case, where Sally knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Greetings: No Rule

Finally, suppose that there was no rule against referring to the judge by her first name, and that Sally referred to the judge by her first name to someone she knew was a lawyer. In this case, with no rule about how to refer to the judges, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Dress Code - Unknowing

At Gregory's school, a dress code is strictly enforced: the rule is that students need to follow the dress code every day but "casual dress day," which occurs once a month and on which students are allowed to wear more casual clothes than usual. The dress code has been in place for decades, and contributes to the unique character of the school. One day, Gregory puts on a new shirt and pants that he finds in his closet and heads to school.

Unbeknownst to Gregory, the new clothing he put on would only be acceptable on casual day. Gregory knew that it wasn't casual day, but he didn't know that what he was wearing would be considered casual.

Because it wasn't casual day that day at school, and Gregory was wearing casual clothes, he was sent to the principal's office.

Dress Code - Knowing

Suppose that Gregory had actually realized, while he was dressing, that the clothes he was about to put on for school violated the rule – that is, that they would be considered casual. And suppose that he decided to wear them anyway. In this case, where Gregory knowingly violated the rule, how you would respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Dress Code: No Rule

Finally, suppose that Gregory's school had no rule against dressing casually every day, and that he knowingly dressed in casual clothes. In this case, with no rule about dress code in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Lunch Table: Unknowing

At Georgia's school, students are assigned to lunch tables at the beginning of each week, and there's a rule that students must sit at their assigned tables. The teachers have found that this reduces the hurt feelings and fights that can arise when students are allowed to choose their own lunch tables. It also increases interaction between students who otherwise might not see much of each other. The policy has been in place for decades and reflects the school's commitment to social and emotional development alongside academics.

On Monday, Georgia didn't realize the lunch table assignments had changed, so she sat at a table other than the one to which she had been assigned. Georgia knew that she had to sit at her assigned table, but she did not realize she was at the wrong table.

Because Georgia was sitting at the wrong lunch table, she was sent to the principal's office.

Lunch Table: Knowing

Suppose that Georgia had actually known the correct table and that the table she chose to sit at violated the rule – that is, she wasn't sitting at her assigned lunch table. And suppose that she decided to sit there anyway. In this case, where Georgia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Lunch Table: No Rule

Finally, suppose that Georgia's school had no rule about where students have to sit for lunch, and that Georgia sat at a particular table. In this case, with no rule about where to sit for lunch, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Tardy: Unknowing

Jane is taking chemistry, which is one of the largest classes at her school, taught in a large lecture hall. As there are so many students in the class, the teacher has established rules to try to make the class run as smoothly as possible. One rule is that students must arrive on time for class on test days so that the tests can be handed out promptly and in an orderly manner. Usually the school has a 5-minute grace period for arriving late before requiring students to obtain a tardy slip from their previous teacher, but on test days arriving even one minute after the bell rings is considered tardy.

On an exam day, Jane started walking to chemistry class, but stopped halfway there to talk to a friend. Jane mistakenly thought the grace period applied, and with the delay, Jane showed up for her test one minute late. However, with the teacher's rule in place, this was actually considered "tardy."

Because Jane was late for class on an exam day, she was sent to the principal's office after the test.

Tardy: Knowing

Suppose that Jane had actually realized, while she was talking to her friend, that there wasn't a grace period for being tardy on exam days. And suppose that she decided to stop and talk for the same amount of time anyway. In this case, where Jane knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Tardy: No Rule

Finally, suppose that Jane's school had no rule modifying the tardiness policy on exam days, and that she knowingly arrived to class one minute after the bell on an exam day. In this case, with no special rule about tardiness on test days, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Moral

The Dollar: Unknowing

One day, Ed helped his mother clean up the house. His mother gave him five dollars because he did such a good job. The next day Ed took the money with him to school so he could buy something for himself after school.

At lunchtime, Joe was bragging about the $5 he had found on the floor, and took out the bill to prove it. Not wanting to be outdone, Ed decided to show *his* $5 to Joe. Ed took it out of his

pocket and put it down on the table. Just then, a teacher began to make an announcement to the lunchroom. Ed turned to listen and left the bill on the table. While Ed was turned, Joe picked up the bill and put it in his pocket, thinking it was his own. When Ed turned back, he couldn't find his dollar anywhere, and he accused Joe of taking it. At Ed and Joe's school, there is a rule against taking the property of others.

Because Joe was accused of taking Ed's money, he was sent to the principal's office.

The Dollar: Knowing

Suppose that Joe had actually realized, while he was picking up the money, that his action violated the rule – that is, that the five dollar bill was not his. And suppose that he decided to pick it up anyway. In this case, where Joe knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

The Dollar: No Rule

Finally, suppose that Joe's school had no rule prohibiting taking the property of others, and that Joe took the $5 bill knowing that it belonged to Ed. In this case, with no rule against taking another student's $5 bill, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Pushing: Unknowing

Sam was standing with a group of girls during recess. John walked up to the girls and started talking to them. Sam doesn't like John and wanted him to leave. Sam turned his back on John, trying to ignore him. When John still didn't leave, Sam swung his backpack onto his back, and walked away. Since his back was turned, Sam didn't realize how close John was or that he had swung his backpack right into John, hitting him and almost knocking him down. When John recovered his balance, he looked upset. At Sam and John's school, there is a rule against hitting or pushing other students.

Because Sam hit John with his backpack, he was sent to the principal's office.

Pushing: Knowing

Suppose that Sam had actually realized, while he swung his backpack onto his back, that his actions violated the rule – that is, that he was going to hit John with his backpack. And suppose that he decided to swing his backpack that way anyway. In this case, where Sam knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Pushing: No Rule

Finally, suppose that Sam's school had no rule against hitting or pushing other students, and that Sam swung his backpack, knowing that it would hit John. In this case, with no rule about how to treat other students, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

The Embezzler: Unknowing

Gloria was the bookkeeper for a club at her school. She was in charge of all the club's money, and Gloria was the only person who regularly managed the club's online accounts. The club had a rule against members mingling private funds with club funds.

One day, Gloria was managing her private accounts after school, and also had the club's accounts open. Gloria was switching back and forth between the two accounts and transferred $5 of the club's money into the account that ended 8624. She thought that it was a club account, but it was actually one of her private accounts.

Because Gloria had transferred some of the club's funds into her private accounts, she was called to the principal's office.

The Embezzler: Knowing

Suppose that Gloria had actually realized, while she was working on the accounts, that her transfer violated the rule – that is, that she would be depositing club money into her account. And suppose that she decided to transfer it anyway. In this case, where Gloria knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

The Embezzler: No Rule

Finally, suppose that Gloria's school had no rule prohibiting the comingling of funds, and that Gloria knowingly deposited $5 of the club's money into her personal account. In this case, with no rule about how to manage the club's money, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Mascot: Unknowing

Gregory plays on his school's soccer team. His school's main rival is Bear Academy, which has a big furry bear as its mascot. Because the two schools have had confrontations in the past, there's a rule that students aren't allowed to hit each other at joint school events.

One day, before a game, Gregory saw the Bear Academy's mascot costume laying on a bench. He thought it would be funny to photograph himself hitting the mascot, so he had a friend snap a photo as he slapped the bear costume in the belly. Unbeknownst to Gregory, however, there was a student in the costume (the student had been taking a nap on the bench), so Gregory unknowingly hit the student.

Because Gregory hit another student at a school event, he was sent to the principal's office.

Mascot: Knowing

Suppose that Gregory had actually realized, while he was hitting the mascot, that his actions violated the rule – that is, that there was a student inside the suit whom he would be hitting. And suppose that he decided to slap the costume anyway. In this case, where Gregory knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Mascot: No Rule

Finally, suppose that Gregory's school had no rule prohibiting hitting other students, and that Gregory slapped the mascot costume knowing that a student was inside. In this case, with no rule about how to behave at school events, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Throwing: Unknowing

At Olivia's school, to avoid injuries, there's a rule that students are allowed to play ball games, but they aren't allowed to throw balls directly *at* each other. One day, Olivia throws a ball on the lawn during lunch, aiming for the center of a group of picnic tables. Unbeknownst to Olivia, another student is sitting there, hidden by the tables.

Olivia knew that she wasn't allowed to throw balls at other students, but she didn't know that there was another student exactly where she aimed the ball.

Because Olivia threw something, and it was aimed at another student, she was sent to the principal's office.

Throwing: Knowing

Suppose that Olivia had actually realized, while she was aiming, that there was a student near the picnic tables at which she was throwing, in the ball's path. And suppose that she decided to throw the ball anyway. In this case, where Olivia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Throwing: No Rule

Finally, suppose that Olivia's school had no rule prohibiting throwing a ball at another student, and that she knowingly threw the ball at the student near the picnic tables. In this case, with no rule about how to play with balls, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Bullying: Unknowing

Brittany's school has been having trouble with bullying and insensitive teasing. To try to improve this trend, the school has had informational assemblies, class discussions, and enacted some new rules. One of the new rules is that students may not read out loud the private papers or writing of another student.

Brittany, a junior, works as a classroom aid for the freshman English class. This means that Brittany helps collect and read assignments and suggests which assignments are good enough that they ought to be shared with the entire class. While reading the assignments one day, Brittany came across one that was especially well-written and moving, discussing a girl who had a stressful home life.

Brittany recommended to the teacher that it be shared with the class as an example of exceptional writing. Brittany read it to the class the next day. Neither Brittany nor the teacher realized that the paper was the private writing of one of the students, which Brittany had picked up by mistake – not the actual class assignment.

Because Brittany read out loud the private writing of another student, she was sent to the principal's office.

Bullying: Knowing

Suppose that Brittany had actually realized, when reading the student's writing, that it was private writing and not the assignment, such that reading it out loud would violate the rule. And suppose that she decided to read it out loud anyway. In this case, where Brittany knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Bullying: No Rule

Finally, suppose that Brittany's school had no rule prohibiting reading private writing out loud, and that Brittany read the students writing out loud knowing it to be private. In this case, with no rule about how to treat the writing of other students, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Supplementary Material 4: Experiment 2 Stimuli

Conventional

Dress Code – Unknowing

At Gregory's school, a dress code is strictly enforced: the rule is that students need to follow the dress code every day but "casual dress day," which occurs once a month and on which students are allowed to wear more casual clothes than usual. The dress code has been in place for decades, and contributes to the unique character of the school. One day, Gregory puts on a new shirt and pants that he finds in his closet and heads to school.

Unbeknownst to Gregory, the new clothing he put on would only be acceptable on casual day. Gregory knew that it wasn't casual day, but he didn't know that what he was wearing would be considered casual. Other students noticed he was wearing casual clothes.

Because it wasn't casual day that day at school, and Gregory was wearing casual clothes, he was sent to the principal's office.

Knowledge Stipulation: Suppose that Gregory had actually realized, while he was dressing, that the clothes he was about to put on for school violated the rule – that is, that they would be considered casual. And suppose that he decided to wear them anyway. In this case, where Gregory knowingly violated the rule, how you would respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Lunch Table – Unknowing

At Georgia's school, students are assigned to lunch tables at the beginning of each week, and there's a rule that students must sit at their assigned tables. The teachers have found that this reduces fights that can arise when students are allowed to choose their own lunch tables. It also helps teachers to know exactly where students will be during lunch hour, in case they need the student for any reason. The policy has been in place for decades and reflects the school's commitment to social and emotional development alongside academics.

On Monday, Georgia didn't realize the lunch table assignments had changed, so she sat at a table other than the one to which she had been assigned. Georgia knew that she had to sit at her assigned table, but she did not realize she was at the wrong table. Other students did notice that Georgia was at the wrong table.

Because Georgia was sitting at the wrong lunch table, she was sent to the principal's office.

Knowledge Stipulation: Suppose that Georgia had actually realized when she sat down, that her actions violated the rule – that is, she wasn't sitting at her assigned lunch table. And suppose that she decided to sit there anyway. In this case, where Georgia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Tardy – Unknowing

Jane is taking chemistry, which is one of the largest classes at her school, taught in a large lecture hall. As there are so many students in the class, the teacher has established rules to try to make the class run as smoothly as possible. One rule is that everyone must be ten minutes early for class on test days so that the tests can be handed out in an orderly manner.

On the day of her first exam, Jane started walking to chemistry class, but stopped half-way there to talk to a friend. Jane mistakenly believed her test was the next day, so she didn't realize the ten minute rule applied. Jane showed up for her test five minutes late which caused a slight disturbance as she got her test and got settled.

Because Jane was late for her exam, she disrupted the test for the other students and was sent to the principal's office.

Knowledge Stipulation: Suppose that Jane had actually known it was test day, and that her actions violated the rule –that is, that she was not going to be 10 minutes early. And suppose that she decided to talk to her friend anyway. In this case, where Jane knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Hall Monitor – Unknowing

Kyle attends a school that believes all students should be involved in keeping the school in good order, and there's a rule that older students must serve as hall monitors. Hall monitors spend their free period walking the halls, making sure everyone is in class, giving instructions to any visitors who may be lost, and generally making sure everything is going as it should. All upper level students are assigned to serve as hall monitors for a few weeks every semester.

Kyle knew that as an older student, he was required to serve as a hall monitor for a few weeks a semester, but one day he mixed up his schedule and did not show up to perform his hall monitor shift. That day, a pipe burst in a bathroom in the area Kyle was supposed to be monitoring. Because the pipe burst during class period, no one noticed for that entire period, at which time the water had spread and the leak had gotten larger. Cleaning up and replacing the things that were damaged will be time consuming and cost money. Had Kyle been monitoring, he would've noticed the leak sooner, reducing the damage.

Because Kyle didn't perform his hall monitor shift, he was sent to the principal's office.

Knowledge Stipulation: Suppose that Kyle had actually known it was his day to serve hall monitor duty, and that his actions violated the rule –that is, that he did not report for his shift. And suppose that he decided not to show up anyway. In this case, where Kyle knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Moral

Throwing – Unknowing

At Olivia school, there's a rule that students are allowed to play ball games, but they aren't allowed to throw balls directly at each other. One day, Olivia was throwing a soft, foam ball on the lawn during lunch, aiming for the center of a group of picnic tables.

Olivia knew she wasn't allowed to throw at another student, but she didn't realize another student was sitting in the center of the picnic tables. The ball hit the other student, but the soft, foam didn't hurt much.

Because Olivia threw something on the playground, and it was aimed at another student, she was sent to the principal's office

Knowledge Stipulation: Suppose that Olivia had actually realized, while she threw the ball, that her actions violated the rule – that is, that there was a student right where she was aiming her throw. And suppose that she decided to throw the ball anyway. In this case, where Olivia knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Candy – Unknowing

At Jeremy's school, there's a special event each year for Halloween. Students wear their costumes to school, and they go from classroom to classroom "trick or treating," with each teacher providing a different kind of treat. The event is informal and boisterous, but there's a rule that no student may take candy that belongs to another student.

After "trick or treating" one year, Jeremy and his friends sat at a lunch table to count out their candies. When they were putting their candies away again, Jeremy noticed a mini-Snickers bar that had fallen on the ground. After counting his candies again he was pretty sure it was one of his. He knew that he wasn't supposed to take candy that belonged to another student, but he thought it was his. However, another student noticed that his mini-Snickers was missing, and the teacher counted everyone's candy and discovered that Jeremy had taken it.

Because Jeremy took a piece of candy that belonged to another student, he was sent to the principal's office.

Knowledge Stipulation: Suppose that Jeremy had actually realized, while he picked up the candy, that his actions violated the rule – that is, that the candy did not belong to him. And suppose that he decided to keep the candy anyway. In this case, where Jeremy knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Burner – Unknowing

Melissa attends a school with an after-school science lab in which students can perform experiments. Some of the experiments involve Bunsen burners, which produce open gas flames. Because the burners are insulated, they cannot actually burn anyone. However, it is surprising and unpleasant to grab hold of one before it has cooled. As a result, there's a rule that students

have to place a sign that says "Beware, may be hot!" next to any Bunsen burner that they are using until it is cool to the touch.

One day Melissa and her lab partner begin a new experiment in the lab that requires a Bunsen burner. Melissa is in charge of clearing up after the experiment, and she puts away the sign in front of the Bunsen burner, because she mistakenly thought the burner was cool enough. Later, another science lab student picks up the burner and notices that it's still warm.

Because Melissa failed to leave the beware sign until the burner was cool to the touch, she was sent to the principal's office.

<u>Knowledge Stipulation</u>: Suppose that Melissa had actually realized, when she put the warning sign away, that her actions violated the rule – that is, the burner was not completely cool. And suppose that she decided to put the warning sign away anyway. In this case, where Melissa knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Pushing – Unknowing

At Sam's school, there is a rule against hitting or pushing other students. Sam was standing with a group of girls during recess. John walked up to the girls and started talking to them. Sam doesn't like John and wanted him to leave. Sam turned his back on John, trying to ignore him.

When John still didn't leave, Sam swung his backpack onto his back, and walked away. As he did this, Sam hit John with his backpack. John was knocked off-balance, and almost fell down. When John recovered his balance, he looked upset.

Sam knew that it was against the rules to hit another student, but he didn't realize that he would hit John with his backpack when he turned to leave.

Because Sam hit John with his backpack, he was sent to the principal's office.

<u>Knowledge Stipulation</u>: Suppose that Sam had actually realized, when he swung his backpack that his actions violated the rule – that is, that he would hit John. And suppose that he decided to swing his bag into John anyway. In this case, where Sam knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Supplementary Material 5: Experiment 3 Stimuli

MOVIES: Mens Rea + Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students of different ages, including a class on horror films, which is known for showing very gory movies. Because the content of the films can be harmful for younger aliens – leading to intense fear and recurring, violent nightmares – there's a rule that students are not allowed to share videos from the class with other students who are younger than 15-years old (in earth years). The school administrators chose the age of 15 somewhat arbitrarily. There isn't any reason why seeing the gory content would harm a 14.5-year old alien much more than a 15.5-year old alien, but they had to choose some age cutoff, and decided on 15.

One day, a student asked Bernice to borrow one of the videos from the class. The student was 14. Bernice knew that she was not allowed to lend the movie to students younger than 15, but Bernice had been in a class with the other student, so mistakenly believed she was also 15. Since Bernice believed the other student was old enough, she lent her the movie.

The younger student's parents found the video in their daughter's backpack, and the daughter told them that Bernice had given it to her. The parents immediately called the school to complain. Because Bernice lent a horror film to a student younger than 15, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, when she gave the video to the other student, that her actions violated the rule – that is, that the student was younger than 15. And suppose that she decided to give the girl the video anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that due to a clerical error, the rule against giving videos to aliens younger than 15 had never been put in the school charter and therefore could not be enforced. That is, suppose that Bernice's school EFFECTIVELY had no rule prohibiting giving the videos to students younger than 15, and Bernice gave the video to the younger student. In this case, with no rule about lending videos in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

MOVIES: Mens Rea + Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students of different ages, including a class on horror films, which is known for showing very gory movies. Because the content of the films can be harmful for younger aliens – leading to intense fear and recurring, violent nightmares – there's a

rule that students are not allowed to share videos from the class with other students who are younger than 15-years old (in earth years). The school administrators chose the age of 15 after careful research on alien development. Aliens of their species undergo an important developmental change such that younger aliens are susceptible to the negative effects of gory and violent content, while older aliens are not. The developmental change is linked to an increase in hormone levels that occurs, like clockwork, at age 15.

One day, a student asked Bernice to borrow one of the videos from the class. The student was 14. Bernice knew that she was not allowed to lend the movie to students younger than 15, but Bernice had been in a class with the other student, so mistakenly believed she was also 15. Since Bernice believed the other student was old enough, she lent her the movie.

The younger student's parents found the video in their daughter's backpack, and the daughter told them that Bernice had given it to her. The parents immediately called the school to complain. Because Bernice lent a horror film to a student younger than 15, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, when she gave the video to the other student that her actions violated the rule – that is, that the student was younger than 15. And suppose that she decided to give the girl the video anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting giving the videos to students younger than 15, and Bernice gave the video to the younger student. In this case, with no rule about lending videos in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

HUMMING: Mens Rea + Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. All aliens of Bernice's species are capable of making loud humming noises of different kinds. As a normal part of their development, young aliens enjoy trying out combinations of these noises to make new and unique hums. Because this humming can be distracting for other students – leading to less attention directed towards lessons, impaired learning, and even headaches – there's a rule that students are not allowed to make humming noises within 15 feet (in earth measurements) of any room currently being used for class. The school administration chose the distance of 15 feet somewhat arbitrarily. There isn't any reason to suspect that the noise would be much more disruptive at 14.5 feet than at 15.5 feet, but they had to choose some distance cutoff, and decided on 15 feet.

One morning, Bernice was at her locker practicing a new combination of humming noises she had been working on. A classroom full of students was 14 feet away. Bernice knew that she was

not allowed to hum within 15 feet of an occupied classroom, but that classroom typically wasn't used in the mornings, so Bernice thought it was empty. Since Bernice believed the classroom was empty, she thought it would be okay to practice her humming.

The class's teacher heard Bernice's humming and reported her to the office. Because Bernice was humming within 15 feet of an occupied classroom, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, before she began humming, that her actions would violate the rule – that is, that the classroom was in fact occupied. And suppose that she decided to start humming anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting humming within 15 feet of an occupied classroom, and Bernice practiced her humming at her locker. In this case, with no rule about humming in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

HUMMING: Mens Rea + Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. All aliens of Bernice's species are capable of making loud humming noises of different kinds. As a normal part of their development, young aliens enjoy trying out combinations of these noises to make new and unique hums. Because this humming can be distracting for other students – leading to less attention directed towards lessons, impaired learning, and even headaches – there's a rule that students are not allowed to make humming noises within 15 feet (in earth measurements) of any room currently being used for class. The school administration chose the distance of 15 feet after careful research on the unique auditory capabilities of their species. Aliens of their species can hear the noises clearly if they are made within exactly 15 feet. Beyond that distance, however, the noises fall below the critical threshold for the alien's auditory system, so they are barely audible and no longer distracting.

One morning, Bernice was at her locker practicing a new combination of humming noises she had been working on. A classroom full of students was 14 feet away. Bernice knew that she was not allowed to hum within 15 feet of an occupied classroom, but that classroom typically wasn't used in the mornings, so Bernice thought it was empty. Since Bernice believed the classroom was empty, she thought it would be okay to practice her humming.

The class's teacher heard Bernice's humming and reported her to the office. Because Bernice was humming within 15 feet of an occupied classroom, she was sent to the principal's office.

Suppose that Bernice had actually realized, before she began humming, that her actions would violate the rule – that is, that the classroom was in fact occupied. And suppose that she decided to start humming anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting humming within 15 feet of an occupied classroom, and Bernice practiced her humming at her locker. In this case, with no rule about humming in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

VITAMINS: Mens Rea + Arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The students at Barry's school are all very concerned about their academic and athletic performance. In order to achieve as much as possible, some of the students at Barry's school have started using a special vitamin to help them stay awake and focused. Because the side effects of this vitamin can be harmful – leading to agitation, aggression, and hot flashes – there's a rule that students are not allowed to possess or give to another student more than 15 milligrams (in earth measurements) of the drug. The school administration chose the dose of 15 milligrams somewhat arbitrarily. There isn't any reason why taking 15.5 milligrams would be more likely to cause the side effects than taking 14.5, but they had to choose some dose cutoff, and decided on 15.

One day, a student asked Barry for some of the vitamin. Barry gave the student four pills totaling 16 milligrams. Barry knew that he was not allowed to give out more than 15 milligrams of the vitamin, but he thought the pills he had were 3 milligrams each when in fact they were 4 milligrams each, so Barry mistakenly believed he had given the other student only 12 milligrams.

A nearby teacher saw Barry give out the pills and stopped the other student to see how much Barry had given him. The teacher discovered it was 16 milligrams. Because Barry gave out 16 milligrams of the vitamin, he was sent to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he gave the vitamin to the other student, that his actions violated the rule – that is, that was giving out more than 15 milligrams. And suppose that he decided to give the boy the vitamins anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting giving more than 15 milligrams of the vitamin to other students, and Barry gave the other student the vitamins. In this case, with no

rule about giving out the vitamins in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

VITAMINS: Mens Rea + Non-arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The students at Barry's school are all very concerned about their academic and athletic performance. In order to achieve as much as possible, some of the students at Barry's school have started using a special vitamin to help them stay awake and focused. Because the side effects of this vitamin can be harmful – leading to agitation, aggression, and hot flashes – there's a rule that students are not allowed to possess or give to another student more than 15 milligrams (in earth measurements) of the drug. The school administration chose the dose of 15 milligrams after careful research on alien physiology. Aliens of their species become susceptible to the negative side effects of the vitamin at precisely 15 milligrams. At lower doses, and the negative side effects are extremely rare.

One day, a student asked Barry for some of the vitamin. Barry gave the student four pills totaling 16 milligrams. Barry knew that he was not allowed to give out more than 15 milligrams of the vitamin, but he thought the pills he had were 3 milligrams each when in fact they were 4 milligrams each, so Barry mistakenly believed he had given the other student only 12 milligrams.

A nearby teacher saw Barry give out the pills and stopped the other student to see how much Barry had given him. The teacher discovered it was 16 milligrams. Because Barry gave out 16 milligrams of the vitamin, he was sent to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he gave the vitamin to the other student, that his actions violated the rule – that is, that was giving out more than 15 milligrams. And suppose that he decided to give the boy the vitamins anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting giving more than 15 milligrams of the vitamin to other students, and Barry gave the other student the drugs. In this case, with no rule about giving out the vitamins in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

DUMPING: Strict Liability + Arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The school offers special electives for students, including a class on hover car maintenance and repair, which produces some oil and other waste. This waste only poses an environmental hazard if it gets into a waterway. Once the waste comes into contact with water, it can be harmful if it

comes in contact with the alien's skin – leading to itching, peeling, and even mild burns. As a result, there's a rule that students are not allowed to dump the waste within 15 feet (in earth measurements) of the river behind the school. The school administration chose the distance of 15 feet somewhat arbitrarily. There isn't any reason to suspect that it would be much riskier to dispose of the waste at 14.5 feet as opposed to 15.5 feet, but they had to choose some distance cutoff, and decided on 15 feet.

One day, Barry was told to dispose of his class waste. Barry dumped the waste 14 feet away from the river. Barry knew that he was not allowed to dump the waste within 15 feet of the river, but he thought he had seen another student dump waste in the same spot, so mistakenly believed that it was 15 feet away. Since Barry believed he was far enough away, he dumped his waste.

Barry's teacher looked out the window and saw Barry dumping the waste. The teacher knew the spot was not 15 feet away. The teacher reported Barry to the principal. Because Barry dumped waste within 15 feet of the river, he was called to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he dumped his waste, that his actions violated the rule – that is, that he was not 15 feet away from the river. And suppose that he decided to dump the waste there anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting dumping within 15 feet of the river, and Barry dumped his waste that close anyway. In this case, with no rule about dumping in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

DUMPING: Strict Liability + Non-arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The school offers special electives for students, including a class on hover car maintenance and repair, which produces some oil and other waste. This waste only poses an environmental hazard if it gets into a waterway. Once the waste comes into contact with water, it can be harmful if it comes in contact with the alien's skin – leading to itching, peeling, and even mild burns. As a result, there's a rule that students are not allowed to dump the waste within 15 feet (in earth measurements) of the river behind the school. The school administration chose the distance of 15 feet after careful research on the waste and the land around the school. The research showed that if the waste traveled more than 15 feet, it was no longer in concentrations that could be harmful when in contact with water, while even slightly shorter distances could permit harmful concentrations to reach the river.

One day, Barry was told to dispose of his class waste. Barry dumped the waste 14 feet away from the river. Barry knew that he was not allowed to dump the waste within 15 feet of the river,

but he thought he had seen another student dump waste in the same spot, so mistakenly believed that it was 15 feet away. Since Barry believed he was far enough away, he dumped his waste.

Barry's teacher looked out the window and saw Barry dumping the waste. The teacher knew the spot was not 15 feet away. The teacher reported Barry to the principal. Because Barry dumped waste within 15 feet of the river, he was called to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he dumped his waste, that his actions violated the rule – that is, that he was not 15 feet away from the river. And suppose that he decided to dump the waste there anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting dumping within 15 feet of the river, and Barry dumped his waste that close anyway. In this case, with no rule about dumping in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

LASERS: Strict Liability + Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class in which the students learn about engineering by building things. One of the big projects in the class is that each student builds his or her own operational laser gun. Because the laser guns can cause harm to other students – leading to pain, burns, and even temporary loss of sight – all students are required to obey strict rules about their laser guns, including registering any laser guns over 15-inches long at the front office, where it becomes associated with a unique ID for each student. The school administrators chose to make registration mandatory for laser guns of over 15-inches somewhat arbitrarily. There isn't any reason why an unregistered gun of 15.5 inches is a greater risk than an unregistered gun of 14.5 inches, but they had to choose some regulations, and decided that registration for guns over 15-inches long should be among them.

One day, another student, who knew Bernice had just completed her laser gun, asked to see it. Bernice took it out of her locker to show the other student. Bernice knew that she was not allowed to have the laser gun of 15-inches or longer in school unless it was registered with the office, but she had made an honest mistake when she measured it, and thought that it was only 14.5-inches long, when in fact it was 15.5.

A teacher saw Bernice showing the other student her laser gun and called the front office to see if Bernice had registered it. The office staff told him that she had not. The teacher immediately reported Bernice to the principal. Because Bernice had an unregistered laser gun of more than 15-inches long on school property, she was sent to the principal's office.

Suppose that Bernice had actually realized, when she took the unregistered laser gun to school, that her actions violated the rule – that is, that her laser gun was over 15-inches long and should have been registered. And suppose that she decided to keep it at school without registering it anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting keeping an unregistered laser gun (of any size) on school grounds, and that Bernice kept hers in her locker. In this case, with no rule about registration in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

LASERS: Strict Liability + Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class in which the students learn about engineering by building things. One of the big projects in the class is that each student builds his or her own operational laser gun. Because the laser guns can cause harm to other students – leading to pain, burns, and even temporary loss of sight – all students are required to obey strict rules about their laser guns, including registering any laser guns over 15-inches long at the front office, where it becomes associated with a unique ID for each student. The school administrators chose to make registration mandatory for laser guns of over 15-inches after careful research. With the technology used by the aliens, laser guns of 15-inches or longer can harness enough energy to cause lasting as opposed to superficial harm.

One day, another student, who knew Bernice had just completed her laser gun, asked to see it. Bernice took it out of her locker to show the other student. Bernice knew that she was not allowed to have the laser gun of 15-inches or longer in school unless it was registered with the office, but she had made an honest mistake when she measured it, and thought that it was only 14.5-inches long, when in fact it was 15.5.

A teacher saw Bernice showing the other student her laser gun and called the front office to see if Bernice had registered it. The office staff told him that she had not. The teacher immediately reported Bernice to the principal. Because Bernice had an unregistered laser gun of more than 15-inches long on school property, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, when she took the unregistered laser gun to school, that her actions violated the rule – that is, that her laser gun was over 15-inches long and should have been registered. And suppose that she decided to keep it at school without registering it anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

Finally, suppose that Bernice's school had no rule prohibiting keeping an unregistered laser gun (of any size) on school grounds, and that Bernice kept hers in her locker. In this case, with no rule about registration in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

SPEEDING: Strict Liability + Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class on how to drive hover cars, which is known for being very strict about particular rules and procedures. Because hover cars have the potential to go quite fast – leading to a higher potential for accidents – there's a rule that students are not allow to go faster than 15 miles per hour (in earth measurements). The school administrators chose the speed limit somewhat arbitrarily. There isn't any reason to believe that driving 15.5 miles per hour is much more dangerous than driving 14.5 miles per hour, but they had to choose some speed cutoff, and decided on 15.

One day, Bernice was driving the school's hover car in the practice field. The hover car was going 16 miles per hour. Bernice knew she was not allowed to drive over 15 miles per hour, but she had misread the speedometer, so mistakenly believed she was going 14.

Bernice's driving instructor noticed that Bernice appeared to be going faster than the speed limit. He took out his speed gun and confirmed that she was going 16 miles per hour. Because Bernice was driving the hover car over 15 miles per hour, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, as she was driving, that her actions violated the rule – that is, that she was going over 15 miles per hour. And suppose that she decided to drive that fast anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting driving faster than 15 miles per hour, and Bernice drove 16 miles per hour. In this case, with no rule about speed limits in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

SPEEDING: Strict Liability + Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class on how to drive hover cars, which is known for being very strict about particular rules and procedures. Because hover cars have the potential to go quite fast – leading to a higher potential for accidents – there's a rule that

students are not allow to go faster than 15 miles per hour (in earth measurements). The school administrators chose the speed limit of 15 miles per hour after careful research on hover car engineering. Due to the way the engines are designed, hover cars become more difficult to control at precisely 15.5 miles per hour.

One day, Bernice was driving the school's hover car in the practice field. The hover car was going 16 miles per hour. Bernice knew she was not allowed to drive over 15 miles per hour, but she had misread the speedometer, so mistakenly believed she was going 14.

Bernice's driving instructor noticed that Bernice appeared to be going faster than the speed limit. He took out his speed gun and confirmed that she was going 16 miles per hour. Because Bernice was driving the hover car over 15 miles per hour, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, as she was driving, that her actions violated the rule – that is, that she was going over 15 miles per hour. And suppose that she decided to drive that fast anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting driving faster than 15 miles per hour, and Bernice drove 16 miles per hour. In this case, with no rule about speed limits in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.).

Supplementary Material 6: Experiment 4 Stimuli

Arbitrary and Non-Arbitrary Stimuli

MOVIES: Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students of different ages, including a class on horror films, which is known for showing very gory movies. Because the content of the films can be harmful for younger aliens – leading to intense fear and recurring, violent nightmares – there's a rule that students are not allowed to share videos from the class with other students who are younger than 15-years old (in earth years).

The school administrators chose the age of 15 based on the recommendation from the studio. The studio chose the age of 15 somewhat arbitrarily. There isn't any reason why seeing the gory content would harm a 14.5-year old alien much more than a 15.5-year old alien, but they had to choose some age cutoff, and decided on 15.

Neither the school administrators nor the students have any idea how the studio chose the age and believe the age was selected arbitrarily. In fact, Bernice once asked a school administrator why the age was 15, and was told that it was an arbitrary cut-off determined by the movie studio.

One day, a student asked Bernice to borrow one of the videos from the class. The student was 14. Bernice knew that she was not allowed to lend the movie to students younger than 15, but Bernice had been in a class with the other student, so mistakenly believed she was also 15. Since Bernice believed the other student was old enough, she lent her the movie.

The younger student's parents found the video in their daughter's backpack, and the daughter told them that Bernice had given it to her. The parents immediately called the school to complain. Because Bernice lent a horror film to a student younger than 15, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, when she gave the video to the other student, that her actions violated the rule – that is, that the student was younger than 15. And suppose that she decided to give the girl the video anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting giving the videos to students younger than 15, and Bernice gave the video to the younger student. In this case, with no rule about lending videos in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

MOVIES: Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students of different ages, including a class on horror films, which is known for showing very gory movies. Because the content of the films can be harmful for younger aliens – leading to intense fear and recurring, violent nightmares – there's a rule that students are not allowed to share videos from the class with other students who are younger than 15-years old (in earth years).

The school administrators chose the age of 15 based on the recommendation from the studio. The studio chose the age of 15 after careful research on alien development. Aliens of their species undergo an important developmental change such that younger aliens are susceptible to the negative effects of gory and violent content, while older aliens are not. The developmental change is linked to an increase in hormone levels that occurs, like clockwork, at age 15. The studio feared that releasing this information would damage their sales and reputation so they have carefully kept it from the public. Only the top executives and a few scientists even know the study was conducted.

Neither the school administrators nor the students have any idea how the studio chose the age and believe the age was selected arbitrarily. In fact, Bernice once asked a school administrator why the age was 15, and was told that it was an arbitrary cut-off determined by the movie studio.

One day, a student asked Bernice to borrow one of the videos from the class. The student was 14. Bernice knew that she was not allowed to lend the movie to students younger than 15, but Bernice had been in a class with the other student, so mistakenly believed she was also 15. Since Bernice believed the other student was old enough, she lent her the movie.

The younger student's parents found the video in their daughter's backpack, and the daughter told them that Bernice had given it to her. The parents immediately called the school to complain. Because Bernice lent a horror film to a student younger than 15, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, when she gave the video to the other student that her actions violated the rule – that is, that the student was younger than 15. And suppose that she decided to give the girl the video anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting giving the videos to students younger than 15, and Bernice gave the video to the younger student. In this case, with no rule about lending videos in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

VITAMINS: Arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The students at Barry's school are all very concerned about their academic and athletic performance. In order to achieve as much as possible, some of the students at Barry's school have started using a special vitamin to help them stay awake and focused. Because the side effects of this vitamin can be harmful – leading to agitation, aggression, and hot flashes – there's a rule that students are not allowed to possess or give to another student more than 15 milligrams (in earth measurements) of the drug.

The school administration chose the dose of 15 milligrams based on the recommendation from the manufacturer. The manufacturer chose the level somewhat arbitrarily. There isn't any reason why taking 15.5 milligrams would be much more likely to cause the side effects than taking 14.5, but they had to choose some dose cutoff, and decided on 15.

Neither the school administrators nor the students have any idea how the manufacturer chose the level and believe the level was selected arbitrarily. In fact, Barry once asked a school administrator why a dose of 15 milligrams was the maximum, and was told that it was an arbitrary cut-off determined by the manufacturer of the vitamin.

One day, a student asked Barry for some of the vitamin. Barry gave the student four pills totaling 16 milligrams. Barry knew that he was not allowed to give out more than 15 milligrams of the vitamin, but he thought the pills he had were 3 milligrams each when in fact they were 4 milligrams each, so Barry mistakenly believed he had given the other student only 12 milligrams.

A nearby teacher saw Barry give out the pills and stopped the other student to see how much Barry had given him. The teacher discovered it was 16 milligrams. Because Barry gave out 16 milligrams of the vitamin, he was sent to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he gave the vitamin to the other student, that his actions violated the rule – that is, that was giving out more than 15 milligrams. And suppose that he decided to give the boy the vitamins anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting giving more than 15 milligrams of the vitamin to other students, and Barry gave the other student the vitamins. In this case, with no rule about giving out the vitamins in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

VITAMINS: Non-arbitrary

Unknowing:
Barry is an alien from Planet Zorg. He attends a high school with other aliens of his species. The students at Barry's school are all very concerned about their academic and athletic performance. In order to achieve as much as possible, some of the students at Barry's school have started using a special vitamin to help them stay awake and focused. Because the side effects of this vitamin can be harmful – leading to agitation, aggression, and hot flashes – there's a rule that students are not allowed to possess or give to another student more than 15 milligrams (in earth measurements) of the drug.

The school administration chose the dose of 15 milligrams based on the recommendation from the manufacturer. The manufacturer chose the level after careful research on alien physiology. Aliens of their species become susceptible to the negative side effects of the vitamin at precisely 15 milligrams. At lower doses, and the negative side effects are extremely rare. The manufacturer feared that releasing this information would damage their sales and reputation so they have carefully kept it from the public. Only the top executives and a few scientists even know the study was conducted.

Neither the school administrators nor the students have any idea how the manufacturer chose the level and believe the level was selected arbitrarily. In fact, Barry once asked a school administrator why a dose of 15 milligrams was the maximum, and was told that it was an arbitrary cut-off determined by the manufacturer of the vitamin.

One day, a student asked Barry for some of the vitamin. Barry gave the student four pills totaling 16 milligrams. Barry knew that he was not allowed to give out more than 15 milligrams of the vitamin, but he thought the pills he had were 3 milligrams each when in fact they were 4 milligrams each, so Barry mistakenly believed he had given the other student only 12 milligrams.

A nearby teacher saw Barry give out the pills and stopped the other student to see how much Barry had given him. The teacher discovered it was 16 milligrams. Because Barry gave out 16 milligrams of the vitamin, he was sent to the principal's office.

Knowledge:
Suppose that Barry had actually realized, when he gave the vitamin to the other student, that his actions violated the rule – that is, that was giving out more than 15 milligrams. And suppose that he decided to give the boy the vitamins anyway. In this case, where Barry knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Barry's school had no rule prohibiting giving more than 15 milligrams of the vitamin to other students, and Barry gave the other student the drugs. In this case, with no rule about giving out the vitamins in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

SPEEDING: Arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class on how to drive hover cars, which is known for being very strict about particular rules and procedures. Because hover cars have the potential to go quite fast – leading to a higher potential for accidents – there's a rule that students are not allow to go faster than 15 miles per hour (in earth measurements).

The school administrators chose the speed limit based on the recommendation from the manufacturer. The manufacturer chose the speed limit somewhat arbitrarily. There isn't any reason to believe that driving 15.5 miles per hour is much more dangerous than driving 14.5 miles per hour, but they had to choose some speed cutoff, and decided on 15.

Neither the school administrators nor the students have any idea how the manufacturer chose the speed and believe the speed was selected arbitrarily. In fact, Bernice once asked a school administrator why the speed limit was 15 mph, and was told that it was an arbitrary cut-off determined by the manufacturer of the hover cars.

One day, Bernice was driving the school's hover car in the practice field. The hover car was going 16 miles per hour. Bernice knew she was not allowed to drive over 15 miles per hour, but she had misread the speedometer, so mistakenly believed she was going 14.

Bernice's driving instructor noticed that Bernice appeared to be going faster than the speed limit. He took out his speed gun and confirmed that she was going 16 miles per hour. Because Bernice was driving the hover car over 15 miles per hour, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, as she was driving, that her actions violated the rule – that is, that she was going over 15 miles per hour. And suppose that she decided to drive that fast anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting driving faster than 15 miles per hour, and Bernice drove 16 miles per hour. In this case, with no rule about speed limits in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

SPEEDING: Non-arbitrary

Unknowing:
Bernice is an alien from Planet Zorg. She attends a high school with other aliens of her species. The school offers special electives for students, including a class on how to drive hover cars, which is known for being very strict about particular rules and procedures. Because hover cars

have the potential to go quite fast – leading to a higher potential for accidents – there's a rule that students are not allow to go faster than 15 miles per hour (in earth measurements).

The school administrators chose the speed limit of 15 miles per hour based on the recommendation from the manufacturer. The manufacturer chose the speed after careful research on hover car engineering. Due to the way the engines are designed, hover cars become more difficult to control at precisely 15.5 miles per hour. The manufacturer feared that releasing this information would damage their sales and reputation so they have carefully kept it from the public. Only the top executives and a few scientists even know the study was conducted.

Neither the school administrators nor the students have any idea how the manufacturer chose the speed and believe the speed was selected arbitrarily. In fact, Bernice once asked a school administrator why the speed limit was 15 mph, and was told that it was an arbitrary cut-off determined by the manufacturer of the hover cars.

One day, Bernice was driving the school's hover car in the practice field. The hover car was going 16 miles per hour. Bernice knew she was not allowed to drive over 15 miles per hour, but she had misread the speedometer, so mistakenly believed she was going 14.

Bernice's driving instructor noticed that Bernice appeared to be going faster than the speed limit. He took out his speed gun and confirmed that she was going 16 miles per hour. Because Bernice was driving the hover car over 15 miles per hour, she was sent to the principal's office.

Knowledge:
Suppose that Bernice had actually realized, as she was driving, that her actions violated the rule – that is, that she was going over 15 miles per hour. And suppose that she decided to drive that fast anyway. In this case, where Bernice knowingly violated the rule, how would you respond to the following questions? (Your responses may be the same as those you just provided, or they may differ.)

No Rule:
Finally, suppose that Bernice's school had no rule prohibiting driving faster than 15 miles per hour, and Bernice drove 16 miles per hour. In this case, with no rule about speed limits in place, how would you respond to the following questions? (Your responses may be the same as those you've provided, or they may differ.)

Chapter 4 Appendices and Supplemental Material

<u>**Appendix A:**</u>
<u>**Experiment 1 Stimuli**</u>

Neurotypical

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over.

John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

—Page Break—

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John sung at him.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's lawyer argued that John was trying to defend himself, saying that the altercation was the result of the other man's actions. John testified that at the time of the fight, he believed the other man posed a threat and could injure him.

Mental Illness, Labeled

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over.

John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

—Page Break—

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with schizophrenia, explaining that he suffers from abnormal thoughts and feelings. John testified that at the time of the fight, he believed the other man had been sent by his enemies to track and injure him.

Mental Illness, Unlabeled

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over.

John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

—Page Break—

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John suffers from abnormal thoughts and feelings. John

testified that at the time of the fight, he believed the other man had been sent by his enemies to track and injure him.

Supplemental Experiment 1

Experiment 1 found significant differences between our *Neurotypical* and *Mental Illness* groups on most of our dependent variables. Most importantly for this study, we found that projection and similarity were significantly lower for our *Mental Illness* group, while stereotyping was significantly higher. We also found that similarity was a partial mediator of these differences.

One natural question is whether these results are specific to the mental illness considered, schizophrenia, or extend more broadly. Another question is what it is about defendants with a mental illness that decreases perceived similarity, decreases projection, increases stereotyping, and has other downstream effects. Prior work finds that stereotypes about mental illnesses vary widely. For instance, stereotypes of people with schizophrenia are far different from stereotypes of people with obsessive compulsive disorder (Sadler, Meagor, & Kaye, 2012; Sadler, Kaye, & Vaughn, 2015). By considering how people evaluate defendants with mental illnesses that vary along the key dimensions of warmth and competence (Sadler, Meagor, & Kaye, 2012), we can simultaneously evaluate how broadly the findings of Experiment 1 generalize to other categories of mental illness, and whether mental illness has an effect because it influences perceptions of warmth and competence.

Methods

Participants. Six-hundred-and-forty adults (369 female, 269 male, 2 other/prefer not to specify, mean age = 34, SD = 11) participated in the study through Amazon Mechanical Turk for monetary compensation. An additional 58 participants were tested, but were excluded for failing catch questions. Participation was restricted to workers with IP addresses in the United States and a prior approval rating of 95% or higher on previous tasks.

Materials & Procedures. The experimental stimuli consisted of four vignettes to which participants were randomly assigned. These four vignettes corresponded to the *Neurotypical*, *Schizophrenia*, *Post-Traumatic Stress Disorder (PTSD)*, and *Obsessive Compulsive Disorder (OCD)* conditions. These three mental illnesses were chosen because prior research has found that they vary significantly in terms of the warmth and competence associated with their corresponding stereotypes (Sadler, Meagor, & Kaye, 2012).[34] The criminal action described for each illness was nearly identical to that used in Experiment 1. The initial portion varied from the one used in Experiment 1 only in that participants were informed that John "always followed the same path, on the same side of the sidewalk, and changing his pattern upset him." This addition was identical in all four conditions, but was included to make it more plausible that the alleged battery related to the defendant's mental illness when the illness was OCD.

As in Experiment 1, the final paragraph of each condition was the only component that differed. The last paragraph for *Neurotypical* and *Schizophrenia* were identical to those used in the *Neurotypical* and *Labeled* conditions in Experiment 1. The relevant text for PTSD and OCD are excerpted below (for full stimuli, see supplemental materials).

---

[34] Our own pilot testing confirmed that the stereotypes corresponding to each of these three mental illnesses differ along these dimensions.

PTSD. "At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with post-traumatic stress disorder, explaining that John suffers from tense feelings, and frightening and negative thoughts about the world. John testified that at the time of the fight, he believed the other man posed a threat and could injure him."

OCD. "At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with obsessive compulsive disorder, explaining that John suffers from fears and obsessive thoughts which lead to his feeling compelled to perform behaviors to quiet the thoughts. John testified that at the time of the fight, he believed the other man posed a threat and could injure him."

After reading their assigned vignette, participants selected verdicts and sentencing in exactly the same way they had in Experiment 1. Next, participants answered a series of evaluative questions designed to assess how unpredictable, competent, warm, and in control of his behavior the defendant is. Competence and warmth had been found to be important in previous research (Sadler, Meagor, & Kaye, 2012), and our own pre-testing indicated that unpredictability and control might additionally be important. Based on prior work and pilot testing, we predicted that our selected mental illness would vary along these four facets, and that these facets might better explain verdict, punishment, and judgment strategy differences than the broader category of mental illness. Each facet had four questions, modeled after those used in Sadler, Meagor, & Kaye (2012). The questions were split across two screens, and presented in a random order. All were rated from 1 (not at all) to 7 (a great deal). Those preceded by an asterisk were reverse scored in analysis.

Unpredictable 1. "In general, how much do Americans believe that people like John display unpredictable behavior?"

Unpredictable 2. "In general, how much do Americans believe that people like John will react erratically?"

*Unpredictable 3. "In general, how much do Americans believe that people like John make decisions that are easy to guess in advance?"

*Unpredictable 4. "In general, how much do Americans believe that people like John will behave in ways we would expect of others?"

Competent 1. "In general, how much do Americans believe that people like John are intelligent?"

Competent 2. "In general, how much do Americans believe that people like John are competent?"

Competent 3. "In general, how much do Americans believe that people like John are skilled?"

Competent 4. "In general, how much do Americans believe that people like John are capable?"

Warm 1. "In general, how much do Americans believe that people like John are friendly?"

Warm 2. "In general, how much do Americans believe that people like John are warm?"

Warm 3. "In general, how much do Americans believe that people like John are good natured?"

Warm 4. "In general, how much do Americans believe that people like John are honest?"
Control 1. "In general, how much do Americans believe that people like John could control their behavior if they tried?"
Control 2. "In general, how much do Americans believe that people like John could control their thoughts and feelings if they tried?"
Control 3. "In general, how much do Americans believe that people like John need some form of treatment in order to change their behavior?"
Control 4. "In general, how much do Americans believe that people like John lack control over their behavior?"

Then, on a separate screen, participants answered the projection, stereotyping, and reasonable person questions from Experiment 1. Next, on a separate screen, participants answered the blame, incapacitation, retribution, and deterrence questions from Experiment 1 in a random order. On another separate screen, in random order, participants answered the similarity questions from Experiment 1, questions meant to assess feelings of fear, anger, and recidivism risk, and two questions that assed whether they believed John's account of the altercation. New questions are quoted below; the first two were rated from 1 (not at all) to 7 (a great deal), while the final two were true/false:

Anger/Fear. "In general, how angry and/or afraid are Americans about people like John?"
Recidivism. "In general, how much do Americans believe that people like John will commit another crime in the future?"
Posed Threat. "At the time of the fight, John believed the other man posed a threat and could injure him."
Saw Gun. "When the man reached for his waistband, John thought he was reaching for a gun."

Then, as in Experiment 1, participants answered a simple true/false question about the vignette, and a reading comprehension check question. Participants who failed either question were excluded from further analysis. Finally, participants provided their age and a gender.

Results

Verdict. In the *Neurotypical* condition, 26% of participants selected a guilty verdict. Across the three mental illness conditions, 33% of participants selected a guilty verdict. A 2 (defendant condition: *Neurotypical*, *Mental Illness*) by 2 (verdict: not guilty, guilty) chi-square test revealed a marginally significant effect, X^2 (1, $N = 640$) = 2.98, $p < .084$. However, the percentage of participants selecting a guilty verdict also varied across the three mental illnesses: 48% for schizophrenia, 29% for PTSD, and 60% for OCD. A 2 (defendant condition: *Neurotypical*, *OCD*) by 2 (verdict, not guilty, guilty) chi-square test found a significant difference, X^2 (1, $N = 320$) = 5.22, $p < .022$; other pairwise comparisons were not significant.

Sentences. As in Experiment 1, we assigned a 0 in cases where participants did not assign a particular punishment, and again evaluated the data using non-parametric tests (see Table 1). Mann-Whitney U tests found only one significant difference between the *Neurotypical* and overarching *Mental Illness* conditions; participants in the *Mental Illness* condition assigned

significantly more counseling, $U = 23658.00$, $p = .001$. Fines, jail time, and service amounts were all higher for *Mental Illness*, but not significantly.

We next analyzed the three mental illness conditions separately. We performed Kruskal Wallis H tests on each sentencing variable to test for significant differences across all four conditions, revealing a significant effect for fines, X^2 (3, $N = 640$) = 7.87, $p < .049$, and counseling, X^2 (3, $N = 640$) = 55.68, $p < .001$. Mann Whitney U tests were then performed to assess which groups significantly varied. For fines, *PTSD* ($M = 200.13$, $SD = 385.28$) was significantly lower than *Neurotypical* ($M = 267.11$, $SD = 396.77$), $U = 10943.50$, $p = .011$, and *OCD* ($M = 324.85$, $SD = 571.78$), $U = 11056.00$, $p = .016$. No other significant differences were found. Counseling results simply mirrored the overarching results: *Neurotypical* was lower than *Schizophrenia*, $U = 7786.00$, $p = .001$, *PTSD*, $U = 7638.50$, $p = .001$, and *OCD*, $U = 8233.50$, $p = .001$, but no significant differences were found across the three *Mental Illness* conditions.

Table S1: Means and standard deviations of sentencing variables for the Mental Illness groups.

	Neurotypical	Schizophrenia	PTSD	OCD
Fines	**267.11**	**299.81**	**200.13**	**324.85**
	(396.70)	**(575.51)**	**(385.28)**	**(571.78)**
Jail Time	2.54 (14.76)	4.90 (34.92)	1.16 (7.99)	2.11 (19.22)
Community Service	39.17 (54.05)	43.91 (59.87)	35.99 (49.84	41.29 (51.88)
Counseling	**3.12 (4.51)**	**7.99 (7.74)**	**7.39 (6.99)**	**7.39 (7.30)**

Judgment Strategies. We first conducted *t*-tests comparing ratings for each judgment strategy question as a function of defendant condition (*Neurotypical*, *Mental Illness*). As in Experiment 1, participants in the *Neurotypical* condition reported higher levels of projection, $t(638) = 3.13$, $p < .002$, $d = .25$, and reliance on a reasonable person standard, $t(637) = 5.68$, $p < .001$, $d = .45$ (corrected for violating Levene's test), while participants in the *Mental Illness* condition reported relying more heavily on stereotyping, $t(638) = 2.98$, $p < .003$, $d = .24$. Similarity was also significantly higher for the *Neurotypical* condition, $t(638) = 4.43$, $p < .001$, $d = .35$ (see Figure S1). No significant differences were found cross the three mental illness conditions.

Figure S1: Means for projection, stereotyping, reasonable person, and similarity as a function of mental condition in Experiment 2. Error bars are one SEM in each direction. All differences are significant.
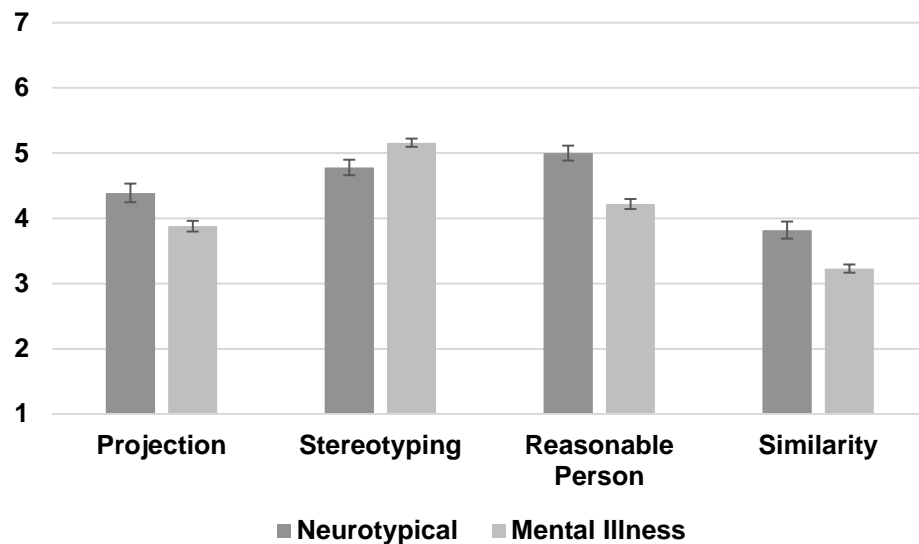
We next conducted a Sobel test to determine whether similarity again mediated levels of projection across our *Neurotypical* and *Mental Illness* conditions, and found a significant result, $t = -4.08$, *SE* $= 0.01$, $p < 0.001$. A stepwise linear progression revealed that similarity fully mediated the effect of defendant condition on projection in Experiment 2 (unlike Experiment 1, in which the mediation was partial). In the first step, mental condition was a significant predictor of projection, accounting for 2% of the variance, $F(1, 638) = 9.80$, $p < .002$, $R^2 = .015$. In the second step, the addition of similarity was a significant improvement, accounting for 16% of the variance, $F(1, 637) = 58.54$, $p < .001$, $R^2 = .155$; R^2 change $= .140$, $p < .001$.

Blame and Punishment Justifications. As with our sentencing variables, we first analyzed each of our blame, punishment justification, anger/fear, and recidivism questions with a t-test comparing *Neurotypical* to *Mental Illness* (see Table 2). Participants' scores for incapacitation, $t(310) = 1.97$, $p < .049$, $d = .22$, recidivism, $t(243) = 4.56$, $p < .001$, $d = .59$ (both corrected for violating Levene's test), and anger/fear, $t(638) = 5.50$, $p < .001$, $d = .44$, were all significantly higher for *Mental Illness*.

Next, we performed ANOVA'S on each variable to determine whether participants gave significantly different scores in the different mental illness conditions. We found a significant effect for anger/fear, $F(1, 636) = 14.83$, $p < .001$, $\eta^2 = .065$. *T*-tests revealed that participants in the *Schizophrenia* condition gave higher scores than those in the *Neurotypical*, $t(299) = 6.63$, $p < .001$, $d = .77$, *PTSD*, $t(314) = 2.98$, $p < .003$, $d = .34$, and *OCD* conditions, $t(307) = 3.69$, $p < .001$, $d = .42$ (all corrected for violating Levene's test), and that participants in the *Neurotypical* condition gave lower scores than participants in the *PTSD*, $t(311) = 3.74$, $p < .001$, $d = .42$ (corrected for violating Levene's test), and *OCD* conditions, $t(318) = 2.86$, $p < .005$, $d = .32$. No other significant differences were found.

The same pattern was observed for recidivism, $F(1, 636) = 10.43$, $p < .001$, $\eta^2 = .047$. Participants in the *Schizophrenia* condition gave higher ratings than participants in *Neurotypical*, $t(309) = 5.37$, $p < .001$, $d = .61$ (corrected for violating Levene's test), *PTSD*, $t(318) = 2.23$, $p < .027$, $d = .25$, and *OCD*, $t(318) = 2.53$, $p < .012$, $d = .28$, and participants in *Neurotypical* gave

significantly lower scores than participants in *PTSD*, $t(307) = 3.38$, $p < .001$, $d = .39$ (corrected for violating Levene's test), and *OCD*, $t(318) = 2.88$, $p < .004$, $d = .32$. No other significant differences were found.

Posed Threat and Saw Gun. To assess the difference between groups on our Posed Threat and Saw Gun questions, we first ran chi square tests on the *Neurotypical* and *Mental Illness* groups. We found no significant differences. Chi square tests considering all four defendant conditions also found no significant differences.

Mental Illness Facets. In Experiment 2, we set out to determine whether the differences we saw between our *Neurotypical* and *Mental Illness* conditions in Experiment 1 were due to overarching beliefs about all mental illnesses, or more specific beliefs about schizophrenia in particular. We identified four facets that vary in the stereotypes of our different mental illnesses, some based on prior research. Each facet was composed of four questions, and received high Cronbach Alpha ratings: warmth ($a = .92$), competence ($a = .92$), predictability ($a = .81$), and lack of control ($a = .71$).

An ANOVA found a significant difference between defendant conditions for warmth, $F(1, 636) = 10.43$, $p < .001$, $\eta^2 = .047$. *T*-tests revealed that participants in the *Schizophrenia* condition gave significantly lower ratings than those in the *Neurotypical*, $t(318) = 4.98$, $p < .001$, $d = .56$, *PTSD*, $t(318) = 3.58$, $p < .001$, $d = .40$, and *OCD* conditions, $t(318) = 3.85$, $p < .001$, $d = .43$. Competence showed a similar pattern, the ANOVA revealed significant differences between the defendant conditions, $F(1, 636) = 13.24$, $p < .001$, $\eta^2 = .059$, and *t*-tests showed that participants in the *Schizophrenia* condition gave lower scores than participants in the *Neurotypical*, $t(318) = 6.26$, $p < .001$, $d = .70$, *PTSD*, $t(318) = 4.48$, $p < .001$, $d = .50$, and *OCD* conditions, $t(318) = 3.14$, $p < .002$, $d = .35$. However, in this case, participants in the *OCD* condition also gave significantly lower scores than those in the *Neurotypical* condition, $t(318) = 2.98$, $p < .003$, $d = .33$.

For unpredictability, we found the opposite pattern. An ANOVA revealed a significant effect of defendant condition, $F(1, 636) = 18.18$, $p < .001$, $\eta^2 = .079$, which was driven by significantly lower scores given by participants in the *Neurotypical* condition than participants in the *Schizophrenia*, $t(318) = 7.43$, $p < .001$, $d = .83$, *PTSD*, $t(318) = 3.69$, $p < .001$, $d = .41$, and *OCD* conditions, $t(318) = 4.54$, $p < .001$, $d = .51$. Additionally, participants in the *Schizophrenia* condition gave significantly higher ratings than the *PTSD*, $t(318) = 3.52$, $p < .001$, $d = .39$, and *OCD* conditions, $t(318) = 2.80$, $p < .005$, $d = .31$. Lack of control showed the same pattern; participants in the *Neurotypical* condition gave significantly lower scores than participants in the *Schizophrenia*, $t(318) = 7.06$, $p < .001$, $d = .79$, *PTSD*, $t(318) = 4.94$, $p < .001$, $d = .55$, and *OCD*, $t(318) = 4.05$, $p < .001$, $d = .45$, conditions, and participants in the *Schizophrenia* condition gave significantly higher scores than participants in the *PTSD*, $t(318) = 2.54$, $p < .012$, $d = .28$, and *OCD* conditions, $t(318) = 2.96$, $p < .003$, $d = .33$. No other significant differences for any of these variables were found. See Figures S2a and S2b.

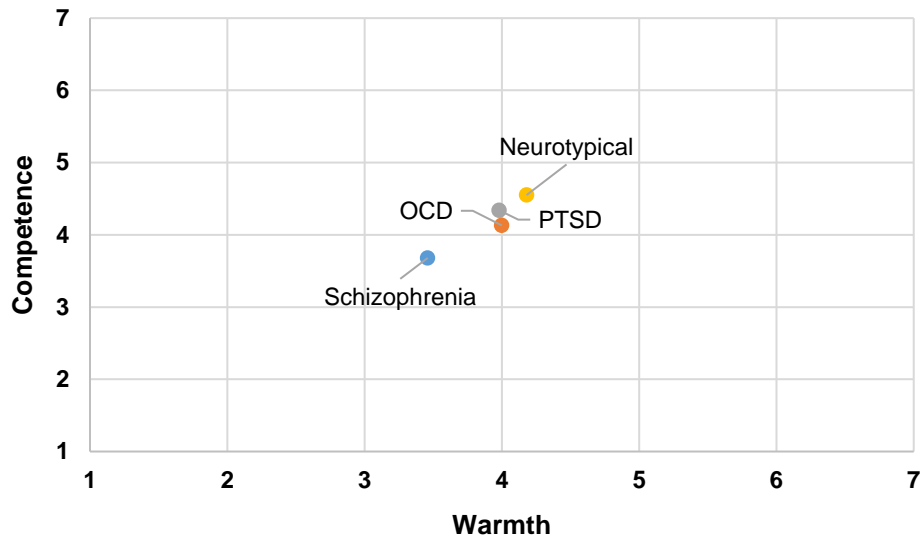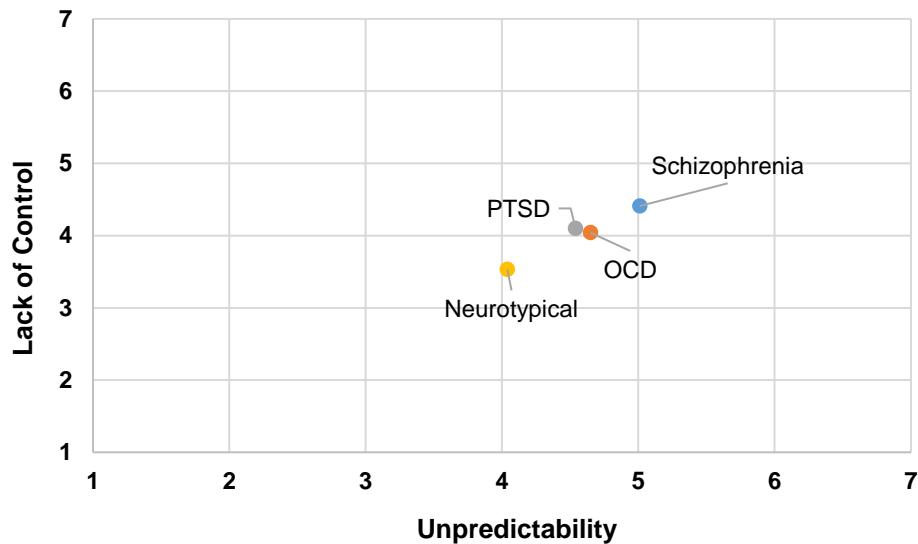Figure S2a: Defendant conditions plotted on warmth and competence facets.

Figure S2b: Defendant conditions plotted on unpredictability and lack of control facets.

Summary of experimental findings. Consistent with Experiment 1 and our predictions, the findings for S1 show that participants judged defendants with a mental illness – any mental illness – more harshly than a defendant without a mental illness. The defendants described as having a mental illness were found guilty more often (significantly more for a defendant with OCD), and given longer, though not significantly longer, sentences. Additionally, we again found that participants reported greater projection and considered themselves more similar to the defendant in our *Neurotypical* condition, than the defendant in any of our *Mental Illness* conditions, and participants justified punishment for the defendant in the *Mental Illness* conditions based more heavily on a desire for incapacitation. New to Experiment S1, we also found that participants in the *Mental Illness* conditions believe that Americans are significantly more angry and afraid of the defendant, as well as believing it is significantly more likely that the defendant will reoffend. The results of our facet questions suggest that while defendants with any of our mental illnesses are viewed differently than a defendant without one, defendants with

schizophrenia may face a particularly difficult road. Our next step was to determine whether mindreading strategy predicted verdicts, and whether the facets that we examined – warmth, competence, control, and predictability – predicted verdicts.

Does judgment strategy predict verdict? A binary logistic regression predicting verdict (1=not guilty, 2=guilty) as a function of projection and stereotyping found that projection was a significant, negative predictor of verdicts, $\beta = -.225$, $p < .001$, while stereotyping was a significant, positive predictor, $\beta = .139$, $p < .030$. This indicates that as in Experiment 1, the more a participant used projection and the less she stereotyped, the more likely she was to choose a not guilty verdict.

Do the individual facets predict verdicts? We hypothesized that different treatment of our defendant conditions could stem, in part, from people's stereotypes about how prevalent each of the four facets of warmth, competence, unpredictability, and lack of control were for each illness. A binary logistic regression predicting verdict (1=not guilty, 2=guilty) as a function of each facet found that warmth was a significant, negative predictor of verdicts, $\beta = -.150$, $p < .025$, while unpredictability was a significant, positive predictor, $\beta = .156$, $p < .029$. Thus, the more warm and predictable a defendant was perceived to be, the less likely he was to be found guilty. Competence and lack of control were not significant predictors.

Discussion

S1 replicated some of the key findings of Experiment 1. We again found support for Hypothesis 1a and b, that projection is significantly higher, and stereotyping significantly lower, in our *Neurotypical* condition than for any *Mental Illness* condition. S1 found greater support for the role of similarity, with similarity fully mediating the effect of defendant condition on projection. Hypothesis 2a was also supported, with higher levels of projection significantly predicting not guilty verdicts. Unlike Experiment 1, S1 also supported Hypothesis 2b, finding that higher levels of stereotyping significantly predicted guilty verdicts. Hypothesis 3 received more mixed support. We found that, overall, participants in our *Mental Illness* conditions were marginally more likely to assign a guilty verdict, and that punishments were generally higher for our *Mental Illness* defendants.

We also found differences across the different mental illnesses that we considered. Our *PTSD* defendant received the lowest fines, jail time, and community service – lower even than *Neurotypical*. *PTSD* participants also gave guilty verdicts only 29% of the time, as opposed to 48% for *Schizophrenia* and 60% for *OCD*. While this pattern of verdicts appears puzzling, we believe it may stem from a simple root: OCD did not provide a satisfying explanation for the behavior described. That is, people's stereotypes of PTSD, and especially schizophrenia, more closely fit the action described in our vignette. Thus, OCD received more guilty verdicts as it had the double hurdle of being a mental illness, with all the negative implications we have found, and also of not being a particularly good (mitigating) explanation. Schizophrenia, while viewed more negatively on our measures, had the slightly mitigating effect of being a credible explanation for the behavior, leading to fewer guilty verdicts. The different mental illnesses also varied along our four facets. *Schizophrenia* was significantly lower in competence and warmth, and higher in unpredictability and lack of control than all other defendant conditions. We found that warmth and unpredictability were significant predictors of verdicts. *Neurotypical* was significantly lower in unpredictability and higher – though only significantly higher than *Schizophrenia* – in warmth

than our *Mental Illness* conditions. This suggests that part of what drives the difference between our *Neurotypical* and *Mental Illness* conditions is perceptions of warmth and unpredictability.

Neurotypical

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. When John walked downtown, he always followed the same path, on the same side of the sidewalk, and changing his pattern upset him.

As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over. John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

—Page Break—

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John swung at him.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's lawyer argued that John was trying to defend himself, saying that the altercation was the result of the other man's actions. John testified that at the time of the fight, he believed the other man posed a threat and could injure him.

Schizophrenia

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. When John walked downtown, he always followed the same path, on the same side of the sidewalk, and changing his pattern upset him.

As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over. John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

<div align="center">—Page Break—</div>

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John swung at him.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with schizophrenia, explaining that he suffers from abnormal thoughts and feelings. John testified that at the time of the fight, he believed the other man posed a threat and could injure him.

PTSD

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

<div align="center">—Page Break—</div>

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. When John walked downtown, he always followed the same path, on the same side of the sidewalk, and changing his pattern upset him.

As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over. John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

<div align="center">—Page Break—</div>

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John swung at him.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with post-traumatic stress disorder, explaining that John suffers from tense feelings, and frightening and negative thoughts about the world. John testified that at the time of the fight, he believed the other man posed a threat and could injure him.

OCD

John is a 21-year-old college student. He enjoys being outside – hiking, swimming, and doing nature photography. He also really likes playing video games and writing short stories. John was a straight A student all through high school and his first year of college, but last year his grades started going down. John is currently home for the summer.

—Page Break—

One night when he was home for summer break, John was walking downtown. The weather was unseasonably warm which had drawn many people outside, making the sidewalk rather crowded. When John walked downtown, he always followed the same path, on the same side of the sidewalk, and changing his pattern upset him.

As John was trying to navigate through the crowd, another man jostled him, nearly knocking him over. John turned around and confronted the man who had run into him. The other man reached for his waistband, and John thought he saw a gun.

—Page Break—

As soon as he thought he saw the gun, John swung at the other man, and the two of them ended up getting into a fight right there on the sidewalk. The police were called, and it turned out that the object the other man had been reaching for was his cell phone. The police arrested John, who was charged with battery. Battery is defined as any willful and unlawful use of force or violence upon the person of another.

The police were able to find a security camera on a nearby business that confirmed the above facts, including that the other man reached for his waistband just before John swung at him.

The above facts were undisputed at trial.

At John's trial, his lawyers plead not guilty. John's psychiatrist, a well-regarded doctor who had been seeing him for years, testified that John had been diagnosed with obsessive compulsive disorder, explaining that John suffers from fears and obsessive thoughts which lead to his feeling compelled to perform behaviors to quiet the thoughts. John testified that at the time of the fight, he believed the other man posed a threat and could injure him.