

Running Head: Person-Centered Morality

A Person-Centered Approach to Moral Judgment

Eric Luis Uhlmann

INSEAD

David A. Pizarro

Cornell University

Daniel Diermeier

Northwestern University

CONTACT:

Eric Luis Uhlmann
INSEAD
Organisational Behaviour Area
1 Ayer Rajah Avenue
138676 Singapore
E-mail: eric.luis.uhlmann@gmail.com

Abstract

Both normative theories of ethics in philosophy and contemporary models of moral judgment in psychology have focused almost exclusively on the permissibility of *acts*, in particular whether acts should be judged based on their material outcomes (consequentialist ethics) or based on rules, duties, and obligations (deontological ethics). However, a longstanding third perspective on morality, virtue ethics, may offer a richer descriptive account of a wide range of lay moral judgments. Building on this ethical tradition, we offer a *person-centered* account of moral judgment, which focuses on individuals as the unit of analysis for moral evaluations rather than on acts. Because social perceivers are fundamentally motivated to acquire information about the moral character of others, features of an act that seem most informative of character often hold more weight than either the consequences of the act, or whether or not a moral rule has been broken. This approach, we argue, can account for a number of empirical findings that are either not predicted by current theories of moral psychology, or are simply categorized as biases or irrational quirks in the way individuals make moral judgments.

There is a growing body of evidence that individuals are fundamentally motivated to evaluate others on a moral dimension—people quickly and easily attribute morally good or bad traits to others, and often do so early in an interaction and with limited information (Goodwin, Piazza & Rozin, 2014; Todorov, Said, Engell, & Oosterhof, 2008). However, most theoretical approaches in moral psychology have focused on the question of how individuals come to believe that certain acts are right or wrong (we refer to these as act-based theories). In contrast in what follows we outline a *person-centered* account of moral judgment (see also Pizarro & Tannenbaum, 2011), arguing that current act-based theories in moral psychology provide an incomplete account of moral judgment to the extent that they do not include the fundamental human motivation to determine the moral character of others. Simply stated, when making moral evaluations it appears as if individuals are often not asking themselves the question “is this act right or wrong?” but rather “is this person good or bad?”

The person-centered approach we propose places a renewed emphasis on the motivation to evaluate the character of others, arguing that it can account for a number of recent empirical findings that appear puzzling or irrational when viewed from the perspective of act-centered approaches. Specifically we will present evidence that judgments of a person’s underlying moral character can be empirically distinguished from judgments about the rightness or wrongness of an act (as demonstrated by evidence that judgments of acts can be dissociated from judgments of character), and that certain transgressions elicit especially negative reactions not because they are unusually wrong in-and-of-themselves, but because they are seen as highly diagnostic of an individual’s moral character. The process of arriving at a moral judgment, then, is often more influenced by what the action reflects about an agent’s moral character than by the degree of harm the act caused, or whether the act violates a set of moral rules. In short, there is growing

evidence that when it comes to moral judgment, human beings appear to be best characterized not as intuitive deontologists or consequentialists, but as intuitive virtue theorists: individuals who view acts as a rich set of signals about the moral qualities of an agent, and not as the endpoint of moral judgment.

In what follows, we will defend a specific set of claims regarding the centrality of character evaluation in moral judgment. Namely, we argue that 1) individuals are motivated to assess the character of others, and not just the rightness or wrongness of an act 2) some acts are perceived as more informative of an individual's moral character than others, and are therefore weighed heavily in moral judgments, 3) moral evaluations of acts and character can diverge, resulting in *act-person dissociations*, 4) judgments of moral character can infuse a host of other judgments that are central to moral evaluations (e.g., judgments of intentionality, agency, and blame), and 5) that a number of recent empirical findings demonstrating apparent inconsistencies in moral judgment may be better interpreted as reasonable for an individual motivated to assess the character of an agent rather than as simple "errors" of moral judgment.

The Person as Naïve Virtue Theorist

Although the goal of moral psychology is to shed light on the question of how and why individuals make judgments about moral right and wrong, not on how individuals *ought* to make these judgments, psychological theories of morality have traditionally relied a great deal on normative ethical theories as a point of departure for understanding moral judgment. In particular, theories in moral psychology have been influenced by two of the most dominant theories of normative ethics within the last century—consequentialism and deontology. Both of these theories specify a procedure for determining the morality of an act. Consequentialist theories hold that an act is right or wrong based solely on whether it maximizes good outcomes

(typically along quantifiable metrics such as happiness, resources spent, or lives saved; Smart & Williams, 1973), whereas deontological theories hold that an act is right or wrong based on whether it adheres to a set of rules, duties, and obligations viewed as foundational to morality (and can be wrong despite bringing about the best consequences; e.g., Kant, 1796/1996). Importantly, both of these theories focus fundamentally on acts as the unit of analysis. That is, the rightness or wrongness of an action is taken as independent of contextual features such as the history of the individual or the decision-making context. (Consequentialism, for instance, would hold that an act is wrong if it brings about negative consequences, regardless of *who* is taking the action).

The most influential descriptive accounts of moral psychology have been fundamentally shaped by these normative ethical traditions. For instance, the theories of Piaget (1932) and Kohlberg (1969), were based largely on deontological normative ethical theories (e.g., Kant, 1785; Rawls, 1975). For Kohlberg, for instance, individuals were characterized as moving from the belief that something was right or wrong based on conventional standards to a more “mature” morality in which they understood morality as fundamentally about deontological moral principles (such as justice). In contrast to these developmental approaches, another set of researchers studying moral judgment built on the consequentialist normative tradition, by focusing on degree to which individuals make moral decisions based on the consequences of an action, and under what conditions they seem to hold beliefs that are insensitive to consequences (e.g., Alter, Kernochan, & Darley, 2007; Baron, 1994; Darley, Solan, Kugler, & Sanders, 2010; Darley, Carlsmith, & Robinson, 2000; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001).

However, there is another ethical tradition that places persons (and not just actions) as the crucial unit of analysis for determining right and wrong (e.g., Aristotle, 4th Century B.C.E./1998; Hume, 1739/1888). These person-based approaches to normative ethics emphasize the character of the agent. Hume, for instance, wrote that “[a]ctions are by their very nature temporary and perishing; and where they proceed not from some cause in the characters and disposition of the person...[he] is not responsible for it: and as it proceeded from nothing in him, that is durable or constant, and leaves nothing of that nature behind it, 'tis impossible he can, upon its account, become the object of punishment or vengeance” (Hume, 1739/1888, as cited in Sripada, 2010, p. 9). As Hobart (1934, p. 84) writes regarding Hume’s perspective, “Morality has its eye upon acts, but an act is fleeting, it cannot be treasured and cherished. A quality can be, it lasts.”

These virtue-based normative ethical approaches may provide a better descriptive fit with the psychology of individuals than deontology and consequentialism in their central emphasis that morality is best understood at the level of persons rather than acts, and that perceived morality is fundamentally about possessing the right sorts of traits. It appears as if the mind is well equipped to make inferences about traits—especially traits that fall along the moral dimension. There is a great deal of evidence that individuals can make these inferences spontaneously and automatically, that the ability to make such evaluations emerges early in development, and that (despite some variability) people across cultures engage in similar sorts of evaluations (Choi & Nisbett, 1998; Fiske, Cuddy, & Glick, 2007; Hamlin, Wynn, & Bloom, 2007; Lieberman, Jarcho, & Obayashi, 2005; Kuhlmeier, Wynn, & Bloom, 2003; Todorov & Uleman, 2003; Willis & Todorov, 2006). Moreover, it appears that evaluating others on the dimensions of trustworthiness and warmth is something that individuals do almost immediately (Bar, Neta, & Linz, 2006; Fiske et al., 2007; Todorov et al., 2008). Individuals seek information

about the moral traits of others through the exchange of social gossip (Foster, 1997), and by looking for emotional signals (Ames & Johar, 2009) and patterns of behavior that may indicate the presence of positive or negative underlying traits (Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009; Schweitzer, Hershey, & Bradlow, 2006).

What sorts of traits fall along the dimension of moral character? Although lay conceptions of moral character are complex and multifaceted (Walker & Hennig, 2004), there is wide consensus that people view personal integrity and empathy as core components of moral character (Aquino & Reed, 2002; Lapsley & Lasky, 2001; Walker & Hennig, 2004; Walker & Pitts, 1998; Wojciszke, Dowhyluk, & Jaworski, 1998). Cues that a person possesses a stable set of traits regarding personal integrity (i.e., trustworthiness and fair treatment of others; Walker & Hennig, 2004) are of value because they suggest that a person can be relied upon to act cooperatively in the future. A lack of trustworthiness suggests a person will defect in joint endeavors when it suits his self-interest, and unfair treatment suggests he will not divide resources equitably. Likewise, the sorts of emotional reactions that seem to indicate care and concern (such as empathic reactions) may be seen as valuable indicators that a person is genuinely motivated toward prosocial action (and would feel constraint against harming others). There is support that these empathic traits may, in fact, serve as reliable indicators of future behavior—deficits in empathy are a hallmark of anti-social tendencies, and avoiding individuals with such tendencies obviously beneficial (Becker, Stuewig, Herrera, & McCloskey, 2004; Haidt, 2001; Walton-Moss, Mangello, Frye, & Campbell, 2005). Even young children are sensitive to this fact when seeking moral counsel, preferring an intellectually deficient but emotionally healthy adult rather than vice versa (Danovitch & Keil, 2008). An action that signals

an individual may have a deficit in empathy would serve as a very useful source of social information.

Perceived Informational Value, Moral Character, and Moral Blame

In setting the stage for a person-centered account of moral judgment we have argued that individuals are motivated to assess the presence of a broad set of traits in others. But what makes a particular behavior subjectively informative about underlying traits? The subjective informational value of any particular behavior (i.e., the probability increase that a behavioral cue provides toward making a strong inference; Nelson, 2005; Nelson et al., 2010) has been shown to depend on a number of characteristics. For example, negative acts (e.g., stealing) are perceived as more diagnostic of traits than positive acts (e.g., contributing to charity) because people often have ulterior motives for pro-social behaviors, such as a desire to create a positive impression on others (Reeder & Brewer, 1979). More generally, acts that can be attributed to multiple plausible motives or causes (i.e., are high in attributional ambiguity; Snyder, Kleck, Strenta, & Mentzer, 1979) tend to be seen as low in informational value. In contrast, behaviors that are statistically rare or otherwise extreme are perceived as highly informative about character traits (Ditto & Jemmott, 1989; Fiske, 1980; Kelley, 1967; McKenzie & Mikkelsen, 2007). In addition, decisions that are taken quickly and easily (Critcher, Inbar, & Pizarro, 2013; Tetlock, Kristel, Elson, Green, & Lerner, 2000; Verplaetse, Vanneste, & Braeckman, 2007), that are accompanied by genuine emotions (Trivers, 1971), and that involve costs for the decision maker (Ohtsubo & Watanabe, 2008) are perceived as especially informative about character.

An example of a seemingly small misdeed that is nonetheless taken as highly informative about an individual's moral character is that of a corporate executive who spends money on what is perceived to be frivolous perks, such as private planes, luxury cars, and country club

memberships (Gasparino, 2009; Hills & Michaels, 2002; Johnson, 2003). Such perks are often met with a high degree of outrage and public condemnation despite the fact that they represent a small proportion of expenditures relative to high corporate salaries more generally. A recent study demonstrated that this response likely occurs because individuals who request perks are assumed to possess a broader number of negative moral traits. When given the hypothetical task of hiring a corporate executive, participants reported a preference for paying an additional \$1 million in salary to a different job candidate just to avoid hiring a candidate whose salary request included a \$40,000 marble table (Tannenbaum, Uhlmann, & Diermeier, 2011, Experiment 2). Mediation analyses indicated that the requested perk was weighed heavily in hiring decisions because it was perceived as higher in informational value about the job candidate's moral integrity than requesting a higher salary. Participants also viewed a candidate who asked for such perks as more likely to act on his own selfish interests rather than the good of the company. Thus, it seems that participants viewed the goal of selecting a leader with good character as worth spending an extra million dollars on when tasked with a hiring decision.

There are cases when an act that causes comparatively less harm is viewed as *more* diagnostic about an agent's underlying character, because of the informational value it provides. In the United States, there are strongly held norms holding that treating individuals poorly based on their ethnicity is not justified (Crandall & Eshleman, 2003; Gaertner & Dovidio, 1986; Sears & Henry, 2005). Acts of racial bigotry speak strongly as to an agent's moral character, and can influence judgments of blame out of proportion to the actual harm caused. In a relevant empirical investigation, participants read about either a bigoted manager who mistreated only Black employees or a misanthropic manager who mistreated *all* of his employees. Even though he harmed far fewer people, participants viewed the bigoted manager's behavior as more

informative about his character. Further, although perceived informational value negatively predicted moral evaluations of the bigot, it was unrelated to evaluations of the misanthrope (Zhu, Uhlmann, & Diermeier, 2014).

A central claim of the person-centered account is that acts are often taken as a source of information regarding an agent's moral character. Importantly, even acts that may not be perceived as especially immoral in-and-of-themselves, are nonetheless seen as shedding light on the goodness or badness of an individual. If so, it should be possible to demonstrate a dissociation between judgments of the wrongfulness of acts and judgments of the moral character of agents who carry them out. It is to evidence of these dissociations that we now turn.

Act-Person Dissociations

News journalist Roland Martin, when discussing perceptions of Michael Vick (the NFL quarterback who was accused of organizing dog fights) argued that if Vick had beaten his girlfriend or murdered a human being, he might have been perceived more positively by the public (Martin, 2007). To investigate such phenomena, Tannenbaum et al. (2011, Experiments 1a and 1b) examined the possibility that although animal cruelty might be viewed as less immoral than violence towards humans, it would serve as a more informative signal that an individual possesses severe deficits in empathy, and would therefore be especially indicative of an agent's character. Consistent with this idea, participants who read a scenario in which a man acted violently toward his girlfriend judged the action as a more immoral act than when they read a similar scenario in which a man beat up his girlfriend's cat. However, participants judged a person who acted violently toward a cat as more coldhearted and sadistic than a person who acted violently toward his girlfriend. Participants may have been accurately judging the cat-beater's character—empirical investigations show animal cruelty is highly predictive of deficits

in empathy and antisocial future behaviors (Becker et al., 2004; Haidt, 2001; Walton-Moss et al., 2005).

Drawing a distinction between moral judgments of acts and persons may help explain negative gut reactions to the harmless-but-offensive transgressions examined in social intuitionist studies (for empirical evidence, see Haidt, Bjorklund, & Murphy, 2011; Haidt & Hersh, 2001; Haidt, Koller, & Dias, 1993; for theoretical reviews, see Haidt, 2001, 2007). Moral interviews find that people have negative gut reactions to harmless transgressions such as eating a dead dog that had been hit by a car and masturbating into a chicken carcass, but find themselves at a loss to explain why (i.e., they are “morally dumbfounded”; Haidt et al., 2011; Haidt et al., 1993). Negative intuitions in such contexts may be driven by a focus on the moral character of the agent rather than the tangible consequences of the act. Supporting this hypothesis, participants viewed harmless-but-offensive acts (e.g., sex with a chicken carcass) as less immoral than harmful acts (e.g., stealing a chicken carcass) but as speaking more strongly to global moral character (Uhlmann & Zhu, 2014; Experiments 1, 2, and 3).

Distinguishing person judgments from act judgments further sheds light on moral dumbfounding effects. As in prior work, participants were more likely to find themselves at a loss to justify their evaluations of harmless-but-offensive acts than harmful acts—most likely because the former are difficult to defend based on “rational” metrics like the monetary damage or physical harm caused. In contrast, the acts’ statistical extremity (Ditto & Jemmott, 1989; Fiske, 1980), low attributional ambiguity (Snyder et al., 1979) and therefore high informational value (Nelson, 2005; Nelson et al., 2010) provide rational reasons for drawing strong character inferences from behaviors like eating a dead dog and sex with a chicken carcass. Accordingly, participants were significantly *less* likely to be dumbfounded by their assessments of the moral

character of agents who engaged in harmless-but-offensive transgressions relative to harmful transgressions (Uhlmann & Zhu, 2014; Experiment 3). Thus, with regard to assessments of personal character, the moral dumbfounding effects made famous by Haidt and colleagues can completely reverse in a manner explained by the informational value of the transgressions involved.

It is important to show that act-person dissociations emerge not just for bizarre and statistically rare acts (e.g., cat-violence and unusual forms of masturbation), but also for incidents people might encounter in their everyday lives. One such context in which act-person dissociations may readily emerge is that of encountering racial bias. Once again, experiments demonstrate a dissociation between judgments of acts and judgments of character. Participants judged referring to a co-worker as a “nigger” as a less immoral act than physically assaulting the co-worker. However, using the racial slur was seen as more indicative of poor global moral character (Uhlmann, Zhu, & Diermeier, 2014; Experiment 1). It is noteworthy that participants drew very negative character inferences about the bigoted agent even though his behavior had no direct victim, in that he muttered the racial slur under his breath and no one heard him (although see Gray, Young, & Waytz, 2012). Further, even though they rated him as having committed the less immoral act, participants were less willing to be friends with the bigoted coworker than with the physically aggressive coworker (Uhlmann, 2013).

Another rich cue that serves as information that an individual possesses poor moral character is if the person appears to actively take pleasure in the suffering of others. Signals regarding the hedonic experience of agents as they carry out moral transgressions are viewed as deeply informative about an individual’s moral character. For instance, “hedonic markers” that suggested the perpetrator took pleasure in a killing (e.g., cutting the body into pieces) led

judgments of the person to become more negative, but not judgments about the wrongness of the act itself (Gromet, Goodwin, & Darley, 2013). The presence of these hedonic cues also made participants more likely to recommend the death penalty as punishment. A similar pattern of results emerged when participants were directly told the agent had experienced positive affect during the commission of the murder. With regard to receiving the ultimate punishment for moral transgressions, what your crime says about you as a person may matter as much as what you actually did.

Can similar dissociations extend to cases involving morally *praiseworthy* acts? In other words, can some acts considered necessary or even admirable nonetheless signal negative global moral character? Uhlmann, Zhu, and Tannenbaum (2013) hypothesized that this often occurs in the context of consequentialist acts. One reason is that consequentialist decisions can be attributionally ambiguous (Bartels & Pizarro, 2011). A decision maker may sacrifice the well-being of one person in order to maximize collective outcomes because they genuinely endorse consequentialist principles, or out of a self-interested desire to share in those collective benefits. In a relevant study, participants felt that throwing an injured man off a sinking lifeboat in order to save the remaining passengers was the morally correct course of action. However, they further perceived lifeboat passengers who actually made such a decision as potentially driven by selfish motives (i.e., a desire to save themselves) and therefore deficient in moral character (Uhlmann, Zhu, & Tannenbaum, 2013, Experiment 1).

A second reason for act-person dissociations in this context is that actually carrying out consequentialist decisions can require suppressing one's empathy for others. Participants viewed a hospital administrator who decided to buy necessary equipment (saving numerous future lives) rather than fund a costly operation for one little boy as having acted on principle and having

made a praiseworthy moral decision. But at the same time, the administrator who let a little boy die for the greater good was perceived as deficient in moral character, an effect mediated by his apparent lack of empathy (Uhlmann, Zhu, & Tannenbaum, 2013, Experiments 2 and 3).

Somewhat ironically given his perceived poor moral character, the consequentialist administrator was rated a better leader than an administrator who chose to forego buying the equipment in order to save one little boy. This suggests that leaders are judged by different criteria than everyday people, perhaps because it allows them to make the difficult trade-offs their positions require.

Of course, moral evaluations of acts and persons do not always diverge. For example, judgments of the act of short selling stocks and of individuals who engaged in such behavior were largely identical (Inbar, Pizarro, & Cushman, 2012). However, this does not mean judgments of a person's character are fully mediated by judgments of their acts. The reverse causal path can also hold, such that character inferences influence moral judgments of acts. Consistent with this idea, when participants were first led to think about the moral character of a person who engaged in a bigoted act (defacing a picture of Martin Luther King, Jr.), moral judgments of the act itself become correspondingly more negative (Uhlmann, Zhu, & Diermeier, 2014; Experiment 2). Thus, in cases where judgments of acts and persons correspond, person inferences may be driving evaluations of acts, rather than the reverse. As discussed in depth in the next section, person inferences can shape not only moral judgments of acts, but also perceptions of the ostensive precursors of moral judgment such as the perceived intentionality and controllability of the act.

Explaining Moral Biases

Psychologists studying the attribution of blame (e.g., Malle, Guglielmo, & Monroe, 2014; Shaver, 1985; Weiner, 1995) have outlined a set of conditions that are generally viewed by lay judges as necessary for the ascription of blame. Building on normative theories of blame from philosophy and legal theory, these accounts portray the attribution of moral blame as a stage-like process, where individuals first judge whether an agent (for instance) caused, desired, and intended an outcome, before finally arriving at a judgment of moral blame. Although these theories are generally taken as descriptive, it is assumed that deviations from this process of blame attribution constitute biases in judgment.

Although lay perceivers do adhere to normative models some of the time, a number of recent findings have demonstrated systematic departures from the process of blame attribution predicted by these models (Knobe & Doris, 2010). Indeed, moral judgments are frequently influenced by factors traditional normative models consider inappropriate or irrelevant, and many of these findings would be interpreted as evidence of the error-prone nature of moral judgments. The person-centered approach to moral judgment explains many supposed errors as the outputs of a moral system whose purpose is to determine the moral character of those around us (Pizarro & Tannenbaum, 2011). In what follows, we discuss two cases (the side-effect effect and asymmetries in judgments of blame and praise), for which there is empirical evidence that directly links the observed pattern of moral judgments to such person-centered concerns. We then discuss a number of other findings for which the person-centered account can offer a potential theoretical explanation, reinterpreting what might be deemed to be errors in moral judgment as reasonable responses given the underlying desire to evaluate the character of others.

The side-effect effect. One widely discussed anomaly of moral judgment is the tendency to judge the negative, but not positive side effects of decisions as intentional (Knobe, 2006, 2010). For example, a profit-minded corporate executive who releases a product he knows will have the side effect of causing environmental damage is viewed as having intentionally harmed the environment. However, when the product had the side effect of helping the environment, this act is not perceived as intentional (Leslie, Knobe, & Cohen, 2006). This pattern of intentionality judgments is observed even though in all conditions, the executive states his only goal is to make a successful product and he does not care one way or the other about the consequences for the environment. The “Side-Effect Effect” may occur because participants draw strong person inferences about a profit-minded executive unconcerned with environmental consequences, and this perception in turn drives their assessments of intentionality. In other words, intentionality judgments—the ostensive precursors of moral condemnation—may be mediated by assessments of the target’s personal values than the reverse. Consistent with this idea, structural equation modeling (SEM) shows that the corporate executive’s apparent values, principles, and characterological dispositions, rather than the goodness or badness of his actions, statistically explains whether the outcomes of his decisions are perceived as intentional (Sripada & Konrath, 2011; see also Sripada, 2012).

Asymmetries in judgments of blame and praise. Normative models of moral judgment stipulate that people should be held less accountable for acts that are difficult to consciously control. In line with such prescriptions, participants discount blame for immoral acts that are committed impulsively and automatically (e.g., physically attacking another person while overwhelmed by an uncontrollable rage). But departing from normative models, participants do *not* discount praise for prosocial acts that are committed impulsively (e.g., helping someone

while overwhelmed by uncontrollable feelings of compassion) (Pizarro, Uhlmann, & Salovey, 2003). This effect occurs because of the inferences lay perceivers have about an agent's "metadesires" (Frankfurt, 1971; Wolf, 1987), or desires about desires. Participants assume that most people do not wish to have negative emotional impulses, and do wish to have positive emotional impulses. This sense that the agent does not "really want" their antisocial impulses is what leads to the discounting of blame for impulsive negative acts (Pizarro et al., 2003).

Thus, as with the side-effect effect, asymmetries in the discounting of blame and praise are driven by the inferences lay perceivers make about the deeper desires and character-based dispositions of the agents involved (see also Sripada, 2010). Such person-centered inferences are not only logically defensible, but likely effective when it comes to determining who might be (or at least might become) a valuable ingroup member. Individuals whose impulsive desires and metadesires are consistent are likely to continue to engage in acts consistent with their emotional impulses in the future. In contrast, individuals who experience internal conflict about their impulses may eventually overcome them, and exhibit quite different (and even redeeming) patterns of behavior.

Culpable causation. Not only judgments of intentionality, but even assessments of causal control over negative events may be influenced by the perceived values of the agent. Research on culpable causation finds that, in a between-subjects design, a driver is more likely to be perceived as the cause of a complicated car accident if he was speeding home to hide a vial of crack cocaine as opposed to speeding home to hide a present for a loved one (Alicke, 1992). The desire to heap blame on an unlikable person involved in an accident appears to influence his perceived control over the event (Alicke, 1992, 2000). Thus, culpable causation effects and person centered morality are very much in harmony with one another: person centered

information about the underlying goodness or badness of the agent's moral character appears to shape assessments of his causal role in good and bad events. If moral judgments were solely based on evaluations of acts and not character, effects such as those reported by Alicke (1992) would not be obtained.

The culpable causation model provides greater generality in understanding moral staining, for instance demonstrating spillover not only between underlying motives and causal attributions, but also between the outcomes of an event and assessments of causal control over the event (Alicke, 2000, 2008; Mazzocco, Alicke, & Davis, 2004). Notably, our person centered approach likewise predicts effects outside the scope of the culpable causation model, such as theoretically meaningful dissociations between moral judgments of action and persons (Tannenbaum et al., 2001; Uhlmann et al., 2013, 2014; Uhlmann & Zhu, 2014). Thus, the two models can complement each other to account for certain phenomena (such as spillover between a person's motives and perceptions of her causal control over events), and at the same time each explain further effects that do not fall under the purview of the other model.

Importantly, there is a significant difference between the culpable causation model (Alicke, 1992, 2000) and our person-centered account as to whether phenomena such as bad motives influencing perceived causality constitute an irrational bias. On our view, character-based inferences spilling over to other judgments reflect a moral system bent on the adaptive goal of identifying good and bad people and using such information to navigate one's social world. Character-based inferences are not assumed to be irrational biases unless there is evidence that the character information being used as an input to understanding an event is irrelevant on some other grounds. There are times in which, especially in the absence of perfect information about what happened, or when there is a dispositional cue present that is known to be a reliable

predictor of immoral behavior, it may be perfectly rational for character influences to spill over to other judgments (for a related point, see Malle et al., 2014).

Identification with externally controlled acts. Inferences about character may influence not only attributions of causality, but also assessments of responsibility for transgressions that are clearly externally caused. Indeed, individuals are judged responsible for acts that were severely constrained by external circumstances, so long as their acts are consistent with their intentions. In one study, when an airline passenger was forced by hijackers to kill another hostage, he was judged as more responsible if he wanted to kill the other man anyway (i.e., if he “identified” with the act; Woolfolk, Doris, & Darley, 2006). Internally embracing a violent act in this way—even one compelled by external circumstances—is highly informative of poor moral character and may therefore influence judgments of responsibility.

Distorted recollections of intentional transgressions. Even knowledge of the basic facts about a moral transgression—the very earliest precursors of moral judgment according to normative models—can be distorted by the perpetrator’s perceived motives. When a restaurant customer intentionally left without paying his bill, participants recalled the cost of the meal as significantly greater than when he had merely forgotten to pay (Pizarro, Laney, Morris, & Loftus, 2006). Thus, recollections of the severity of an act can be influenced by the internal motivations of the person who carried it out.

Omission bias. Person centered inferences may further underlie the tendency to assign greater blame for crimes of commission than omission. For example, a tennis player who orders a rival a meal containing an allergen is judged more harshly than if he simply fails to warn the rival about an allergen in his meal (Spranca, Minsk, & Baron, 1991). Lay perceivers may reason

that it “takes a worse person” to actively harm others than to passively allow harm to take place, and this assumption may shape their assessments of blame and responsibility.

Summary. Act-based theories struggle to account for the tendency to judge unintended negative side effects as intentional, assign greater causal control over negative events to a dislikeable person, hold people responsible for externally compelled acts they internally identify with, misremember the severity of intentional harms, and assign greater blame for commissions than omissions. These findings are not predicted by most normative models of blame, and therefore tend to be interpreted as instances of bias. However, they are understandable as a reasonable set of evaluations from the products of a moral system bent on assessing moral character.

Conclusion

According to the person-centered account of moral judgment, human beings are intuitive virtue theorists who view acts as signals of underlying moral traits such as integrity and empathy for others. Relatively harmless actions high in informational value regarding character are therefore weighed heavily in moral judgments. Indeed, striking dissociations emerge between moral evaluations of acts and the persons who carry them out, such that some acts speak strongly to moral character despite not being condemned as especially harmful or immoral in-and-of themselves. Many putative biases and error of moral judgment may be the products of a moral system designed to determine the character of those around us. It is time for psychological theories of moral judgment to rediscover Hume’s insight that while acts are fleeting, the lasting qualities of moral character are to be treasured and cherished. Bringing this insight to contemporary theories of moral judgment will lead to a more comprehensive and complete psychology of human morality.

References

- Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, 63, 368-378.
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126, 556-574.
- Alicke, M. (2008). Blaming badly. *Journal of Cognition and Culture*, 1-2, 179-186.
- Alter, A. L., Kernochan, J., & Darley, J. M. (2007). Transgression wrongfulness outweighs its harmfulness as a determinant of sentence severity. *Law and Human Behavior*, 31, 319-335.
- Ames, D. R., & Johar, G. V. (2009). I'll know what you're like when I see how you feel. *Psychological Science*, 20, 586-593.
- Aquino, K. & Reed, A. II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83, 1423-1440.
- Aristotle (4th Century, B.C.E./1998). *The Nicomachean ethics*. Oxford: Oxford University Press.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, 6, 269-278.
- Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, 17, 1-10.
- Bartels, D., & Pizarro, D.A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154-161.
- Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? *Advances in Experimental Social Psychology*, 20, 65-122.
- Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality. *Behavioral and Brain Sciences*, 36, 59-122.
- Becker, K. D., Steuwig, J., Herrera, V. M., & McCloskey, L. A. (2004). A study of firesetting and

- animal cruelty in children: Family influences and adolescent outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 905–912.
- Choi, I. & Nisbett, R. E. (1998). Situational salience and cultural differences in correspondence bias and the actor-observer bias. *Personality and Social Psychology Bulletin*, 24, 949–960.
- Crandall, C.S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, 129, 414-446.
- Critcher, C., Inbar, Y. & Pizarro, D.A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science*, 4, 308-315.
- Danovitch, J. H., & Keil, F. C. (2008). Young Humeans: The role of emotions in children’s evaluation of moral reasoning abilities. *Developmental Science*, 11, 33-39.
- Darley, J.M., Carlsmith, K.M., & Robinson, P.H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, 24, 659-683.
- Darley, J. M., Solan, L. M., Kugler, M. B., & Sanders, J. (2010). Doing wrong without creating harm. *Journal of Empirical Legal Studies*, 7, 30–63.
- Ditto, P. H., & Jemmott, J.B. III (1989). From rarity to evaluative extremity: Effects of prevalence information on evaluations of positive and negative characteristics. *Journal of Personality and Social Psychology*, 57, 16–26.
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889–906.
- Fiske, S., Cuddy, A., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.
- Foster, E. K. (2004). Research on gossip: Taxonomy, methods, and future directions.

Review of General Psychology, 8, 78-99.

- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5-20.
- Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61-89). Orlando, FL: Academic Press.
- Gasparino, C. (2009, January 22). John Thain's \$87,000 rug. *The Daily Beast*. Retrieved from <http://www.thedailybeast.com>.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 1-21.
- Gray, K., Young, L., Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23, 101-124.
- Greene, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., & Cohen, J.D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107, 1144-1154.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral Judgment. *Science*, 293, 2105-2108.
- Gromet, D.M., Goodwin, G.P., & Darley, J.M. (2013). *Taking pleasure in doing harm: The influence of hedonic states on judgments of immorality and evil*. Unpublished manuscript.
- 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., Bjorklund, F., & Murphy, S. (2011). *Moral dumbfounding: When intuition finds no reason*. Unpublished manuscript.

- Haidt, J., & Hersh, M. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology, 31*, 191-221.
- Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613-628.
- Hamlin, J., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559.
- Hills, A. & Michaels, A. (2002, September 18). Paw taste condemns Kozlowski: Report says Tyco bought \$15,000 dog umbrella stand for chief’s apartment. *Financial Times*, First section, p. 1.
- Hobart, E. (1934). Free will as involving determination and inconceivable without it. *Mind, 43*, 1-27.
- Hume, D. (1739/1888). *A treatise of human nature*. Oxford: Oxford University Press.
- Inbar, Y., Pizarro, D.A., & Cushman, F. (2012). Benefiting from misfortune: When harmless actions are judged to be morally blameworthy. *Personality and Social Psychology Bulletin, 38*, 52-62.
- Johnson, C. (2003, October 28). Tyco trial turns to ‘birthday bash.’ *The Washington Post*, p. E01.
- Kant, I. (1796/2002). *Groundwork for the metaphysics of morals* (tr. Arnulf Zweig). New York: Oxford University Press.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp. 192–238). Lincoln: University of Nebraska Press.
- Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies, 130*, 203-231.

- Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, *33*, 315-329.
- Knobe, J., & Doris, J. (2010). Responsibility. In J. Doris and the Moral Psychology Research Group (Eds.) *The Moral Psychology Handbook* (pp. 321-354). Oxford, UK: Oxford University Press.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: Rand McNally.
- Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402-408.
- Lapsley, D. K. & Lasky, B. (2001). Prototypic moral character. *Identity*, *1*, 345-363.
- Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, *17*, 421-427.
- Lieberman, M. D., Jarcho, J. M. & Obayashi, J. (2005). Attributional inference across cultures: Similar automatic attributions and different controlled corrections. *Personality and Social Psychology Bulletin*, *31*, 889-901.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*, 147-186.
- Martin, R. S. (2007, Sept. 12). Commentary: Where is the outrage when humans are abused? *CNN*. Retrieved from <http://www.cnn.com/>
- Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, *26*, 131-146.
- McKenzie, C. R. M., & Mikkelsen, L. A. (2007). A Bayesian view of covariation assessment.

- Cognitive Psychology*, 54, 33–61.
- Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, 112, 979–999.
- Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters: Information acquisition optimizes probability gain. *Psychological Science*, 7, 960-969.
- Ohtsubo, Y., & Watanabe, E. (2008). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior*, 30, 114-123.
- Piaget, J. (1932/1965). *The moral judgment of the child*. (M. Gabain, Trans.) New York: Free Press. (Original work published 1932).
- Pizarro, D.A., Laney, C., Morris, E., & Loftus, E. (2006). Ripple effects in memory: Judgments of moral blame can distort memory for events. *Memory and Cognition*, 34, 550-555.
- Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In P. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil*. New York: APA books.
- Pizarro, D. A., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived meta-desires. *Psychological Science*, 14, 267-272.
- Rand, D. G., Dreber A., Ellingsen, T., Fudenberg, D., & Nowak, M.A. (2009). Positive interactions promote public cooperation. *Science*, 325, 1272-1275.
- Rawls, J. (2005). *A theory of justice* (2nd ed.) Harvard, MA: Belknap Press.
- Reeder, G. D. & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, 86, 61–79.

Schweitzer, M., Hershey, J., & Bradlow, E. (2006). Promises and lies: Restoring violated trust.

Organizational Behavior and Human Decision Processes, 101, 1-19.

Sears, D. O., & Henry, P. J. (2005). Over thirty years later: A contemporary look at symbolic

racism and its critics. *Advances in Experimental Social Psychology*, 37, 95-150.

Smart, J. J. C., & Williams, B. (1973). *Utilitarianism: For and against*. Cambridge, UK:

Cambridge University Press.

Snyder, M.L., Kleck, R.E., Strenta, A., & Mentzer, S.J. (1979). Avoidance of the handicapped:

an attributional ambiguity analysis. *Journal of Personality and Social Psychology*, 37,

2297-2306.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice.

Journal of Experimental Social Psychology, 27, 76-105.

Sripada, C.S. (2010). The deep self model and asymmetries in folk judgments about intentional

action. *Philosophical Studies*, 151, 159-176.

Sripada, C.S. (2012). Mental state attributions and the side-effect effect. *Journal of*

Experimental Social Psychology, 48, 232-238.

Sripada, S., & Konrath, S. (2011). Telling more than we can know about intentional action. *Mind*

and Language, 26, 353-380.

Tannenbaum, D., Uhlmann, E.L., & Diermeier, D. (2011). Moral signals, public outrage, and

immaterial harms. *Journal of Experimental Social Psychology*, 47, 1249-1254.

Tetlock, P. E., Kristel, O., Elson, B., Green, M., & Lerner, J. (2000). The psychology of the

unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal*

of Personality and Social Psychology, 78, 853-870.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding

- evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*, 455-460.
- Todorov, A., & Uleman, J.S. (2003). The efficiency of binding spontaneous trait inferences to faces. *Journal of Experimental Social Psychology*, *39*, 549-562.
- Trivers, R. (1971). Evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35-57.
- Uhlmann, E.L. (2013). The logic of moral outrage. *Behavioral and Brain Sciences*, *36*, 38.
- Uhlmann, E.L., & Zhu, L. (2014). Acts, persons, and intuitions: Person-centered cues and gut reactions to harmless transgressions. *Social Psychological and Personality Science*, *5*, 279-285.
- Uhlmann, E.L., Zhu, L., & Diermeier, D. (2014). When actions speak volumes: The role of inferences about moral character in outrage over racial bigotry. *European Journal of Social Psychology*, *44*, 23-29.
- Uhlmann, E.L., Zhu, L., & Tannenbaum, D. (2013). When it takes a bad person to do the right thing. *Cognition*, *126*, 326-334.
- Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: the sequel. A kernel of truth in predictive cheating detection. *Evolution and Human Behavior*, *28*(4), 260-271.
- Walker, L. J. & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, *86*, 629-647.
- Walker, L. J. & Pitts, R. C. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology*, *34*, 403-419.
- Walton-Moss, B. J., Mangello, J., Frye, V., & Campbell, J. C. (2005). Risk factors for intimate

- partner violence and associated injury among urban women. *Journal of Community Health, 30*, 377–89.
- Willis, J. & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592–598.
- Wojciszke, B., Dowhyluk, M., & Jaworski, M. (1998). Moral and competence-related traits: How do they differ? *Polish Psychological Bulletin, 29*, 283–294.
- Wolf, S. (1987). Sanity and the metaphysics of responsibility. In F. Schoeman (Ed.), *Responsibility, character, and the emotions: New essays in moral psychology* (pp. 363–373). Cambridge, UK: Cambridge University Press.
- Woolfolk, R. L., Doris, J. M., & Darley, J. M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility. *Cognition, 100*, 283–301.
- Zhu, L., Uhlmann, E.L., & Diermeier, D. (2014). *Moral evaluations of bigots and misanthropes*. Study report available on the Open Science Framework at: <https://osf.io/a4uxn/>