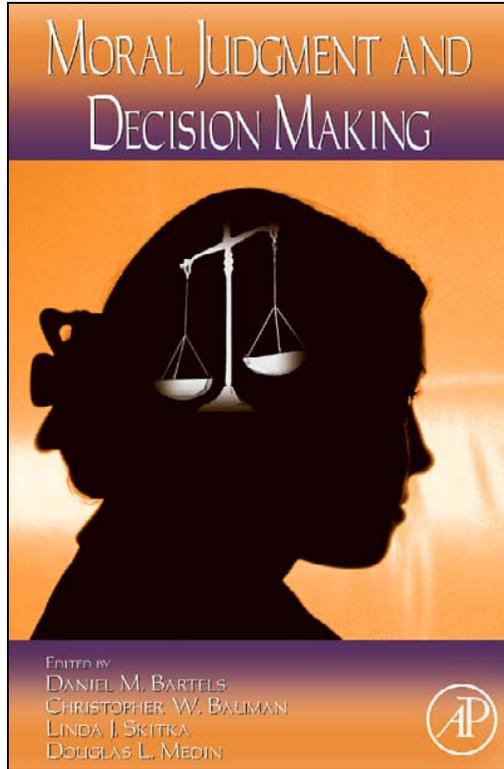


**Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.**

This chapter was originally published in the book *The Psychology of Learning and Motivation, Vol 50*, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who know you, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

From: Peter H. Ditto, David A. Pizarro, and David Tannenbaum,
Motivated Moral Reasoning.

In Daniel M. Bartels, Christopher W. Bauman, Linda J. Skitka, and Douglas
L. Medin, editors, *The Psychology of Learning and Motivation, Vol 50*,
Burlington: Academic Press, 2009, pp.307-338.

ISBN: 978-0-12-374488-3

© Copyright 2009 Elsevier Inc.
Academic Press.

MOTIVATED MORAL REASONING

Peter H. Ditto, David A. Pizarro, *and* David Tannenbaum

Contents

1. Introduction	307
2. Motivated Reasoning	309
3. Motivated Moral Reasoning	312
4. Motivated Assessments of Moral Accountability	315
4.1. Controllability and Culpability	316
4.2. Memory for the “Facts”	317
4.3. Intentionality	317
5. Motivated Use of Moral Principles	322
5.1. The Use and Abuse of Principle	322
5.2. Tales of Political Casuistry	325
5.3. Trolleys, Lifeboats, and Collateral Damage	326
5.4. Is Reliance on Moral Principles Really Motivated?	330
6. Motivated Moral Reasoning and Views of the Moral Thinker	332
References	334

Abstract

Moral judgments are important, intuitive, and complex. These factors make moral judgment particularly fertile ground for motivated reasoning. This chapter reviews research (both our own and that of others) examining two general pathways by which motivational forces can alter the moral implications of an act: by affecting perceptions of an actor's moral accountability for the act, and by influencing the normative moral principles people rely on to evaluate the morality of the act. We conclude by discussing the implications of research on motivated moral reasoning for both classic and contemporary views of the moral thinker.

1. INTRODUCTION

A system of morality which is based on relative emotional values is a mere illusion, a thoroughly vulgar conception which has nothing sound in it and nothing true.

Socrates as quoted in *Phaedo* (360 BCE)

Civilization's oldest written documents attest that evaluating and regulating moral conduct were among the earliest concerns of human social life (Haidt, 2008). And we have been fretting and fighting about it ever since.

Few topics inflame passions like issues of right and wrong, and few things drive our impressions of others more than moral virtues or moral failings. We respond viscerally to acts that affirm or offend our moral sensibilities. Acts of compassion inspire us, acts of self-sacrifice humble us, acts of injustice outrage us, and acts of indecency disgust us. Infuse any pragmatic question with moral significance and the emotional stakes are immediately raised (Skitka et al., 2005; Tetlock, 2003). A cocktail party conversation about the effectiveness of some government program will quickly bore all but your most wonkish acquaintances and send them scurrying for the bar. But question that program's morality, and your audience is likely to swell and, before long, voices are likely to rise. In short, despite the distaste expressed by Socrates (and many other moral philosophers) for mixing the two, morality and emotion are inextricably intertwined. Humans care deeply about right and wrong, and will go to extraordinary lengths to protect and promote their moral beliefs.

Because moral judgments are so frequently made against a backdrop of emotion, it seems obvious to ask how the passionate feelings we often have about moral issues affect our judgments about them. Interestingly, however, modern psychology's first foray into the study of moral reasoning paid little attention to the role of affect. Reacting in part to Freud's depiction of moral thinking as ensnared in a web of conflicted and largely unconscious motivational forces (Freud, 1923/1962), Kohlberg built on Piaget's cognitive-developmental infrastructure to depict moral reasoning as an essentially analytical and rational enterprise, that grew ever more so as an individual's reasoning abilities became more sophisticated (Kohlberg, 1969, 1984). This view of moral judgment as a reasoning task, albeit one often marred by faulty or simplistic logic, was in perfect sync with the cold, information processing perspective that was the driving force behind the "cognitive revolution" that set the agenda for psychological research throughout the last few decades of the twentieth century.

But psychology is now in the throes of an affective revolution (Forgas and Smith, 2003; Haidt, 2007). The dominant view in contemporary social cognitive research is one in which affect, intuition, and analytical thinking are all recognized as necessary characters in a comprehensive narrative of mental life, and considerable research attention is devoted to redressing the previous lack of attention both to affective forces as crucial determinants of how information is processed (Forgas, 1992; Kunda, 1990; Slovic et al., 2007), and to the importance of hedonic as well as judgmental outcomes as central topics of psychological science (Gilbert, 2006; Kahneman et al., 1999). Accordingly, understanding the role of affect in moral reasoning has now become a prominent focus of research in moral psychology (Greene et al., 2001; Haidt, 2001; Monin et al., 2007).

In this chapter, we will explore one aspect of the morality-affect interface that has not received extensive empirical or theoretical attention: how the motivated reasoning processes that have been well documented in other spheres of judgment affect thinking about moral issues. That is, our specific goal in this chapter is to examine how moral reasoning is perturbed when individuals have an affective stake in perceiving an act as moral or immoral. Given that moral beliefs are among the most strongly held beliefs we possess (Skitka et al., 2005), it should come as no surprise that people often have deep feelings about what and who is moral or immoral and that these feelings may affect how they think about moral issues. In the sections that follow, we will first lay out a general view of motivated reasoning processes and discuss why moral judgment should be particularly susceptible to their influence. We will then review research examining two general pathways by which motivational forces can alter the moral implications of an act: by affecting perceptions of an actor's moral accountability for the act, and by influencing the normative moral principles people rely on to evaluate the morality of the act. We conclude by discussing the implications of research on motivated moral reasoning for both classic and contemporary views of the moral thinker.

2. MOTIVATED REASONING

Passion and prejudice govern the world; only under the name of reason.

John Wesley from a letter to Joseph Benton (1770)

Imagine for a moment that you are in the Supreme Court of the United States. At the front of the high-ceilinged room, august in their long black robes and seated behind a tall and imposing bench, are the nine Justices. Facing them on opposite sides of the room are the two opposing attorneys and their supporting legal teams. Everyone in the room has a prescribed part to play.

The two teams of attorneys each have a conclusion that they want the Justices to reach, and their job is to muster any evidence or argument they can to support their preferred conclusion. In our adversarial system of justice, it is not up to the attorneys to present a balanced view of the strengths and weaknesses of their case, nor to fairly present the merits of the arguments they put forth. Their job is to advocate; to build a compelling case for whatever conclusion they have been handsomely paid to defend.

It is the Justices, of course, whose job it is to fairly, analytically, and objectively adjudicate “the truth.” As the embodiment of “blind” justice, their charge is to be free of a priori preferences for one conclusion over

another, or at least not allow any biases or preconceptions they do have to influence their evaluation of the merits of the case. Whereas attorneys accepted role is to work from a particular conclusion “backward” to construct a convincing foundation for their case, the Justices are to go about their work in precisely the opposite direction, working “forward” to combine fact patterns and legal principles in a way that leads them agnostically to whatever conclusion these facts and principles seem to demand. To use a different but more common spacial metaphor, attorneys can be top-down, but Justices, and justice more generally, must always be bottom-up.

The conflicting motives that characterize the highest courtroom of the United States provide a helpful metaphor to frame the nuances of social psychological research on motivated reasoning. In the world of everyday judgment, most of us recognized ourselves more in the visage of a Supreme Court Justice than a hired gun attorney. When faced with a decision like who should be hired for a position at our workplace, we perceive ourselves as proceeding with judicial objectivity. We look at each applicant’s qualifications, weight them by the criteria we believe best predict success at the job, and choose the applicant whose qualifications best exemplify the most important criteria. We understand the potential to be biased by irrelevancies like race, gender, or physical attractiveness, but we perceive ourselves as able to rise above their influence. After all, the facts on the resumes and their fit with our stated standards seem to us clear evidence for the impartial, bottom-up nature of our decision.

And yet, a wealth of social psychological research suggests that in many judgment situations, particularly those that involve people and issues we care about deeply, people act more like lay attorneys than lay judges (Baumeister and Newman, 1994). Although the motivation to be accurate can sometimes improve the quality of inferences (Kunda, 1990; Lerner and Tetlock, 1999), people (like attorneys) often have a preference for reaching one conclusion over another, and these directional motivations (Kunda, 1990) serve to tip judgment processes in favor of whatever conclusion is preferred. Much of this research examines responses to self-relevant feedback, showing that people tend to perceive information that supports their preferred images of themselves as smart, well-liked, and healthy as more valid than information that challenges these flattering self-conceptions (Ditto and Lopez, 1992; Ditto et al., 1998; Wyer and Frey, 1983). But judgments about other people and about the facts of the world have also been shown to be subject to motivational influence. People make more charitable attributions for the behavior of people they like than those they dislike (Reeder et al., 2005), and perceive evidence that supports cherished social attitudes as more valid and compelling than evidence that challenges those attitudes (Lord et al., 1979; MacCoun, 1998; Munro and Ditto, 1997).

But how do people accomplish this “sleight of mind”? How can people act like attorneys yet perceive themselves as judges?

First, it is important to recognize that maintaining this “illusion of objectivity” (Kunda, 1990; Pyszczynski and Greenberg, 1987) is essential if motivated reasoning processes are to affect genuine belief. Explicitly, most people, most of the time desire an accurate view of the world. We are naïve realists (Ross and Ward, 1996), who believe that truth exists and that our senses and intellect are the way that truth reveals itself. In matters of truth, of course, anything an attorney says is suspect because everyone knows they have both a professional duty and a financial stake in arguing for one particular conclusion. Attorneys are often reviled, in fact, because their motivation to construct a particular reality by the selective use of facts and strategic manipulation of reasoning seems naked and unrepentant. While most of us accept this kind of backward reasoning as a necessary evil of our adversarial system of justice, we reject it as a way to approach everyday decisions. If we approached our judgments like an attorney with an explicit goal of reaching a particular conclusion, or even if we recognized that our judgments could be biased by our preferences, their value to us as reflections of the true state of the world would be sorely compromised.

But even when an individual's conscious motivation is accuracy, one conclusion can still be preferred over another because it supports a desired view of self or others, or the validity of a cherished belief. In this case, we use the term “preference” not in the sense of an explicit judgment goal, but rather as a set of implicit affective contingencies that underlie how we process information related to the judgment. That is, we say that someone has a preference for a particular judgment conclusion when that person would be happier if that conclusion were true than if it were false. Consequently, as people consider information relevant to a judgment where they have a preferred conclusion, they experience positive affect if that information seems to support their preferred conclusion, and negative affect if it seems to challenge their preferred conclusion (Ditto et al., 2003; Munro and Ditto, 1997). These affective reactions are quick, automatic, and ubiquitous (Winkielman et al., 2005; Zajonc, 1980) and can exert a host of subtle organizing effects on the processing of preference-relevant information.

A number of studies have shown, for example, that people test more favorable hypotheses when considering preference-consistent than preference-inconsistent information (Dawson et al., 2002; Kunda, 1990; Pyszczynski and Greenberg, 1987), are more likely to perceive ambiguous information in preference-consistent rather than preference-inconsistent ways (Balcetis and Dunning, 2006), apply less rigorous judgmental standards to preference-consistent than preference-inconsistent information (Ditto and Lopez, 1992; Ditto et al., 1998, 2003), and weight most heavily general

decision criteria that are most consistent with preference-consistent conclusions (Dunning et al., 1995; Norton et al., 2004; Uhlmann and Cohen, 2005).

The crucial aspect of these motivated reasoning mechanisms is that their subtlety allows them to operate well within the confines of what people perceive as the dictates of objectivity. So returning to the job hiring scenario we described above, liking for applicant X (based let us say on an implicit preference to hire a male) may lead the individual in charge of the hiring decision to begin the evaluation process with the hypothesis that applicant X is a strong rather than a weak candidate for the position, to interpret vague aspects of applicant X's record in a favorable rather than an unfavorable light, to accept applicant X's positive qualifications at face value while carefully considering alternative explanations for the weak spots in his resume, and to weight prior work experience (of which applicant X has a lot) as a more important criterion for the position than a high-quality educational background (where applicant X's credentials are less stellar). In each case, the decision maker's preference for the male candidate operates implicitly to bend but not break normative rules of decision making, leaving behind little introspective evidence of an untoward decision process (Pronin, 2008). Instead, the decision maker is likely to perceive himself as carrying out the hiring process in a judicious, bottom-up fashion — testing hypotheses, scrutinizing evidence, and comparing qualifications to standards — failing to recognize that his implicit preference for a male employee subtly shaped his decision process from the top-down, and tipped the scales toward his preferred verdict.

3. MOTIVATED MORAL REASONING

All morality depends upon our sentiments; and when any action or quality of the mind pleases us after a certain manner we say it is virtuous. . .

David Hume from *A Treatise of Human Nature* (1739–1740)

Motivated reasoning has been examined in a wide range of judgment domains in which the motivation to reach a particular conclusion derives from a number of different sources. We use the term *motivated moral reasoning* to describe situations in which judgment is motivated by a desire to reach a particular moral conclusion. That is, in this chapter we are interested in situations in which an individual has an affective stake in perceiving a given act or person as either moral or immoral, and this preference alters reasoning processes in a way that adjusts moral assessments in line with the desired conclusion.

There are a number of characteristics of moral judgment that should make it particularly fertile ground for motivated reasoning processes. First,

moral judgment is deeply evaluative, so people should have strong preferences about whether certain acts and certain people are perceived as moral or immoral. The evaluative significance of moral judgment stems from its fundamentally social nature. Most contemporary treatments of moral judgment situate its development in the evolutionary pressures of group living (e.g., Gintis et al., 2005; Haidt, 2001). Coordinated group activity requires that we can trust others not to harm us, to deal with us fairly, and to generally act in the best interests of the larger ingroup (Fehr and Gächter, 2000; Haidt and Graham, 2007). It also requires others to trust us in a similar fashion. This social component of moral reasoning explains not only why we have a deep affective stake in wanting others to perceive us as actors of good moral character (and thus to perceive ourselves in similar fashion), but also why we care deeply about the moral character of others.

Moreover, as researchers have noted for years, people desire a world that makes good moral sense (e.g., Lerner and Miller, 1977). A just world is one that is predictable and controllable, a world that can be trusted to reward good people and good behavior (including our own). This again suggests that we should have a strong preference to believe in our own moral worthiness, but also that we should be motivated to believe that the people we like and the groups we identify with are good moral actors, while often assuming that people and groups we feel negatively about have moral qualities consistent with our overall negative feelings about them.

Second, moral judgment is inherently intuitive, and so should be particularly amenable to affective and motivational influence. An influential paper by Haidt (2001) argued that contrary to Kohlberg's rationalist view of moral judgment in which reasoning is thought to precede and determine moral evaluation, moral reasoning more typically follows from moral evaluations rather than precedes them. Building on the philosophy of Hume (1739–1740/1969) and the psychology of Zajonc (1980), Haidt argued that moral evaluations most typically arise through an intuitive, and generally affective, process. Certain acts just “feel” wrong to us, and this realization comes in a form more akin to aesthetic judgment (“Lima beans disgust me!”) than reasoned inference (“Having sex with one's sibling is wrong because it increases the probability of potentially dangerous recessive phenotypes if pregnancy should result”). Haidt's point was not to say that reasoned moral analysis never occurs or can never override intuitive moral reactions (Haidt, 2007; Pizarro and Bloom, 2003), but rather that in sharp contrast to the Kohlbergian view of moral judgment, the primary sources of our moral evaluations are relatively automatic and affective as opposed to thoughtful and cognitive.

From an intuitionist perspective, moral reasoning is, fundamentally, motivated reasoning. Rather than being driven from the bottom-up, by data or reason, moral judgments are most typically top-down affairs, with the individual generating moral arguments with intuitions about the

“correct” moral conclusion already firmly in place. These moral intuitions can derive from factors that most everyone would see as morally relevant, such as whether an act causes harm to others or violates basic principles of justice. But they can also derive from factors that are more controversially “moral” in nature, such as whether an act is perceived as disgusting (Haidt et al., 1993; Wheatley and Haidt, 2005), or even from factors that would consensually be viewed as irrelevant to a rational moral analysis, such as one’s liking for the actor or the actor’s social group, or how the morality of the act reflects on valued beliefs and attitudes held by the perceiver.

Finally, moral judgment is complex and multifaceted, and so motivational forces should have considerable latitude to perturb moral reasoning in ways that support affective preferences while still allowing individuals to maintain a veneer of objectivity. Part and parcel of the legal analogy we presented in the previous section is the notion that motivated reasoning processes are constrained by plausibility (Ditto and Lopez, 1992; Kunda, 1990; Pyszczynski and Greenberg, 1987). People only bend data and the laws of logic to the point that normative considerations challenge their view of themselves as fair and objective judges, and motivated reasoning effects are most pronounced in situations where plausibility constraints are loose and ambiguous (Ditto and Boardman, 1995; Dunning et al., 1995).

As we will expand upon in subsequent sections, moral evaluation of an act involves a complicated set of judgments that can be profitably parsed into two categories: those focused on the “actor,” with the goal of assessing his or her moral accountability for the act and its consequences, and those focused on the “act itself,” with the goal of evaluating how the act and its consequences conform with general normative principles regarding what constitutes moral and immoral behavior. Judgments of moral accountability involve inferences about the actor’s motives, intentions, and knowledge; internal states of mind that can only be inferred indirectly, and thus are seldom unambiguous. Similarly, despite centuries of philosophical bickering, little consensus exists regarding the appropriate normative principles by which an act’s morality should be judged, and individuals often have conflicting intuitions about which moral principle should prevail, especially in classic “moral dilemmas” in which seemingly unsavory acts also bring about positive moral consequences. What this means is that the morality of any given act is seldom self-evident, and is notoriously difficult to “prove” in any uncontroversial way with either data or reason (Sunstein, 2005). This ambiguity should leave a motivated moral judge considerable flexibility to construct plausible justifications for preferred moral conclusions without offending their sense of their own objectivity, either by adjusting perceptions of an actor’s accountability for a moral act, or by altering the principles brought to bear in evaluating the morality of the act itself.

We will take up these two different pathways, in turn and in detail, in [Sections 4 and 5](#).

4. MOTIVATED ASSESSMENTS OF MORAL ACCOUNTABILITY

Even a dog distinguishes between being stumbled over and kicked.

Oliver Wendell Holmes, Jr. from *The Common Law* (1881)

Many of our moral judgments are concerned with determining whether an act is morally permissible or impermissible — such as whether having an abortion, eating meat, or cheating on our federal taxes is wrong. But because most of us already possess a mental list of sins and virtues, much of our daily moral judgment ends up being about whether or not to hold a particular individual responsible for one of these moral breaches. Not only is this a common concern in daily life (when a slightly neurotic friend insults you, should you blame her or excuse her because of her personality quirks?), but we are also constantly faced with tough cases in the media and popular culture in which it is unclear whether or not an individual should be held responsible for a moral transgression. Take one distressing example: should Andrea Yates, the mother who drowned her five young children in a bathtub while suffering from severe depression, have been held responsible for murder, or simply institutionalized for mental illness? We go to great lengths to try to arrive at the right answer to these questions. A wrong answer, after all, could mean punishing a person who simply does not deserve it (or failing to punish someone who does).

Accordingly, a pressing concern for psychologists, legal theorists, and philosophers alike has been to specify how we *ought* to arrive at a judgment that an individual should be held fully responsible for an act. Across these disciplines, theorists have posited a variety of necessary conditions for the ascription of responsibility (Aristotle, 1998; Kant, 1795/1998; Shaver, 1985; Weiner, 1995). The most influential psychological theories of moral responsibility suggest that for an individual to be held fully responsible for an act, that act should have been caused, controllable, and intended by the actor (Shaver, 1985; Weiner, 1995).

While psychological theories of responsibility draw heavily from normative philosophical and legal theories, they also often assume that everyday moral decision makers actually adhere to the prescriptions of these normative theories. As it turns out, there is a great deal of evidence suggesting that we often do go about making judgments of responsibility in such a careful fashion. If an individual is believed to have contracted HIV because of a blood transfusion, for instance, people are less likely to blame the individual for his plight than if he contracted the disease through licentious sex (a presumably controllable act; Weiner, 1995). Similarly, individuals make fine-grained distinctions about causality in determining whether or not to

hold an individual responsible. In fact, even when a negative act was clearly caused by an agent, people do not hold the agent as responsible for the act if the causal chain failed to proceed in the precisely intended manner (cases of so-called “causal deviance”; Pizarro et al., 2003a).

Finally, individuals are very sensitive to the presence of intentionality — whether or not a bad action was performed purposefully. When there is reason to believe that the ability to formulate intentions might have been compromised (such as in the case of some forms of mental illness), individuals reduce judgments of responsibility. For instance, relatives of schizophrenic patients reduce blame for negative (and even extremely harmful) acts caused by the presence of delusions and hallucinations (although less so for negative acts stemming from negative symptoms such as flattened affect; Provencher and Fincham, 2000). Even in cases of emotional impulsivity among presumably normal individuals, if a strong emotional impulse (such as extreme anger) leads to a negative act, individuals discount blame for that act (Pizarro et al., 2003b). Presumably, most people’s lay theories of emotion hold that under the influence of strong feelings, one loses the ability to formulate intentions and act upon them.

And yet, despite the evidence that individuals are capable of making sensitive, objective distinctions among the features of various acts to arrive at a judgment of blame, there is also good evidence that this process can be distorted by motivational factors. Indeed, a fairly clear picture is emerging that judgments that an individual is “bad” or “good” often come *prior* rather than *as a product of* more fine-grained judgments of intentionality, controllability, and causality, and these component assessments are then “bent” in a direction consistent with the overall moral evaluation. It is to this evidence that we now turn.

4.1. Controllability and Culpability

One set of findings that demonstrates the role that motivation can play in judgments of responsibility are those of Alicke (1992) showing that individuals are more likely to judge that an individual possesses causal control over an outcome if they are motivated to blame that individual. In one example, when participants were told that a man was speeding home in a rainstorm and got in an accident (injuring others), they were more likely to say that he had control over the car if he was speeding home to hide cocaine from his parents than if he was speeding home to hide their anniversary gift (Alicke, 1992). Alicke (1992, 2000) argues that spontaneous judgments of blame lead participants to distort judgments of control and causality in order to justify the initial blame — a process he refers to as “blame validation.” According to this blame validation model, the spontaneous assessment of blame made when first evaluating a man who would be despicable enough to hide drugs

in his parents house, leads us to distort the evidence in the direction of finding him more culpable for the accident.

4.2. Memory for the “Facts”

Judgments of control and causality are inherently subjective, making them excellent candidates for motivated distortion. But there is evidence that even more “objective” facts about an event can be distorted in the right direction, given sufficient time. Pizarro and colleagues (Pizarro et al., 2006) demonstrated that the degree of blame given to an individual could influence memory for the facts of the blameworthy act itself. Participants were given a description of a man who dined by himself at a fine restaurant one night, and then finished his meal by walking out on the check. One group of participants (the low-blame condition) was told that he had received a phone call that his daughter had been in an accident and left the restaurant in such a hurry that he forgot to pay his bill. Another group (the high-blame condition) was told that the reason the man walked out on his check was that he disliked paying for things and was happy whenever he could get away with stealing.

Approximately one week later, participants were asked to recall the price of the items the man had purchased for dinner, as well as the price of the total bill (which had been explicitly stated in the original scenario). Those in the high-blame condition recalled the values as significantly higher than those in the low-blame condition, and significantly higher than they actually were. Those in the low-blame condition, by comparison, had, on average, an accurate memory of the prices a week later. In addition, the degree of blame participants gave the man initially upon reading the description of the act was a significant predictor of the magnitude of inflation in memory a week later. These findings provide initial evidence that the motivation to justify blameworthiness may even lead to a distortion of the objective facts surrounding an event.

4.3. Intentionality

Perhaps the most compelling set of evidence that motivations can shift judgments of moral responsibility comes from recent work on judgments of intentionality. In an influential set of findings by Knobe and his colleagues, it has been shown that people are more inclined to say that a behavior was performed intentionally when they regard that behavior as morally wrong (Leslie et al., 2006; see Knobe, 2006, for a review). For instance, when given a scenario in which a foreseeable side effect results in a negative outcome, individuals are more likely to say that the side effect was brought about intentionally than if the side effect results in a positive outcome. In the most common example, the CEO of a company is told that implementing a

new policy will have the side effect of either harming or helping the environment. In both cases, the CEO explicitly states that he only cares about increasing profits, not about the incidental side effect of harming or helping the environment. Nonetheless, participants perceive that the side effect of harming the environment was intentional — but not the side effect of helping the environment. This pattern of findings (with simpler scenarios) is evident in children as young as six and seven years old. One plausible account for this pattern of findings, consistent with Alicke's "blame validation" approach, is that the motivation to blame the bad guy — someone who would be capable of such an immoral act — directly leads to a distortion in judgments of intentionality.

If this were indeed the case, it should be possible to show that individuals with divergent moral views would perceive acts (whose moral status they disagreed upon) as more or less intentional depending on their beliefs about that particular act. Such findings would provide evidence that the so-called "side-effect effect" is not simply due to the nonconventional nature of the infractions (i.e., since most people obey rules and norms most of the time a harmful or infrequent act might be inferred to be more intentional simply because it deviates from modal behavior; [Kahneman and Miller, 1986](#)).

Accordingly, we sought to provide evidence that an individual's moral preferences could influence the very criteria for what is thought to constitute intentional action ([Tannenbaum et al., 2008](#)). In order to test this hypothesis, we first looked at individuals who had protected values within a certain domain, compared to individuals who did not. By protected values, we mean absolute beliefs about the impermissibility of certain acts that serve as a strong form of moral motivation to prevent any moral breach. Protected values are thought to be both nonfungible and to motivate behavior in numerous ways ([Baron and Spranca, 1997](#); [Skitka et al., 2005](#); [Tetlock, 2003](#)). For example, if an individual has a protected value regarding the sanctity of animal life, that individual would be very reluctant to trade-off the lives of animals in exchange for another good (food for the poor), or even to agree that cutting down on meat consumption is a virtue (it would seem about as reasonable as a pedophile cutting down on the number of children he molests).

In our initial experiment, we presented a large group of subjects with two versions of the scenario originally used by [Knobe \(2003\)](#), in which a corporate executive decides to undertake a business venture that will bring him a direct profit, but which also has the side effect of (depending on scenario version) either harming or helping the environment. Afterwards, we examined the degree to which participants were willing or unwilling to make trade-offs on the environment ([Baron and Spranca, 1997](#); [Tanner and Medin, 2004](#)), as well as some straightforward attitude items about the relative priority for the environment over the economy. Using these items, we identified a group of participants who expressed both a preference

for the environment and also a complete unwillingness to make trade-offs on the environment, even if doing so led to a more beneficial outcome (environmental absolutists). For our control group, we selected individuals who expressed a similar preference for the environment, but were willing to make trade-offs on the environment if doing so led to some beneficial outcome (nonabsolutists). Because environmental absolutists placed greater moral worth on outcomes that affect the environment, we expected that (compared to nonabsolutists) they would be more likely to judge the CEO's decision as intentionally harming the environment in the harm condition, but less likely to see the CEO's decision as intentionally helping the environment in the help condition.

And this was indeed the case. Overall, we replicated the basic asymmetry in intentionality judgments documented by [Knobe \(2003\)](#): 75% of our subjects reported that the CEO intentionally harmed the environment, but only 11% reported that the CEO intentionally helped the environment. But we also found that this asymmetry was greater for environmental absolutists compared to nonabsolutists: 81% of absolutists saw the harm to the environment as intentional compared to 62% of nonabsolutists, and only 7% of absolutists saw the help to the environment as intentional compared to 21% of nonabsolutists. This pattern supports a motivational account of the side-effect effect, in that individuals who presumably had particularly strong evaluative reactions to harming or helping the environment also showed an exacerbated effect across conditions.

Even more interestingly, we asked all participants to justify why they believed the action to be intentional or unintentional, and a clear pattern emerged. Of the subjects who reported that the CEO acted intentionally, they quite often (61%) made reference to the CEO's foreknowledge of the policy's side effects (e.g., "I believe it was intentional because he knew ahead of time that it was going to harm the environment yet he chose to do it anyways"). Conversely, subjects who reported that the CEO acted unintentionally almost always (99%) focused on goal-directed mental states (e.g., "The chairman said that he did not care about the environment. He was only motivated by monetary gain"). That is, participants seemed to be using two different definitions of intentionality, one strict (requiring that the CEO desire the effects for them to be considered intentional) and one loose (requiring only that the CEO have foreknowledge that the side effects would occur). Both of these definitions are plausible (in fact, they generally correspond with two basic legal standards of criminal culpability; direct and oblique intent; [Duff, 1990](#)), and it may have been this conceptual flexibility that allowed participants to define "intentional" in a way that fit their desire to blame or not blame the CEO for his actions.

In a follow-up study, we replicated this finding using protected values concerning the economy (a value that cut quite differently across liberal-conservative lines). We presented subjects with a scenario modeled closely

on [Knobe \(2003\)](#), but this time describing the chairman of an environmental protection organization whose decision to preserve the environment also had the effect of harming or helping the economy (in the form of increasing or decreasing unemployment, respectively). We once again selected only those who expressed a genuine preference for the economy (when pitted against the environment), and who also expressed an unwillingness to make trade-offs on economic welfare (market absolutists) versus those who expressed a willingness to make trade-offs if doing so promoted other beneficial outcomes (nonabsolutists). Consistent with the results of our first study, 75% of market absolutists reported that the chairman intentionally harmed the economy, whereas only 44% of nonabsolutists reported the harm as intentional. Moreover, only 16% of market absolutists reported that the environments intentionally helped the economy, whereas 28% of non-absolutists reported such help to the economy as intentional. Participants again justified their judgments by using convenient definitions of “intentional.” When reporting that the chairman unintentionally harmed or helped the economy, almost all participants (96%) focused on the agent’s goal-directed mental states (i.e., his desire to merely preserve the environment). When reporting that the chairman intentionality harmed or helped the economy, on the other hand, they typically (74%) focused on the chairman’s foreknowledge of the effect it would have on the economy.

In a final follow-up study, we sought to investigate attributions of intentionality in a more realistic context. A controversial moral issue is whether collateral damage in the conduct of war is permissible. Is it morally wrong, for example, for the U.S. military to bomb a village where it is believed a suspected terrorist is hiding, even though it will also result in the death of innocent civilians? Proponents of such military strikes note that the collateral damage caused to innocent civilians is not an act of intentional harm — the goal is to strike at the enemy, and any harm to innocent civilians is both unwanted and unfortunate. Opponents of such military strikes might argue, however, that collateral damage to innocent civilians although not desired is clearly foreseeable, and thus such decisions knowingly (and by their definition, intentionally) bring about harm to innocent civilians. Given that the issue of collateral damage contains the right conceptual structure of knowing but not wanting bad outcomes when making decisions with moral gravity, we chose to utilize it in an experimental context.

We presented participants with one of two military scenarios. Half of the participants received a scenario describing American military leaders deciding to carry out an attack to stop key Iraqi insurgent leaders in order to prevent the future deaths of American troops. The other half read about Iraqi insurgent leaders deciding to carry out an attack to stop key leaders of the American military in order to prevent future deaths of Iraqi insurgents. In both cases, it was explicitly stated that the attackers (whether American or

Iraqi) did not want nor intend to cause civilian casualties, but in both cases the attack did. After reading the scenario, participants were asked to indicate whether the military leaders had intentionally harmed the innocent civilians, and to indicate their political ideology on a standard 7-point liberal-conservative scale.

Our motivational prediction was simple. We suspected that political conservatives would be more motivated than political liberals to view American military action in a positive light, especially in comparison to the action of Iraqi insurgents. For many conservatives, patriotism in general, and support for the American military in particular, take on the quality of protected values. This is consistent with research suggesting the conservative's moral judgments are more influenced than those of liberals by issues of ingroup loyalty (Haidt and Graham, 2007), and suggests that conservatives should be more likely than liberals to make a moral distinction between the acts and lives of Americans and those of a disliked outgroup like Iraqi insurgents.

Consistent with this analysis, the study showed that liberal participants showed no significant difference in their intentionality judgments depending on the nationality of the perpetrators. Conservative participants, however, were significantly more likely to see the killing of innocent civilians as intentional when the deaths were caused by Iraqi insurgents than when they were caused by the American military. Moreover, we once again found a very similar pattern in how participants justified their intentionality judgments. When subjects reported the harm to innocent civilians as intentional, they typically (77%) justified their response by focusing on the military leader's foreknowledge of harm to the innocent civilians. When subjects saw the harm as unintentional, however, they almost always (95%) made reference to the agent's goal-directed mental states — that they were only trying to target key enemy leaders.

To recap, our studies of perceived intentionality suggest that people with divergent moral values are likely to make different intentionality judgments. If an action poses little affront to our moral sensibilities, we tend to think of intentional behaviors as only those that are directly intended or desired. Foreseeability is not enough. The more we view an act as morally offensive, however, the more we seem to loosen our criteria for intentionality to include actions that may not have been desired but should have been foreseen. Utilizing differing definitions of intention is a powerful way of maintaining an illusion of objectivity. In fact, this analysis suggests that two people could view the same act in exactly the same way — they could agree on what the actor did and did not desire and what the actor did and did not foresee — and nevertheless disagree on whether that act was performed intentionally, simply because their differing moral responses to the act led them to think about intentional behavior in qualitatively different ways.

5. MOTIVATED USE OF MORAL PRINCIPLES

You can't be the President unless you have a firm set of principles to guide you as you sort through all the problems the world faces.

George W. Bush during a Presidential Press Conference
(December 20, 2007)

Politicians are famous for portraying themselves as men and women of principle, and infamous for adjusting their principles for the sake of political expediency. But while politicians may be the most shameless and visible manipulators of principle, there is good reason to believe that they are not alone in adopting a flexible approach to principled argumentation.

In [Section 4](#), we illustrated how motivational factors can influence perceptions of an actor's moral responsibility for his or her actions. To the extent that an action provokes negative feelings in the perceiver, that perceiver is likely to interpret the act as more controllable, more intentional, and with time may even have a more negative memory of the act itself. But moral evaluation does not stop at moral accountability. Even when the evidence is clear that someone is morally responsible for an action, assessment of the morality of the act still requires the application of normative principles. This is most clearly illustrated when an act has both negative and positive consequences, as in the classic moral dilemmas that have captured the imagination of both philosophers and psychologists for generations ([Foot, 1967](#); [Greene et al., 2001](#); [Kohlberg, 1969](#); [Thompson, 1986](#)). The iconic example here is the so-called "trolley problem" in which one has to decide whether it is morally justified to sacrifice the life of one individual to stop a runaway trolley car that will otherwise kill five. How one resolves this dilemma is generally accepted to depend on whether one endorses a *deontological* ethic, in which certain acts are thought to be "wrong" in and of themselves and no matter their consequences, or a *consequentialist* ethic, in which an act's morality is judged solely based on the extent to which it maximizes positive consequences. An individual relying on deontological principles should conclude that killing the innocent individual is morally wrong even though the act would save five others, whereas an individual relying on consequentialist principles should reach the opposite conclusion; that killing the one, while obviously unfortunate and regrettable, is morally justified based on the net gain in positive consequences that results from saving the other five.

5.1. The Use and Abuse of Principle

Principle-based reasoning plays an important role in moral judgment because moral reasoning occupies a peculiar middle ground between aesthetic preference and fact-based inference. Although people sometimes

press us for the reasons *why* we like a painting or hate lima beans, ultimately, that we simply like or hate with no articulable reason is acceptable when it comes to matters of personal taste. Because moral judgments, like aesthetic ones, often come to us intuitively, people frequently find themselves similar dumbstruck when asked to justify why a given act is thought to be right or wrong (Haidt, 2001; Haidt et al., 1993). But the pressure to provide a rationale for moral assessments is stronger than it is for aesthetic preferences because most of us see morality as something more than just a matter of taste. Our tendency toward naïve realism extends into the moral realm as well (Goodwin and Darley, 2008). While people are certainly capable under some circumstances of accepting a position of moral relativism, most people most of the time have a strong intuitive sense that moral claims, like factual ones, are either true or false. This naïve moral realism consequently leads us to feel that moral claims require some rational basis to establish their validity. If a moral claim is to be taken as universally “true,” rather than a mere personal preference or social convention, some reasonable justification for the claim must be provided.

But how does one justify a moral belief? Factual beliefs (e.g., cigarette smoking causes cancer) can be supported by data, but what data can be mustered to prove one’s assertion that terrorism is immoral or that a white lie is morally justified if it leads to a greater good?

What often takes the place of data-based inference in moral reasoning is grounding a specific moral belief as an instantiation of a general moral principle. Principles can be understood as foundational rules that, while not always fully universal, are at least widely applicable across a defined set of situations. Reasoning one’s way from a general moral principle to a specific moral belief allows moral reasoning (e.g., I believe terrorism is immoral *because* it is immoral to deliberately sacrifice innocent life even for a greater good) to take a form very much like fact-based inference (e.g., I believe smoking causes cancer *because* epidemiological research provides overwhelming support for the link). Kohlberg (1984), of course, viewed principle-based reasoning as the hallmark of mature moral judgment, and most moral philosophers would recognize it as the primary way they go about their business.

Fundamental to principle-based reasoning is the idea that principles are general rather than case-specific, and that they should not be applied (or ignored) selectively. The power of principles as explanatory mechanisms derives precisely from their generality, but this is also what makes them so darn inconvenient. The same general rule that can provide justification for a desirable course of action in one case will often compel a less palatable course in another. If one relies on principle only when it is convenient, however, the door is opened to charges of hypocrisy or casuistry, and the normative status of the principle as a justification for the validity of any specific moral claim is correspondingly weakened.

Nonetheless, there is good evidence from outside the realm of moral reasoning that people use general judgment standards in just this kind of selective fashion. Dunning and colleagues (Dunning and Cohen, 1992; Dunning et al., 1995), for example, showed that if people are asked to identify general criteria of excellence in a given domain, they typically endorse standards that put their own idiosyncratic credentials in the best possible light. Studies examining mock hiring and legal decisions have similarly shown that evaluators tend to inflate the value of general decision criteria (e.g., the importance of particular job credentials or types of legal evidence) that tend to favor preferred conclusions (Norton et al., 2004; Uhlmann and Cohen, 2005; Simon et al., 2004a,b).

Somewhat closer to the domain of moral reasoning, a recent series of studies by Furgeson et al. (2008, in press) suggest that reliance on principles of constitutional interpretation can be affected by one's preferred legal conclusion. The irony of our use of Supreme Court Justices as icons of objectivity is that legal scholars have long noted the tendency for political ideology to influence even this highest level of judicial reasoning (e.g., Bork, 1990; Brennan, 1990). While judges like to couch their specific judicial decisions as guided by broad constitutional principles (such as originalism or expansive interpretation) it seems frequently the case that principles are favored or ignored depending on their fit with politically palatable conclusions. The classic anecdotal example is the Supreme Court's 2000 decision in *Bush v. Gore* (Dershowitz, 2001). In brief, the essential decision in that case concerned whether to let stand the decision of the Florida State Supreme Court to allow vote recounting to continue (knowing that if recounting was stopped, George W. Bush would almost inevitably be awarded Florida's electoral votes and consequently the Presidency of the United States). Interestingly, the five most conservative Justices, whose previous court decisions frequently favored state-sovereignty over federal intervention, decided in this case that it was appropriate to overturn the Florida State Supreme Court's ruling, while the four more liberal and historically more federalism friendly Justices favored allowing the state court's ruling to stand. This interpretation of the Justices's reasoning is obviously speculative and unsurprisingly controversial (Dionne and Kristol, 2001), but Furgeson et al. (2008, in press) have demonstrated just this sort of politically motivated reliance on constitutional principle in experiments using both college undergraduates and law students.

Despite its dubious normative status, selective reliance on general standards is likely a particularly common and particularly effective form of motivated reasoning, first, because in most domains more than one standard can plausibly be seen as appropriate, and second, because motivated reliance on a particular standard is difficult to detect unless the individual is confronted with multiple cases where the affective implications of reliance on the standard conflict. In any single case, the preferential weighting of

judgment standards occurs through a subtle, intuitive (and likely bidirectional; [Simon et al., 2004a,b](#)) process. At an explicit level, individuals should thus experience little that would lead them to question the objectivity of their reliance on one particular (and seemingly plausible) standard. It is only if an individual is quickly confronted with a second case in which the same general standard obviously compels a more objectionable conclusion that awareness of conflict should arise, and the individual will be forced to struggle with the normative implications of their decision-making strategy.

5.2. Tales of Political Casuistry

Like almost all judgmental biases, however, selective reliance on principle is more easily recognized by observers than actors ([Pronin, 2008](#)), especially when it is motivated by preferences that the observer does not share. And in fact, our initial inspiration to study the motivated use of moral principles arose from our observations of moral rationales offered by politicians in support of politically conservative policy positions. For example, many political conservatives (including President Bush) have staked out a principled deontological stand in opposition to government support for embryonic stem cell research, arguing that the potential lives saved by any technology generated by this research does not justify the sacrificing of innocent fetal life. Their moral assessment of the extensive civilian death toll caused by the invasion of Iraq, however, has been decidedly more consequentialist in tone, suggesting that in this case the sacrificing of innocent life is a necessary cost to achieve a greater good. More generally, although political conservatism is often seen as a bastion of deontological thinking (involving lots of uncompromising rules about the moral permissibility of various sexual and reproductive behaviors in particular), its recent approach to issues of national security (e.g., the use of electronic surveillance and harsh interrogation techniques as means of combating terrorism) is firmly rooted in consequentialist logic.

Of course, it should not take long for a truly objective observer to recognize that on virtually every issue mentioned above where conservatism suggests a deontological position, liberals swing consequentialist, and on the issues where conservatives adopt a consequential position, liberalism favors a less-forgiving deontological stance. Although the intuitive “triggers” that engage motivation for liberals and conservatives may differ ([Haidt and Graham, 2007](#)), there is little reason to believe that any particular political ideology is more or less conducive to motivated reasoning ([Munro and Ditto, 1997](#); but see [Jost et al., 2003](#) for a different view). An equally obvious limitation of our informal observations is that real-life anecdotes are a weak basis on which to draw inferences about motivated inconsistency. Issues like stem cell research and collateral war casualties, although arguably comparable in general moral structure, differ in

numerous subtle and not so subtle ways that could form a legitimate basis for what might superficially seem to be inconsistent and motivated moral assessments.

Accordingly, we have now conducted a number of controlled laboratory studies to examine the motivated use of moral principles by comparing the judgments of political liberals and conservatives to scenarios that we believed would invoke in them differing moral intuitions (Uhlmann et al., 2008). There are two things to note before we describe these studies. First, the studies all use political ideology as their motivational source, but this does not mean that we believe the effects are limited to political judgments. Rather, it simply reflects our personal interest in the intersection of political and moral reasoning and our belief that a political context provides both sufficient motivation and testable predictions for examining selective principle use. Second, the studies all examine scenarios in which consequentialist principles are pitted against deontological ones, but this should again not be taken to indicate that any effects are necessarily limited to this context. We chose to utilize the distinction between deontological and consequentialist reasoning because of the keen interest it has received in both moral philosophy and moral psychology (Greene, 2007), and because many moral dilemmas (both in the laboratory and in the real world) present individuals with conflicting choices of action based on these two moral ethics. Moreover, there is good reason to believe that people harbor intuitions consistent with both views, in that relatively subtle manipulations can lead people to shift from near consensual endorsement of deontological action (e.g., when a man must be pushed onto the tracks to stop an oncoming trolley from killing five innocent workmen) to near consensual endorsement of consequentialist action (e.g., when the trolley can be redirected to kill one person rather than five by simply flipping a switch; Greene et al., 2001). Access to multiple plausible intuitions is conducive to motivated reasoning as it allows people to comfortably draw upon whichever set of principles seems to best justify the moral conclusion they find most emotionally satisfying, without also challenging their view of themselves as logical and well-meaning moralists.

5.3. Trolleys, Lifeboats, and Collateral Damage

In the first study we conducted, college students were presented with a modified version of the trolley/footbridge dilemma, in which the morality of pushing one man into the tracks to save the lives of many others must be assessed. Our key modification was to include in the scenario extraneous information that we believed would evoke differing affective reactions depending on students' political ideology. There is a strong disdain among American college students in general, but among politically liberal Americans in particular, for harboring feelings that may be considered prejudiced

(Monin and Miller, 2001; Norton et al., 2004; Plant and Devine, 1998; Tetlock, 2003). We therefore decided to vary the race of the characters in the trolley scenario (subtly and between-subjects) to see whether this would influence participants' judgments concerning the appropriate moral action. Specifically, half of the participants were faced with a decision about whether to push a man named "Tyrone Payton" onto the tracks to save "100 members of the New York Philharmonic," while the other half had to decide whether to push a man named "Chip Ellsworth III" onto the tracks to save "100 members of the Harlem Jazz Orchestra." It should be clear that our goal was to lead our subjects, without using actual racial labels, to infer that in the first case their decision involved whether to sacrifice one African-American life to save 100 that were mostly White, and in the second case whether to sacrifice one White life to save 100 that were mostly African-American (our scenario used a larger than typical number of people to be saved in order to minimize floor effects caused by people's general reluctance to endorse the consequentialist action in the footbridge dilemma). After reading the scenarios, participants completed a series of scales assessing their beliefs about whether sacrificing Chip/Tyrone was the morally appropriate course of action, and measuring their endorsement of consequentialism as a general moral principle (e.g., "It is sometimes necessary to allow the death of an innocent person in order to save a larger number of innocent people"). In this study, as in all of the others in this series, these items were combined into an overall consequentialism index, and participants also indicated their political ideology on a standard liberal-conservative scale.

The results revealed that, as predicted, our race manipulation significantly affected participant's moral judgments. First, there was a general tendency for participants presented with the scenario in which Tyrone Payton was sacrificed to save 100 members of the New York Philharmonic to be less likely to endorse consequentialism than participants who read the version in which Chip Ellsworth III was sacrificed to save members of the Harlem Jazz Orchestra. In other words, participants were generally more likely to invoke a consequentialist justification for sacrificing a man with a stereotypically White American name than one with a stereotypically Black American name. When participant's political orientation was entered into the regression, however, a significant interaction effect was found. The tendency to view Chip's life in more consequentialist terms than Tyrone's was limited to political liberals. Political conservatives showed no hint of this effect.

The problem with this study, despite its intriguing results, is that it raises as many questions as it answers. Were the effects driven by the race of the individual being sacrificed or of the people being saved? Did people really draw inferences about race from the stereotypically Black and White names we used? Why did only liberals show changes in their use of moral

principles and is this pattern due to the relatively unique (and particularly liberal) political perspective of a college student sample?

To address these questions, we replicated the study, this time using the “lifeboat” dilemma in which a decision has to be made about whether or not to throw a severely injured person off of a crowded lifeboat in order to prevent the drowning of all of the others aboard. As in the first study, the injured person was named Tyrone Payton for half of the participants, Chip Ellsworth III for the other half. No information was provided about the race of those who would be saved. After participants completed the key dependent measures, they were asked to guess the race of Chip or Tyrone. The data were collected in an outdoor shopping mall. The average participant was 37 years old and the sample had a better mix of liberals and conservatives than the college student participants in our first study.

Participants’ moral responses in this study were very similar to those found in the first. Although in this study there was no overall effect for Tyrone to evoke less consequentialism than Chip, a significant interaction pattern was found of the same form as that found in our trolley study. When considering the morality of sacrificing an injured man to save the other occupants of a lifeboat, liberals were more likely to evoke a consequentialist justification for sacrificing the man if he was named Chip than Tyrone. Conservatives, on the other hand, although seeming to show something of the opposite tendency, revealed no statistical difference in their moral responses to the Chip and Tyrone scenarios. When asked to guess the race of the injured man, 79% of those in the Chip condition believed that Chip was White, and 64% of those in the Tyrone condition believed that Tyrone was Black. We suspect that the percentage assuming Tyrone was Black was suppressed by a general political correctness bias (people believing that it is wrong to guess someone’s race based solely on an ethnic sounding name), but nonetheless, the pattern of results is identical if only participants making the desired racial inferences are examined.

Taken together then, the results of our Chip and Tyrone studies show good evidence of motivated recruitment of moral principles, at least among political liberals. But why were the effects limited to our liberal participants (in two different studies using two different moral dilemmas and two different study samples)? Our speculation is that egalitarian considerations, especially those relevant to race, play a greater role in influencing liberals’ judgments compared to conservatives. A recent meta-analysis by [Jost et al. \(2003\)](#) indicates that one of the fundamental differences between liberals and conservatives lies in conservative’s greater tolerance for social inequality. Research on the moral foundations underlying liberal and conservative ideologies also suggests that fairness concerns are particularly acute for political liberals ([Haidt and Graham, 2007](#)), and race is likely the key symbol evoking these concerns in contemporary America. As such, we believe that

this particular situation simply held more motivational power for liberals than conservatives. Our Chip–Tyrone manipulation faced liberals with choices sure to alert their sensitivity to inequality, and they likely felt more negative affect when asked to sacrifice a Black life than a White life (especially a White person with a vaguely aristocratic-sounding name). Conservatives, on the other hand, not overtly prejudiced but simply lacking liberals' highly accessible intuitions regarding inequality, tended to respond in a more evenhanded fashion (both affectively and cognitively). This pattern is consistent with a number of recent studies (e.g., Norton et al., 2004) showing that college student samples (which are often skewed liberal) tend to show what might be called a “political correctness” bias in racial issues.

We were confident, however, that other moral dilemmas could be found that would be more likely to push conservative's motivational buttons. We also wanted to examine our ideas in the context of a moral issue that held more real-life import than those involved in the fanciful world of trolleys and lifeboats. So, we chose to return to the military scenarios we had previously used to examine attributions of intention.

In addition to involving questions about intentionality, collateral damage scenarios pit deontological and consequentialist ethics against one another in much the same way as the trolley or lifeboat dilemmas. Is the possibility (or certainty) of innocent civilians inadvertently being killed by a military operation justified by any greater good that might be achieved by military victory? This situation has clear real-world relevance, particularly in the context of U.S. involvement in military operations in Afghanistan and Iraq, and as we saw in our intentionality study, tends to evoke particularly strong moral intuitions among political conservatives.

We thus presented a sample of undergraduates with our collateral damage scenarios, one involving American military and the other involving Iraqi insurgents, and asked them questions measuring their tendency to offer consequentialist principles as justification for civilian casualties. Perhaps not surprisingly, there was an overall tendency for conservatives to take a more permissive (i.e., consequentialist) view of collateral military damage than did liberals. In addition, however, conservatives endorsed more consequentialist justifications for American-caused casualties than Iraqi-caused casualties. Liberals, on the other hand, showed a nonsignificant trend in the opposite direction.

Importantly, this study confirms that both liberals and conservatives engage in the motivated recruitment of moral principles. It also reveals, however, the moderated nature of the phenomenon. For motivated reasoning effects to occur, information must evoke significant affective stakes, and in the domain of moral reasoning, this requires an understanding of the intuitive infrastructure underlying moral judgment, and how moral intuitions may vary across individuals and groups.

5.4. Is Reliance on Moral Principles Really Motivated?

One of the constants of research on motivated reasoning is the challenge of ruling out cognitive counterexplanations (Ditto, in press; Tetlock and Levi, 1982). It is almost always possible to construct an amotivational and at least semirational explanation for putatively motivationally driven effects, usually based on a subtle informational confound buried in the motivational manipulation (Erdeylyi, 1974; Miller and Ross, 1975). The current situation is no exception. In this case, it might be argued that the seeming inconsistency in our participants' reliance on moral principles is not really inconsistency at all because our participants actually view the "extraneous" information we provided (about the race and nationality of individuals involved in the scenarios) as morally relevant. That is, perhaps our participants are not recruiting general normative principles via an implicit motivational process, but rather are simply naïve particularists (Dancy, 1993, 2004) who are cognizant of and comfortable with the fact that they are using different moral rules to evaluate Black and White or American and Iraqi lives.

There are three good reasons, however, to reject this interpretation of our results. First, our participants explicitly denied that race and nationality were relevant factors in their decisions. When asked at the end of our lifeboat study, 92% of participants in both the Chip and Tyrone conditions believed that their responses would not have been different if the target person was of a different race. In another study, we asked 238 students to evaluate the relevance of a host of factors in making life-or-death moral decisions. Participants overwhelmingly rated both race (87%) and nationality (87%) as morally irrelevant factors. Importantly, they did not hold the belief that *all* contextual factors were irrelevant to such life-or-death decisions. For example, 49% indicated that the health of the person was relevant and 62% thought the potential victim's age was relevant. Participants' political orientation was not reliably associated with their responses to any of these questions.

Second, if our participants believed that differential reliance on consequentialist versus deontological principles based on race or nationality was defensible, then they should report differing responses across our experimental scenarios if given both scenarios together (just as they gave differing responses when this information was manipulated between subjects). If, however, participants believe that these factors hold no moral weight, and that the moral principles they are invoking are normatively invariant, then their responses to the first scenario they receive should be an effective predictor of their responses to a second. In fact, when we did just this kind of within-subjects version of the Chip and Tyrone study, the pattern of responses clearly showed this latter pattern. For their first scenario, half

of the participants got the Chip version and half got the Tyrone version. The pattern seen in the original between-subjects study was replicated. Liberals gave more consequentialist justifications for sacrificing Chip than Tyrone, while conservatives showed little difference. When participants then received the alternative scenario to evaluate immediately afterward, their responses remained remarkably consistent (in fact, the correlation between the consequentialism index scores in the two scenarios was .98!). This effect produced the most striking pattern for liberals, as it led to a complete reversal of their initial bias. Whereas liberals were more consequentialist toward Chip than Tyrone in the first scenario, they were more consequentialistic toward Tyrone than Chip in the second. Participants seemed to perceive a strong constraint to remain consistent in their use of moral principles across the two scenarios, even when their initial choice of principle was evoked by motivational factors. This “carry-over” pattern again supports the motivational account of our effects, and suggests that the moral principles our participants were in fact using selectively, were ones that they believed were general in nature.

Finally, a key weakness of all of the studies we have described so far is that their designs rely on a nonmanipulated factor (political orientation). While this is a perfectly reasonable way of examining individuals with differing moral preferences, a more ideal approach would be to manipulate individuals' favored moral judgments and demonstrate that this manipulation affects their endorsement of moral principles. We sought to provide such a test by nonconsciously priming participants (Bargh and Chartrand, 1999) with words related to either patriotism (e.g., patriots, American, loyal) or multiculturalism (e.g., multicultural, diversity, equal) and then examining their reactions to our collateral damage scenarios. It was thought that these two primes mapped roughly onto intuitive reactions conservatives and liberals are likely to have when considering issues involving American troops and Iraqi insurgents. Consistent with this notion, we found these primes to produce a pattern of moral responses that closely matched those seen in our original collateral damage study. Participants exposed to patriotic words mimicked the pattern of judgments shown by political conservatives, endorsing a more consequentialist view of American-caused collateral damage than when the casualties were inflicted by Iraqi insurgents. Individuals exposed to multicultural words on the other hand, tended to show the opposite pattern, consistent with the judgments made by political liberals. The experimental nature of this evidence, and particularly its use of a nonconscious priming procedure, provide a final piece of evidence that reliance on normative moral principles can be driven by the kind of intuitive affective processes posited by a motivated reasoning account (Haidt, 2001).

6. MOTIVATED MORAL REASONING AND VIEWS OF THE MORAL THINKER

There is nothing so bad but it can masquerade as moral.

Walter Lippmann from *A Preface to Politics* (1914)

We have now reviewed research suggesting a variety of ways that moral reasoning is subject to motivational influence. Several characteristics of moral judgment make it particularly hospitable to motivated reasoning processes, but perhaps the clearest conclusion that emerges from our review regards the essential similarities between motivated reasoning mechanisms in moral and nonmoral domains. In either case, when individuals have an affective stake in reaching a particular conclusion, it can affect both descriptive beliefs about the evidence for that conclusion (e.g., whether a poor test grade really reflects poor underlying ability or an act that brings about negative consequences really reflects a negative underlying intention) as well as the normative criteria individuals use to evaluate that evidence (e.g., whether a hiring decision should be based on educational attainment or prior work experience or a moral decision should be based on deontological or consequentialist logic). What our review suggests more than anything else, however, is that additional research is needed to fully explore the territory of motivated moral reasoning, including what may make it unique from or similar to motivated reasoning in its various other manifestations.

More generally, thinking about moral judgment through the lens of motivated reasoning research has a number of implications for both classic and contemporary views of the moral thinker. The normative role of affect in moral thinking has been an issue of long-standing debate. Plato, Kant, and a host of other scholars have all viewed emotion as an irrational and essentially corrupting force on moral judgment, its influence something that is best eradicated or at least channeled toward virtuous goals by the powers of rationality and reason. Others, however, have challenged this view (mostly prominently Hume), arguing that our emotions reflect an inherent “moral sense,” and that reason without emotion would leave people with no moral compass by which to divine ethical behavior.

Contemporary research on moral reasoning suggests a nuanced view that captures elements of both sides of this ancient debate. Considerable research now supports the Humean position that moral judgment is inherently affective (Greene et al., 2001; Haidt, 2001; Moll et al., 2003; Valdesolo and Desteno, 2006; Wheatley and Haidt, 2005). Moreover, the intuitive infrastructure guiding these affective reactions is thought by many to have a functional logic grounded in evolved group living (Haidt, 2007; Haidt and Graham, 2007; Hauser, 2006). And yet, we would argue that it is precisely

this intuitive quality of moral judgment that leaves it open to potentially biasing motivational and affective influences. If moral reasoning is guided by affective intuitions, any factor that contributes to an individual's initial affective response — feelings about the actor, self-serving motivations, incidental affect — could potentially perturb how that reasoning proceeds (see [Valdesolo and Desteno, 2006](#) and [Wheatley and Haidt, 2005](#) for examples). It is important to point out again, however, that any form of judgmental bias is subject to the constraints of plausibility. Although the normative vagaries of moral judgment may give affective influences a relatively wide berth within which to operate, Lippman's assertion that there is no behavior so abhorrent that it cannot be recast as moral would seem an overstatement, or at least limited to cases of clinical psychopathy. There are likely certain behaviors that provoke in us such a strong visceral reaction (e.g., compunction or revulsion) that they simply cannot be rationalized as moral without offending our emotional sensibilities. Thus, the interesting thing about moral reasoning is that affect may play an important role both in promoting bias and restraining it.

Research on motivated moral reasoning also has implications for recent work equating deontological judgment with affective processing and consequentialist judgment with more systematic reasoning ([Greene, 2007](#); [Greene et al., 2004](#)). Although there is certainly some insight to this observation, and some support for it from functional brain imaging studies ([Greene et al., 2001, 2004](#)), it almost as certainly underestimates the complexity of affective influences on moral reasoning ([Bartels, 2008](#)). Our work, for example, shows that motivational factors can lead people toward reliance on either consequentialist or deontological principles depending on whether sacrificing innocent lives is consistent or inconsistent with an individual's ideologically based preferences. Research on moral reasoning will clearly benefit as more attention is directed toward situating it within the large body of work in social psychology documenting the multifaceted roles of affect, mood, motivation, and emotion on judgment.

A related limitation of current moral judgment research is a tendency to view deontological and consequentialist reasoning as qualitatively different. Our research, and that of others represented in this volume, increasingly suggests that rather than representing two distinctly different modes of reasoning, deontology and consequentialism reflect competing intuitions deriving from the belief that both acts and their consequences have moral import. The world is not populated with Immanuel Kants or Peter Singers, individuals who through sheer dint of will and intellect are able to advocate unforgivingly-consistent deontological or consequentialist moral systems. For most of us, intuition tells us that in some cases the ends do not justify the means, but in other cases they do. Thus, it would seem more apt to characterize people as having at their disposal a “moral toolbox” that can

be flexibly drawn upon depending on which moral intuitions are “primed” by any number of motivational or cognitive (e.g., attentional) factors.

Finally, another point we hoped to highlight in this review concerns the essentially social nature of moral thinking. This point, of course, has been made quite eloquently by others (e.g., Haidt, 2001), but the research reviewed here reminds us that central to moral judgments are assessments of people’s motives, thoughts, and intentions. Actions only become fully eligible for moral evaluation to the extent that an individual or group is perceived to have intentionally enacted them with reasonable knowledge of their good or bad consequences — that is, when the act can be fully attributed to the knowing desires of a social agent. A motivated reasoning perspective also highlights the fact that moral judgments can be affected by a host of social factors — our feelings about other people, the social groups to which we and other’s belong, etc. As psychological researchers increasingly utilize the complex puzzle cases drawn from normative ethics (sometimes referred to as “trolleyology”), it is important to remember that moral judgment is more than just an abstract reasoning task. It is a deeply social enterprise, subject to all the complexities, passions, and pressures that characterize human life across a wide variety of domains. At the end of the day, most moral judgments are not based on written descriptions of the implausible plights of “Jones” and “Smith,” but rather are made about friends, enemies, criminals, and politicians — people about whom we cannot help but feel strongly.

REFERENCES

- Alicke, M. D. (1992). Culpable Causation. *Journal of Personality and Social Psychology*, 63, 368–378.
- Alicke, M. (2000). Culpable Control and the Psychology of Blame. *Psychological Bulletin*, 126, 556–574.
- Aristotle, M. (1998). *The Nicomachean Ethics*. (D. Ross, Trans.). Oxford University Press, Oxford, England.
- Baron, J. and Spranca, M. (1997). Protected Values. *Organizational Behavior and Human Decision Processes*, 70, 1–16.
- Balcetis, E. and Dunning, D. (2006). See What You Want to See: Motivational Influences on Visual Perception. *Journal of Personality and Social Psychology*, 91, 612–625.
- Bargh, J. A. and Chartrand, T. L. (1999). The Unbearable Automaticity of Being. *American Psychologist*, 54, 462–479.
- Bartels, D. M. (2008). Principled Moral Sentiment and the Flexibility of Moral Judgment and Decision Making. *Cognition*, 108, 381–417.
- Baumeister, R. and Newman, L. (1994). Self-Regulation of Cognitive Inference and Decision Processes. *Personality and Social Psychology Bulletin*, 20, 3–19.
- Bork, R. H. (1990). *The Tempting of America: The Political Seduction of the Law*. Free Press, New York.
- Brennan, W. (1990). *The Great Debate: Interpreting Our Written Constitution*. The Federalist Society, Washington, DC.

- Dancy, J. (1993). *Moral Reasons*. Blackwell, Cambridge, MA.
- Dancy, J. (2004). *Ethics Without Principles*. Clarendon Press, Oxford.
- Dawson, E., Gilovich, T. and Regan, D. T. (2002). Motivated Reasoning and Performance on the Wason Selection Task. *Personality and Social Psychology Bulletin*, 28, 1379–1387.
- Dershowitz, A. (2001). *Supreme Injustice. How the High Court Hijacked the Election 2000*. Oxford University Press, New York.
- Dionne, E. J. and Kristol, W. (2001). *Bush v. Gore: The Court Cases and the Commentary*. Brookings Institution Press, Washington, D.C.
- Ditto, P. H. (in press). Passion, Reason, and Necessity: A Quantity of Processing View of Motivated Reasoning, in edited by Bayne, T. and Fernandez, J. (eds.), *Delusion, Self-Deception, and Affective Influences on Belief Formation*, Psychology Press, New York. in press.
- Ditto, P. H. and Boardman, A. F. (1995). Perceived Accuracy of Favorable and Unfavorable Psychological Feedback. *Basic and Applied Social Psychology*, 16, 137–157.
- Ditto, P. H. and Lopez, D. F. (1992). Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions. *Journal of Personality and Social Psychology*, 63, 568–584.
- Ditto, P. H., Munro, G. D., Apanovitch, A. M., Scepansky, J. A. and Lockhart, L. K. (2003). Spontaneous Skepticism: The Interplay of Motivation and Expectation in Responses to Favorable and Unfavorable Medical Diagnoses. *Personality and Social Psychology Bulletin*, 29, 1120–1132.
- Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M. and Lockhart, L. K. (1998). Motivated Sensitivity to Preference-Inconsistent Information. *Journal of Personality and Social Psychology*, 75, 53–69.
- Duff, R. A. (1990). *Intention, Agency, and Criminal Liability: Philosophy of Action and the Criminal Law*. Basil Blackwell, Oxford.
- Dunning, D. and Cohen, G. L. (1992). Egocentric Definitions of Traits and Abilities in Social Judgement. *Journal of Personality and Social Psychology*, 63, 341–355.
- Dunning, D., Leuenberger, A. and Sherman, D. A. (1995). A New Look at Motivated Inference: Are Self-Serving Theories of Success a Product of Motivational Forces? *Journal of Personality and Social Psychology*, 69, 58–68.
- Erdeyli, M. H. (1974). A New Look at the New Look: Perceptual Defense and Vigilance. *Psychological Review*, 81, 1–25.
- Fehr, E. and Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, 90, 980–994.
- Foot, P. (1967). The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review*, 5, 5–15.
- Forgas, J. P. (1992). Affect in Social Judgments and Decisions: A Multiprocess Model, in edited by Zanna, M. (ed.), *Advances in Experimental Social Psychology* (Vol. 25, pp. 227–275), Academic Press, New York.
- Forgas, J. P. and Smith, C. A. (2003). Affect and Emotion, in edited by Hogg, M. and Cooper, J. (eds.), *The Sage Handbook of Social Psychology* (pp. 161–189), Sage, London.
- Freud, S. (1923/1962). The Ego and the Id. The Standard Edition of the Complete Psychological Works of Sigmund Freud, W. W. Norton & Co., New York.
- Furgeson, J. R., Babcock, L. and Shane, P. M. (2008). Do a Law's Policy Implications Affect Beliefs about Its Constitutionality? An Experimental Test. *Law and Human Behavior*, 32, 219–227.
- Furgeson, J. R., Babcock, L. and Shane, P. M. (in press). Behind the Mask of Method: Political Orientation and Constitutional Interpretive Preferences. *Law and Human Behavior*.
- Gilbert, D. (2006). *Stumbling on Happiness*. Alfred A. Knopf, New York.
- Gintis, H., Bowles, S., Boyd, R. T. and Fehr, E. (2005). *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge, MA.

- Goodwin, G. and Darley, J. (2008). The Psychology of Meta-Ethics: Exploring Objectivism. *Cognition*, 106, 1339–1366.
- Greene, J. D. (2007). The Secret Joke of Kant's Soul, in edited by Sinnott-Armstrong, W. (ed.), *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*. MIT Press, Cambridge, MA.
- Greene, J., Nystrom, L., Engell, A., Darley, J. and Cohen, J. (2004). The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, 44, 389–400.
- Greene, J., Sommerville, R., Nystrom, L., Darley, J. and Cohen, J. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, 293, 2105–2108.
- Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108, 814–834.
- Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, 316, 998–1002.
- Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, 3, 65–72.
- Haidt, J. and Graham, J. (2007). When Morality Opposes Justice: Conservatives Have Moral Intuitions That Liberals May Not Recognize. *Social Justice Research*, 20, 98–116.
- Haidt, J., Koller, S. and Dias, M. (1993). Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog? *Journal of Personality and Social Psychology*, 65, 613–628.
- Hauser, M. D. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. Harper Collins, New York.
- Hume, D. (1739). *A Treatise of Human Nature*. Penguin, London.
- Jost, J. T., Glaser, J., Kruglanski, A. W. and Sulloway, F. (2003). Political Conservatism as Motivated Social Cognition. *Psychological Bulletin*, 129, 339–375.
- Kahneman, D., Diener, E. and Schwarz, N. (1999). *Well-Being: The Foundations of Hedonic Psychology*. Russell Sage Foundation, New York.
- Kahneman, D. and Miller, D. (1986). Norm Theory: Comparing Reality to Its Alternatives. *Psychological Review*, 93, 136–153.
- Kant, I. (1785/1998). *Groundwork for the Metaphysics of Morals*, in edited by Gregor, M. J. (ed.), Cambridge University Press, New York.
- Knobe, J. (2003). Intentional Action and Side-Effects in Ordinary Language. *Analysis*, 63, 190–193.
- Knobe, J. (2006). The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology. *Philosophical Studies*, 130, 203–231.
- Kohlberg, L. (1969). Stage and Sequence: The Cognitive-Developmental Approach to Socialization, in in edited by Goslin, D. A. (ed.), *Handbook of Socialization Theory and Research*, Rand McNally, Chicago.
- Kohlberg, L. (1984). *Essays in Moral Development: Vol. 2. The Psychology of Moral Development*. Harper & Row, New York.
- Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, 108, 480–498.
- Lerner, M. J. and Miller, D. T. (1977). Just World Research and the Attribution Process: Looking Back and Ahead. *Psychological Bulletin*, 85, 1030–1051.
- Lerner, J. S. and Tetlock, P. E. (1999). Accounting for the Effects of Accountability. *Psychological Bulletin*, 125, 255–275.
- Leslie, A. M., Knobe, J. and Cohen, A. (2006). Acting Intentionally and the Side-Effect Effect. *Psychological Science*, 17, 421–427.
- Lord, C., Ross, L. and Lepper, M. (1979). Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of Personality and Social Psychology*, 37, 2098–2109.
- MacCoun, R. J. (1998). Biases in the Interpretation and Use of Research Results. *Annual Review of Psychology*, 49, 259–287.
- Miller, D. T. and Ross, M. (1975). Self-Serving Biases in Attribution of Causality: Fact or Fiction? *Psychological Bulletin*, 82, 213–225.

- Moll, J., de Oliveira-Souza, R. and Eslinger, P. (2003). Morals and the Human Brain: A Working Model. *NeuroReport*, 14, 299.
- Monin, B. and Miller, D. T. (2001). Moral Credentials and the Expression of Prejudice. *Journal of Personality and Social Psychology*, 81, 33–43.
- Monin, B., Pizarro, D. A. and Beer, J. S. (2007). Deciding versus Reacting: Conceptions of Moral Judgment and the Reason-Affect Debate. *Review of General Psychology*, 11, 99–111.
- Munro, G. D. and Ditto, P. H. (1997). Biased Assimilation, Attitude Polarization, and Affect in Reactions to Stereotyped-Relevant Scientific Information. *Personality and Social Psychology Bulletin*, 23, 636–653.
- Norton, M. I., Vandello, J. A. and Darley, J. M. (2004). Casuistry and Social Category Bias. *Journal of Personality and Social Psychology*, 87, 817–831.
- Pizarro, D. A. and Bloom, P. (2003). The Intelligence of Moral Intuitions: Comment on Haidt (2001). *Psychological Review*, 110, 197–198.
- Pizarro, D. A., Laney, C., Morris, E. K. and Loftus, E. F. (2006). Ripple Effects in Memory: Judgments of Moral Blame Can Distort Memory for Events. *Memory and Cognition*, 34, 550–555.
- Pizarro, D., Uhlmann, E. and Bloom, P. (2003a). Causal Deviance and the Attribution of Moral Responsibility. *Journal of Experimental Social Psychology*, 39, 653–660.
- Pizarro, D., Uhlmann, E. and Salovey, P. (2003b). Asymmetry in Judgments of Moral Praise and Blame: The Role of Perceived Metadesires. *Psychological Science*, 14, 267–272.
- Plant, E. A. and Devine, P. G. (1998). Internal and External Sources of Motivation to Respond without Prejudice. *Journal of Personality and Social Psychology*, 75, 811–832.
- Pronin, E. (2008). How We See Ourselves and How We See Others. *Science*, 320, 1177–1180.
- Provencher, H. and Fincham, F. (2000). Attributions of Causality, Responsibility and Blame for Positive and Negative Symptom Behaviours in Caregivers of Persons with Schizophrenia. *Psychological Medicine*, 30, 899.
- Pyszczynski, T. and Greenberg, J. (1987). Toward an Integration of Cognitive and Motivational Perspectives on Social Inference: A Biased Hypothesis-Testing Model. *Advances in Experimental Social Psychology*, 20, 297–340.
- Reeder, G., Pryor, J., Wohl, M. and Griswell, M. (2005). On Attributing Negative Motives to Others Who Disagree With Our Opinions. *Personality and Social Psychology Bulletin*, 31, 1498–1510.
- Ross, L. and Ward, A. (1996). Naive Realism in Everyday Life: Implications for Social Conflict, in edited by Reed, E. S., Turiel, E. and Brown, T. (eds.), *Values and Knowledge. The Jean Piaget Symposium Series* (pp. 103–135), Lawrence Erlbaum Associates Inc., Hillsdale, NJ.
- Shaver, K. G. (1985). *The Attribution of Blame*. Springer-Verlag, New York.
- Simon, D., Krawczyk, D. C. and Holyoak, K. J. (2004a). Construction of Preferences by Constraint Satisfaction. *Psychological Science*, 15, 331–336.
- Simon, D., Snow, C. J. and Read, S. J. (2004b). The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction. *Journal of Personality and Social Psychology*, 86, 814–837.
- Skitka, L. J., Bauman, C. W. and Sargis, E. G. (2005). Moral Conviction: Another Contributor to Attitude Strength of Something More? *Journal of Personality and Social Psychology*, 88, 895–917.
- Slovic, P., Finucane, M., Peters, E. and MacGregor, D. (2007). The Affect Heuristic. *European Journal of Operational Research*, 177, 1333–1352.
- Sunstein, C. R. (2005). Moral Heuristics. *Behavioral and Brain Sciences*, 28, 531–573.
- Tannenbaum, D., Ditto, P. H. and Pizarro, D. A. (2008). Different Moral Values Produce Different Judgments of Intentional Action. Poster presented at the annual meeting of the Society for Personality and Social Psychology, Albuquerque, NM.

- Tanner, C. and Medin, D. L. (2004). Protected Values: No Omission Bias and No Framing Effects. *Psychonomic Bulletin and Review*, 11, 185–191.
- Tetlock, P. (2003). Thinking the Unthinkable: Sacred Values and Taboo Cognitions. *Trends in Cognitive Science*, 7, 320–324.
- Tetlock, P. E. and Levi, A. (1982). Attribution Bias: On the Inconclusiveness of the Cognition–Motivation Debate. *Journal of Experimental Social Psychology*, 18, 68–88.
- Thompson, J. J. (1986). *Rights, Restitution, and Risk: Essays in Moral Theory*. Harvard University Press, Cambridge, MA.
- Uhlmann, E., Pizarro, D. A., Tannenbaum, D. and Ditto, P. H. (2008). The Motivated Use of Moral Principles. *Unpublished manuscript*.
- Uhlmann, E. L. and Cohen, G. L. (2005). Constructed Criteria: Redefining Merit to Justify Discrimination. *Psychological Science*, 16, 474–480.
- Valdesolo, P. and Desteno, D. (2006). Manipulations of Emotional Context Shape Moral Judgment. *Psychological Science*, 17, 476–477.
- Wheatley, T. and Haidt, J. (2005). Hypnotically Induced Disgust Makes Moral Judgments More Severe. *Psychological Science*, 16, 780–784.
- Weiner, B. (1995). *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. Guilford Press, New York.
- Winkielman, P., Berridge, K. C. and Wilbarger, J. L. (2005). Emotion, Behavior, and Conscious Experience: Once More without Feeling, in edited by Feldman-Barrett, L., Niedenthal, P. and Winkielman, P. (eds.), *Emotion and Consciousness*, Guilford Press, New York.
- Wyer, R. and Frey, D. (1983). The Effects of Feedback about Self and Others on the Recall and Judgments of Feedback–Relevant Information. *Journal of Experimental Social Psychology*, 19, 540–559.
- Zajonc, R. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist*, 35, 151–175.