Empathy is a moral force

Jamil Zaki
Department of Psychology, Stanford University

FORTHCOMING in Gray, K. & Graham, J. (Eds.), The Atlas of Moral Psychology

Address correspondence to:
Jamil Zaki
Department of Psychology
Stanford University
Stanford, CA 94305
jzaki@stanford.edu

Word count: 3,473
Abstract word count: 69
ABSTRACT

What is empathy’s role in driving moral behavior? Here I argue that empathy is a noisy but useful moral compass, and in particular that (i) empathy’s limits in guiding morality often reflect empathizer’s motives, not their capacities, (ii) motivating people to empathize can overcome these limits, and (iii) empathy lends affective “force” to morality, such that empathy-based moral behavior produces benefits that other forms of moral action do not.
People share each other’s emotional lives. We react not only to events that befall us, but also to the experiences of those around us. The tendency to share, understand, and care about others’ inner lives together constitute empathy (Davis, 1994; Decety & Jackson, 2004; Zaki & Ochsner, forthcoming).

Psychologists and neuroscientists have produced a deluge of research on empathy in recent decades. This trend reflects interest not only in empathy’s characteristics, but also in its power to encourage prosocial and moral action. Individuals help the targets of their empathy (Batson, 1991, 2011) and “humanize” those targets, for instance by resisting stereotypes about them or treating them fairly (Galinsky & Moskowitz, 2000). These behaviors likely reflect empathy’s effects in prompting aversion to others’ pain and pleasure in others’ success (Lamm, Decety, & Singer, 2011; Morelli, Lieberman, & Zaki, in press; Morelli, Sacchet, & Zaki, 2014), which in turn provide an intuitive “compass” guiding moral action. Individuals with psychopathy, who often lack empathy, provide a striking example of how important this compass is (Blair, 2005).

Does empathy’s compass always guide people towards ideal moral behavior? Probably not. Like other affective responses involved in moral decision-making (Haidt, 2001), empathy is noisy and biased. It can spur concern for the well being of some people, but not others, for instance skewing prosocial behavior unfairly towards ingroup members. Empathy can even generate clearly immoral choices, for instance when empathy for one’s own community encourages aggression towards other groups (Lickel, Miller, Stenstrom, Denson, & Schmader, 2006).

These biases and constraints have led theorists to propose that empathy constitutes a suboptimal, and even dangerous, source of moral behavior (Bloom, 2013;
Prinz, 2011a). For instance, Bloom (2013) highlights the case of “baby Jessica,” who became trapped in a well. Her case was captured on television and subsequently produced a worldwide outpouring of empathy. Donors provided hundreds of thousands of dollars to support baby Jessica, while ignoring the simultaneous suffering of countless others.

How could an emotional state that produces such misguided moral behavior ever be trusted? Critics of empathy suggest that morality can better serve the greater good if it is guided by utilitarian principles (i.e., doing the most good for the most people), as opposed to emotion. This viewpoint is important and clearly right in many cases. It is also incomplete, and risks discarding the baby with the proverbial well water.

Here I offer a counterpoint to recent criticisms of empathy, in two parts. First, I suggest that the limits of empathy are not stable, and instead reflect individuals’ motivation to connect with or avoid others’ experiences. These motives shift dynamically across situations, and strategies that increase empathic motivation can also reduce biases associated with empathy. Second, although utilitarian principles best guide the behavior of large groups, individuals who act morally “with feeling” are likely to be more committed to and fulfilled by their behaviors. Thus, to the extent that people can align their principles and affect, empathy can lend emotional meaning to moral actions.

**Historical context**

The modern concept of empathy is tied at the roots to moral philosophy. Adam Smith, in his *The Theory of Moral Sentiments* (1790/2002), famously described the “fellow feeling”—vicarious experience of others’ emotions—as a source of civilizing and moral
action. In the intervening centuries, empirical data have borne out Smith’s insight. In particular, Batson systematically demonstrated that empathy encourages moral actions, including the maintenance of equity and kindness to people in need (e.g., Batson, 2011; Batson et al., 2003; Batson & Shaw, 1991). At a macro level, key moments in a culture’s moral development often follow a shift in popular empathy. For instance, Uncle Tom’s Cabin, which prompted widespread empathy for the struggles of slaves, also intensified support for the abolitionist movement of the 19th century (Appiah, 2010).

This and other work has inspired scientific accounts under which empathy provides a vital, evolutionarily old emotional foundation for moral action (De Waal, 2010), and the expansion of moral values (Pinker, 2011). On such accounts, social turmoil such as political polarization reflects people’s failure to empathize with members of other groups (Trout, 2009), and remedying such problems requires either reinstating lost empathy or building empathic concern for ever wider swaths of the population (Krznaric, 2014; Rifkin, 2009; Singer, 2011).

More recently, however, a growing countercurrent has questioned the utility of empathy in driving moral action. This argument builds on the broader idea that emotions provide powerful but noisy inputs to people’s moral calculus (Haidt, 2001). Affective reactions often tempt people to make judgments that are logically and morally indefensible. Such emotional static famously includes moral dumbfounding, under which people’s experience of disgust causes them to judge others’ actions as wrong when they have no rational basis for doing so (Cushman, Young, & Hauser, 2006). Emotion drives other irrational moral judgments, such as people’s tendency to privilege physical force (a “hot” factor) over more important dimensions such as harm when judging the moral
status of an action (Greene, 2014; Greene et al., 2009). Even incidental, morally irrelevant feelings alter moral judgment, further damaging the credibility of emotion in guiding a sense of right and wrong. (Wheatley & Haidt, 2005).

In sum, although emotions play a powerful role in moral judgment, they need not play a useful role. Instead, capricious emotion-driven intuitions often attract people towards internally inconsistent and wrong-headed judgments. From a utilitarian perspective aimed at maximizing well being, these biases render emotion a fundamentally mistaken moral engine (cf. Greene, 2014).

Does this criticism apply to empathy? In many ways, it does. Like other affective states, empathy arises in response to evocative experiences, often in noisy ways that hamper objectivity. For instance, people experience more empathy, and thus moral obligation to help, in response to the visible suffering of others, as in the case of Baby Jessica described above. This empathy leads people to donate huge sums of money to help individuals whose stories they read about or see on television, while ignoring widespread misery that they could more efficaciously relieve (Genevsky, Västfjäll, Slovic, & Knutson, 2013; Slovic, 2007; Small & Loewenstein, 2003). Empathy also collapses reliably when sufferers and would-be empathizers differ along dimensions of race, politics, age, or even meaningless de novo group assignments (Cikara, Bruneau, & Saxe, 2011; Zaki & Cikara, in press).

Even when people experience empathy, the causal link between affect and moral action can be circuitous and noisy. At an inter-individual level, empathy fails to predict sensitivity to just versus unjust outcomes (Decety & Yoder, 2015). Worse, in some cases empathy inspires expressly unjust behavior. For instance, close connection with ingroup
members can prompt aggression towards outgroup members (Gilead & Liberman, 2014; Waytz & Epley, 2012). Even when empathy generates rapport across group boundaries, it can do so at the cost of justice. Low status group members who empathize with higher status individuals grow reticent to criticize unfair structural norms such as unequal access to education and resources (Dixon, Tropp, Durrheim, & Tredoux, 2010; Saguy, Tausch, Dovidio, & Pratto, 2009). In these cases, empathy improves surface-level relations between groups at the cost of more meaningful social change (Zaki & Cikara, in press).

Empathy thus falls prey to the same limitations as other emotions in driving moral behavior: it is hot-headed, short-sighted, and parochial. We would begrudge these qualities in policy makers trying to render the most good for the largest number of people. So why should people ever rely on empathy when making moral judgments?

**Theoretical stance**

Empathy is not a perfect source of moral choice, but I think its recent critics have dismissed it too readily. This is for two reasons.

*Empathy is motivated, and limits on empathy are not stable.* Critics of empathy characterize it as biased: responding to morally irrelevant content (e.g., visibility) while missing morally relevant content (e.g., total suffering). Must moral “mistakes” like these always characterize empathy? In order to answer this question, we must first ask another: to what extent can people control their experience of empathy? Psychological theory and lay intuition converge to suggest that empathy tends to fall out of our control. Imagine, for instance, witnessing someone suffer a horrific industrial accident. In cases like this, it doesn’t seem as though observers select their level of empathy; vicarious distress simply
happens to them. The assertion that empathy is automatic runs through the philosophy (Goldman, 2006), psychology (Hatfield, Forbes, & Rapson, 2013), and neuroscience (Gallese, 2007) of empathy (for a comprehensive review of this account, see Zaki, 2014).

If empathy is automatic, then people can control neither when they feel empathy nor when they do not. Under such a state of affairs, the biases that characterize empathy—such as in-group favoritism—are as uncontrollable as the experience of empathy itself. This assumption underlies arguments for minimizing the role of empathy in moral decision-making. For instance, Greene (2014) suggests that “automatic empathy programs” lead people towards poor moral choices, and that “it would be foolish to let the inflexible operating characteristics of our empathy gizmos serve as foundational moral principles” (pg. 264).

This take on empathy reflects a broader view of emotions as fundamentally distinct from and inaccessible to cognition. This model is at least as old as Plato’s account of reason and passion. It is also incorrect. Decades of data from affective science demonstrate that logic and emotion interact pervasively. Moral theorists focus largely on one side of this interaction: ways that emotions inadvertently color thinking, producing irrational but strongly held judgments. The opposite causal direction—from thinking to affect—matters just as much. Cognition takes part in “constructing” the emotions that people feel (Barrett, 2013; Schachter & Singer, 1962) and judge others to feel (Ong, Zaki, & Goodman, in press). Cognition-affect interactions also characterize emotion regulation, through which people alter their feelings in response to their goals (Gross, 2015; Inzlicht, Bartholow, & Hirsh, 2015; Ochsner, Silvers, & Buhle, in press).
Emotion regulation often comprises people’s attempts to feel better by reducing negative affect or maximizing positive affect, but people sometimes want to feel badly. For instance, prior to conflicts, people up-regulate their experience of anger, and prior to making a request for support, they up-regulate sadness (Tamir, 2009). In these contexts, negative emotions help people accomplish their goals, and thus become the target of regulation strategies. “Social” emotions such as gratitude, righteous anger, and guilt are particularly useful in driving interpersonal outcomes such as cooperation (DeSteno, 2015; Trivers, 1971). As such, people regulate their experience of these states in response to social goals.

I propose that empathy follows suit. Instead of succumbing to their experience (or non-experience) of empathy, people often choose to engage with or avoid others’ emotions. This choice can be conscious or non-conscious, and—like other forms of emotion regulation—(i) reflects people’s goals in a given context and (ii) can be carried out through multiple strategies (for a thorough review of this model, see Zaki, 2014).

A motivated account recasts the “empathic failures” described above. When people exhibit blunted empathy to strangers or outgroup members, this does not mean they are incapable of empathizing; instead they might be unmotivated to do so (Keysers & Gazzola, 2014). This is important because it suggests that limits of empathy are not “baked in” to the nature of empathy itself. Instead, they signal local features of a situation that reduce people’s propensity to empathize. To the extent this is the case, empathic limits can be overcome by increasing motivation to empathy, and empathy can be harnessed to build moral concern on a broad scale (more on this in Extension and expansion below).
Empathy and Morality

Empathy-based action confers unique benefits. If empathy is motivated, people should be capable of empathizing in “smarter” ways that supersede group boundaries and other morally irrelevant factors. Still, empathy will always be subject to some noise, and cannot match the optimal moral principles that emerge from a utilitarian approach (Singer, 2015). Empathy will never provide anyone with a perfect moral compass. As such, even if empathic limits can be overcome, why bother?

One important reason is that prosocial and moral action driven by empathy might differ from action based solely on principle. These differences confer at least some advantages to empathy-based action, many of which reflect the added force emotions lend to action.

First, emotional goals often take precedent over non-emotional goals, and people pursue such goals with urgency and immediacy. Principles are difficult to abide on an empty stomach, or under other states that tax people’s psychological energy. For instance, cognitive load interferes with utilitarian moral judgments (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Emotions, by contrast, guide behavior efficiently and in ways that are robust to limits on people’s cognitive bandwith. As such, to the extent that people can tune empathy to match their principles, they gain access to a “hot,” emotional engine for powering prosocial behavior (DeSteno, 2009).

Second, emotion-based moral behavior might confer benefits that other moral behaviors do not. Prosocial actions “help the helper,” such that acting kindly renders people healthier and longer-lived (Dunn, Aknin, & Norton, 2014; Thoits & Hewitt, 2001). These benefits likely reflect boosts in subjective well being—such as increased happiness and decreased stress—that prosociality provides (Dunn, Aknin, & Norton,
2008; Zaki & Mitchell, 2013). I propose that these boosts most often follows prosocial acts driven by passion, not principle. The benefits of emotion-driven moral action likely transcend single individuals. Recipients of others’ support, for instance, benefit most from emotion-driven help, and emotion strengthens the reputational and relational benefits associated with helping others.

In sum, emotions in general—and empathy in particular—adds weight both to the efficiency of prosocial actions and to their benefits. If this is the case, then cultivating empathy-based morality stands as a worthwhile goal.

Evidence

Decades of research support a motivated model of empathy (see Zaki, 2014 for review). In particular, it is clear that situational factors reliably increase and decrease people’s desire to empathize. This is reflected across self report, behavior, and brain activity (Cameron & Payne, 2011; Keysers & Gazzola, 2014; Tamir, 2013). When empathy is goal inconsistent—for instance when carries heavy financial or emotional costs, or interfere with people’s ability to compete with others—people avoid empathy-provoking situations and cues (Davis et al., 1999; Pancer, McMullen, Kabatoff, Johnson, & Pond, 1979; Shaw, Batson, & Todd, 1994).

By contrast, manipulations that render empathy more goal-relevant cause people to expand their empathic experience. People who are lonely or desire social connection, as compared to people who are more socially “sated,” play closer attention to others’ internal states (Gump & Kulik, 1997; Pickett, Gardner, & Knowles, 2004). Likewise, people who believe that empathy is socially desirable or common among their peers also
Empathy and Morality

act empathically themselves, even towards outgroup members (Nook, Ong, Morelli, Mitchell, & Zaki, under review; Tarrant, Dazeley, & Cottom, 2009; Thomas & Maio, 2008). In these cases, goal relevance increases empathic effort, or the extent that people pursue empathy. Empathy and its consequences—including prosocial and kind action—follow suit.

Interestingly, the very notion that empathy is out of people’s control might hinder empathic effort. This follows from the broader idea of “lay theories,” or beliefs people hold about psychological constructs. Dweck (2006; Dweck & Leggett, 1988) has demonstrated that people vary in their belief about whether characteristics such as intelligence, prejudice, and personality are “fixed” and out of their control, or “malleable” and within their control. People who hold malleable, as compared to fixed, theories embrace challenges and difficulties as opportunities to grow valued psychological skills (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, Chiu, & Hong, 1995; Rattan & Dweck, 2010; Yeager, Trzesniewski, & Dweck, 2012).

Recently, we found that lay theories concerning empathy likewise affect empathic motivation (Schumann, Zaki, & Dweck, 2014). People who held a malleable empathic theory—or in whom we induced such a theory—exhibited more willingness to empathize than people holding a fixed lay theory. This difference was especially stark in cases where people might not otherwise feel motivated to empathize, such as intergroup settings or when empathy promised to be painful.

These data dovetail to support the idea that empathy is far from automatic, and often reflects people’s motives to connect with or avoid others’ emotions. This further suggests, crucially, that features of empathy that often render it a poor moral compass can
be reversed. Increasing empathic motives can also expand the *scope* of empathy, even to cases in which it typically fails.

Evidence also supports the contention that empathy lends weight to moral actions and the benefits they confer. Consider the effect of prosociality on well being (Dunn et al., 2014; Thoits et al., 2001). Kindness pays dividends to those who engage in it, but more recent work suggests that such benefits are strongest for individuals who are affectively engaged in their prosocial acts. For instance, volunteering decreases mortality risks in older adults, but only if their service is driven by “other-oriented” motives such as empathic concern for people in need (Konrath, Fuhrer-Forbis, Lou, & Brown, 2012). Likewise, college students experience increased happiness and reduced stress after providing practical help to others, but these effects are strongest in people who empathically engage with the targets of their help (Morelli, Lee, Arnn, & Zaki, in press).

Empathy-based prosociality maximally benefits not only helpers, but also the recipients of their help. When people turn to each other under both difficult and happy circumstances, they seek out not only concrete help, but also emotional connection (Rimé, 2007; Zaki & Williams, 2013). As such, people feel closer to support providers who experience and exhibit empathy, as opposed to those who provide less emotionally responsive support (Gable, Gonzaga, & Strachman, 2006; Gable & Reis, 2010).

Finally, emotion serves as a broader social signal about the meaning of prosocial actions. Groups elevate the moral status of people who act prosocially (Hardy & Van Vugt, 2006). Here, too, emotion matters. Individuals who express emotion when acting prosocially are perceived by others as more genuinely motivated to help others, further building the social capital kindness provides (Barasch, Levine, Berman, & Small, 2014).
Broadly, this work demonstrates that people benefit most not from moral acts alone, but rather from those that are imbued with affective force. Empathy is a messy source of prosociality, and at a broad policy level utilitarian principles provide a clearer moral compass than affect. But at an individual level, not only can people flexibly align empathy with their principles; doing so renders their subsequent moral action more powerful.

Extension and expansion

President Obama routinely refers to an “empathy deficit” that threatens the cohesion of our social fabric (Obama, 2006). Consistent with his view, college students’ self-reported empathy has dwindled over the last 30 years (Konrath, O’Brien, & Hsing, 2011). Empathy deficits pervade crucial social problems, such as the increasing polarization that characterizes our political system (Prior, 2013), the rise of bullying among adolescents (Wang, Iannotti, & Nansel, 2009), and medical professionals’ lack of connection to their patients (Haque & Waytz, 2012).

The ideas laid out above offer two broad points about this state of affairs. First, they make novel suggestions about how we can address the empathy deficit (cf. Zaki & Cikara, in press). A small but growing number of interventions focus on building empathy across settings including medical training (Riess, Kelley, Bailey, Dunn, & Phillips, 2012), education (Şahin, 2012), conflict resolution (Todd & Galinsky, 2014), and the treatment of clinical populations in which empathy is impaired, such as autism spectrum disorders (Golan & Baron-Cohen, 2006; Hadwin, Baron-Cohen, Howlin, & Hill, 1996). The majority of these interventions focus on two strategies: building
empathic skills such as emotion recognition, and inducing people to think more about social targets. This strategy is effective overall (van Berkhout & Malouff, in press), but a motivated model of empathy suggests that it is also incomplete. In many cases, people fail to empathize not because they are incapable of doing so, but because they are unmotivated to share, understand, or generate concern for others’ internal lives. As such, interventions should complement training in empathic skills with “psychological levers” (Miller & Prentice, 2010) that can build people’s desire to empathize in the first place.

Second, the work covered here offers evidence about why we should care about empathic deficits. On some accounts, a lack of empathy—although alarming—might be inconsequential. This conclusion is predicated on the idea that empathy is at best a dubious source of moral behavior, and at worst a barrier to broad moral progress (Bloom, 2014; Greene, 2014; Prinz, 2011b). Here I propose that this perspective obscures both people’s flexibility to grow their empathy beyond its typical limits, and also the unique power of empathy in rendering moral acts more beneficial.

Empathy is noisy, but scientists should not be so quick to dismiss it as a moral force. Although unstable when compared to principles, emotion can lend those principles deeper psychological meaning.
REFERENCES


