

COMMENTS

The Intelligence of the Moral Intuitions: Comment on Haidt (2001)

David A. Pizarro and Paul Bloom
Yale University

The social intuitionist model (J. Haidt, 2001) posits that fast and automatic intuitions are the primary source of moral judgments. Conscious deliberations play little causal role; they are used mostly to construct post hoc justifications for judgments that have already occurred. In this article, the authors present evidence that fast and automatic moral intuitions are actually shaped and informed by prior reasoning. More generally, there is considerable evidence from outside the laboratory that people actively engage in reasoning when faced with real-world moral dilemmas. Together, these facts limit the strong claims of the social intuitionist model concerning the irrelevance of conscious deliberation.

Within social psychology, the nineties was the decade of automaticity. Influential work from researchers such as Bargh (1994) and Greenwald and Banaji (1995) inspired a host of research documenting the unconscious nature of much of human beings' social judgment and behavior. In addition, research over the past 20 years has added vastly to psychologists' knowledge of emotion's role in judgment and decision-making (e.g., Forgas, 1995; Petty & Cacioppo, 1986; Schwarz & Clore, 1983). However, the study of moral judgment has remained relatively uninfluenced by these findings, and psychologists studying this process have continued to focus mainly on the role played by factors such as rational deliberation and the development of cognitive abilities (e.g., Turiel & Neff, 2000).

Recently, as an alternative to "rationalist" models of moral thought, Haidt (2001) defended a social intuitionist model. He proposed that moral judgments are typically the direct products of intuitions—fast, effortless, and automatic affective responses that present themselves to consciousness as immediate judgments. These intuitions are to some extent the products of a Darwinian "moral sense" that has evolved through natural selection, but they are also shaped by the cultural context in which an individual is raised, and, in particular, by the beliefs and practices of the individual's peer group.

In support of this view, Haidt (2001) pointed to evidence from social psychology showing that certain judgments are strongly affected by unconscious factors. He also provided some striking examples of *moral dumbfounding*—instances in which an individual is unable to generate adequate reasons for his or her moral judgment yet stands fast in believing it to be true. For instance,

when told a story in which there is consensual sex between adult siblings, many people (though not all) will insist that it is wrong, even though they cannot articulate their reasons for this view. Haidt argued that this research supports the more general conclusion that our moral judgments are not based on moral reasoning.

We are sympathetic to many of the claims that Haidt (2001) made. We agree that psychology has suffered from an overreliance on reasoning at the expense of affective processes (Pizarro, 2000). We also agree that there are basic moral intuitions that cannot themselves be justified by reason—"self-evident" truths as Thomas Jefferson put it. (Indeed, this is the case for all domains of reasoning, including deductive inference and inductive generalization.) And there is little doubt that many of our emotional assessments are the products of natural selection. The love we feel toward our children and our anger at those who cheat us can reasonably be thought of as biological adaptations that exist because of the selective advantages they gave to our ancestors over the course of evolution (e.g., Darwin, 1872/1998; Trivers, 1971; see Pinker, 1997, for a review).

Finally, Haidt (2001) was surely right when he concluded that we often have no conscious understanding of why we feel what we feel. William James (1890/1950) put it as follows:

Not one man in a billion, when taking his dinner, ever thinks of utility. He eats because the food tastes good and makes him want more. If you ask him why he should want to eat more of what tastes like that, instead of revering you as a philosopher he will probably laugh at you for a fool. . . . It takes, in short, what Berkeley calls a mind debauched by learning to carry the process of making the natural seem strange, so far as to ask for the why of any instinctive human act (pp. 1007–1008).

In fact, when contemporary philosophers try to explore precisely what is so wrong about killing babies or having sex with chickens, the reaction is often ridicule and anger (see Saletan, 2001, for discussion). Indeed, for many, the mere act of contemplating the morality of certain acts is deeply aversive and, potentially, a cause for atonement (Fiske & Tetlock, 1997).

All of these facts are consistent with social intuitionism, but they also mesh with most other theories of moral cognition. The primary contrast between social intuitionism and other theories of

David A. Pizarro and Paul Bloom, Department of Psychology, Yale University.

We thank the following individuals for very helpful comments on an earlier version of this article: Antonio Freitas, Peter Salovey, Wayne Steward, Eric Uhlmann, and especially Jonathan Haidt.

Correspondence concerning this article should be addressed to David A. Pizarro, who is now at Department of Psychology and Social Behavior, University of California, Irvine, 3340 Social Ecology II, Irvine, California 92697-7085. E-mail: dpizarro@uci.edu

moral cognition is over the role of deliberative reasoning. The distinctive claim of social intuitionism, as Haidt (2001) described it, is that “moral reasoning does not cause moral judgment; rather, moral reasoning is usually a post hoc construction, generated after a judgment has been reached” (p. 814), and that moral intuitions drive moral reasoning “just as surely as a dog wags its tail” (p. 830). With the notable exception of professional philosophers who have been “extensively trained and socialized to follow reasoning even to very disturbing conclusions” (p. 829), conscious deliberation plays little role in determining our moral judgments. To use Haidt’s metaphor, when we reason about moral issues, we are not like judges, considering the evidence and arguments in an objective search for the truth—we are like lawyers, trying to make a persuasive case for a preestablished point of view.

In contrast, rationalist theories, such as the one we defend below, agree that people possess intuitively given (and potentially sacrosanct) first principles—but posit that these serve as a starting point for deliberative reasoning, which can play an important role in the formation of moral judgments. Contrary to Hume, it is the rational dog that wags the emotional tail, not vice versa.

Educating the Moral Intuitions

Haidt (2001) considered two ways in which moral reasoning could, in principle, affect moral judgment. One is *reasoned judgment*, wherein a decision is made by sheer force of logic, overriding moral intuition. The second is *private reflection*, wherein, in the course of thinking about a situation, one activates a new intuition that contradicts an earlier judgment. The resulting moral judgment is based on which of the two intuitions is stronger or, alternatively, on the conscious application of a rule or principle. Haidt suggested that both processes exist, but that they play a small role in real-world moral judgment.

However, there is another potential process by which reasoning can influence intuitions. Prior reasoning can determine the sorts of output that emerge from these intuitive systems. This can happen through shifts in cognitive appraisal, as well as through conscious decisions as to what situations to expose oneself to. In both of these regards, prior controlled processes partially determine which fast, unconscious, and automatic intuitions emerge.

Cognitive Appraisal

There is a large literature pointing to the importance of cognitive appraisals in the arousal of quick involuntary responses (e.g., Lazarus, 1991). For instance, finding a telephone number in the pocket of one’s spouse can engender intense jealousy in one individual but mere curiosity in another, depending on how they construe the situation. You do not normally respond with fear when you hear someone start to whistle—but you might do so if it was 3 a.m. and you had thought you were alone in the house.

One of the most effective ways to change one’s intuitive moral responses, then, is to change one’s thoughts or appraisals about an issue. As an empirical demonstration of this, Dandoy and Goldstein (1990) showed participants three films of accidents that occurred in a factory setting. Participants who were instructed to adopt a detached, analytical attitude toward the films showed less physiological distress (as measured by galvanic skin response) than did individuals who received introductory statements that

merely informed them of the films’ content. Such changes of appraisal need not be directed by an external force but may be the result of a motivation to discover the facts of the matter or the result of a desire not to empathize with certain people.

One important instance of human beings’ cognitive flexibility is the ability to take the perspective of others. (This is quite similar to what Haidt (2001) called *private reflection*, though, unlike his proposed mechanism, it need not conflict with any preexisting intuitive judgment.) Usually, taking the perspective of a victimized other leads to an arousal of empathic emotions (see Batson, 1998, for a review). For instance, anger that a student failed to show up for an exam can quickly turn to sympathy when one discovers that the cause of the absence was a death in the family (Betancourt, 1990). In experimental contexts, one can get participants to be more empathic and more willing to help merely by asking them to try hard to take another person’s perspective (Batson et al., 1988).

This ability to shift one’s point of view has significant consequences. The empathy that occurs as a result of perspective taking may in turn fuel one’s initial reasons for taking the other’s perspective and could thereby lead to increased deliberation about moral principles. Indeed, empathy and reasoning about justice are linked from early on in the moral development of children (Hoffman, 2000). More generally, according to many philosophers, true ethical reasoning involves the capacity to transcend self-interest. Singer (1981) observed that this is a feature that all ethical systems appear to share. Examples (as cited in Singer, 1981) include the Golden Rule, attributed to Rabbi Hillel and repeated by Jesus—“Do unto others as you would have them do unto you.” Philosophers such as David Hume appealed to an “impartial spectator” as the test of a moral judgment. Utilitarians have argued that, in the moral realm, “each counts for one and none for more than one.” And the contemporary philosopher John Rawls presented a recipe for the creation of a just society, which is that one start with a “veil of ignorance,” not knowing which person one might become.

In this regard, Wright (1994) was only partially correct that “our ethereal intuitions about what’s right and what’s wrong are weapons designed for daily, hand-to-hand combat among individuals” (p. 328). This is true for the initial, adapted moral sense. However, as humans, we can modify our intuitions so that they move us in directions that actually oppose our material interests, as when we choose not to favor our own group over another or when we give up resources to help starving children thousands of miles away.

Control Over the Input

Aristotle (*Nicomachean Ethics*, 350 B.C.E., trans. 1998) first pointed out that individuals can construct the contingencies of their lives in order to exert control over their emotional reactions, a proposal also endorsed by James (1890/1950). This second-order control over emotional reactions and automatic judgments has recently received much empirical attention (e.g., Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Rudman, Ashmore, & Gary, 2001) and has called into question many of the original strong claims regarding the inability to exert control over automatic reactions made early on (Greenwald & Banaji, 1995).

One way that individuals may exert distal control over automatic reactions is through selective exposure to environments that

“educate” the moral intuitions. Researchers studying implicit attitudes have empirically documented the ease with which our implicit (i.e., fast, intuitive) judgments are manipulated by using a variety of common-sense techniques. For instance, participants exposed to positive African American exemplars, either through an experimental manipulation (Dasgupta & Greenwald, 2001) or because they chose to take a course on racism taught by an African American professor (Rudman et al., 2001), showed a significant reduction in implicit negative attitudes toward African Americans.

A more direct way to control the input is simply by choosing what to attend to. Our daily lives are filled with instances in which individuals exert second-order control over their impulsive, uncontrollable, or automatic reactions. Dieters may choose not to walk down the ice cream aisle of the grocery store for fear that they might succumb to temptation; smokers trying to quit might tell their friends not to give them any cigarettes, no matter how much they beg them later on; and so on (see Schelling, 1984). Even young children, when told that they can get a larger reward later on if they resist a smaller immediate reward, will consciously engage in tactics so as to distract themselves from the immediate temptation (Mischel & Ebbesen, 1970). In the moral domain, we often try to find situations that evoke our sympathies and avoid those that will elicit less positive reactions. When it comes to our loved ones, for instance, we often seek out information that would elicit our pride and approval; we would rather not be exposed to information that would evoke moral disgust. On the other hand, when it comes to gossip about our enemies, we sometimes have the opposite desire.

This last point illustrates that although second-order desires often serve a positive moral purpose, they sometimes do not, as when human beings (literally) avert their eyes to situations that might elicit their sympathy and intervention. Indeed, when participants are told that they are going to be shown a movie and then will be asked to help the person in the movie, they prefer to watch an “objective” movie rather than one that is more emotional and hence would appeal more directly to their sympathies (Shaw, Batson, & Todd, 1994).¹

The Scope of Moral Reasoning

We have discussed two ways in which reasoning can affect moral intuitions—by shaping the sorts of intuitions that occur and by controlling the situations that would elicit these intuitions. Haidt’s (2001) model of moral reasoning does not preclude these sorts of processes—indeed, his model allows for virtually all conceivable relationships between environment stimuli, deliberative reasoning, moral intuitions, and moral judgments. However, these processes pose a challenge to Haidt’s (2001) more general conclusions about the irrelevance of deliberative reasoning, as they raise the possibility that deliberative reasoning can affect moral judgment, albeit in an indirect fashion.

This raises the more general question as to the importance of these processes in real-world moral thought. After all, these processes do require prior moral reasoning and, according to the social intuitionist, the moral reasoning that takes place is typically after-the-fact justification.

When one looks outside the laboratory, however, there is considerable evidence that people do struggle with moral issues, trying to determine the right thing to do. Coles (1986), for instance,

documented the moral struggles faced by Black and White children in the American South during the Civil Rights movement, recounting the decisions that they made; Gilligan (1982) did similar work looking at young women who are deciding whether to get an abortion. And although many of our views might be shaped by our cultural context, there are innumerable instances in which people—not necessarily professional philosophers—take moral stands that put themselves very much at odds with members of their community. Examples include “righteous Gentiles” in Nazi Germany, children who insist on becoming vegetarians within nonvegetarian families, college professors who defend the abolition of tenure, and many pacifists during wartime. Also, modern humans often face moral issues that have not been anticipated by evolution or by the culture in which they were raised: Should research on stem cells be permitted? Should graduate students be permitted to unionize? Should the American government pay reparations to the descendants of slaves? These are not questions about empirical fact; they are questions of right and wrong.

A cynic might say that only a small minority make decisions about these issues—the rest of us just follow what they say. However, many moral issues are personal and have to be addressed by each individual in the course of his or her life: How much should I give to charity? What is the proper balance of work and family? What are my obligations to my friends? There are no “off-the-shelf” answers to these questions, no immediate gut reactions as to what is right and wrong. Haidt (2001) was likely correct that we do have quick and automatic responses to certain situations—killing babies, sex with chickens, and so on—and although these responses can be modified and overridden by conscious deliberation, they need not be. But most moral cognition is not about such simple cases; in the real world, moral reasoning is essential.²

In sum, the social intuitionist theory has considerable merits, but it misses some central aspects of moral cognition: Our immediate moral intuitions can be (and are) informed by conscious deliberation, and this deliberation plays a central role in our moral judgments.

¹ Along the same lines, people sometimes choose to act quickly so as not to allow these second-order desires to override their baser impulses. After World War II, the Belgians realized, on the basis of experience in the previous World War, that the punishments for collaborators would be more measured if the trials did not take place immediately. Collaborators who were tried immediately were often executed; this was less likely to occur after some time had passed and passions had cooled. For this reason, the Belgians wanted the trials to proceed as quickly as possible (Elster, 2000).

² All of this is fully consistent with Haidt’s (2001) repeated claim that deliberative reasoning is statistically a rare occurrence, that “most of our behaviors and judgments are in fact made automatically” (p. 819). Once a person has thought about, for instance, the morality of stem cell research, all subsequent responses to this issue might be fast and automatic, independent of any conscious reasoning. It is therefore possible that deliberative moral judgments are less frequent and occupy less of our time than nondeliberative ones. However, frequency and duration are poor cues to the importance of an event. A 30-second assault can profoundly change a person’s emotional life; a flash of inspiration can transform a society forever. The average American adult spends just over 4 minutes a day engaged in sex—which is almost exactly the time spent doing paperwork for the American government and 2% of the time spent watching television (Gleick, 1999).

References

- Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.
- Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, pp. 282–316). Boston: McGraw-Hill.
- Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., & Griffit, C. (1988). Five studies testing two new egoistic alternatives to the empathy–altruism hypothesis. *Journal of Personality and Social Psychology*, *55*, 52–77.
- Betancourt, H. (1990). An attribution–empathy model of helping behavior: Behavioral intentions and judgments of help-giving. *Personality & Social Psychology Bulletin*, *16*, 573–591.
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*, 828–841.
- Coles, R. (1986). *The moral life of children*. Boston: Atlantic Monthly Press.
- Dandoy, A. C., & Goldstein, A. G. (1990). The use of cognitive appraisal to reduce stress reactions: A replication. *Journal of Social Behavior & Personality*, *5*, 1275–1285.
- Darwin, C. (1998). *The expression of the emotions in man and animals*. Oxford, England: Oxford University Press. (Original work published 1872)
- Dasgupta, N., & Greenwald, A. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.
- Elster, J. (2000). *Ulysses unbound*. New York: Cambridge University Press.
- Fiske, A. P., & Tetlock, P. E. (1997). Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. *Political Psychology*, *18*, 255–297.
- Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, *117*, 39–66.
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Gleick, J. (1999). *Faster: The acceleration of just about everything*. New York: Random House.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
- Haidt, J. (2001). The emotional dog and its rational tail. *Psychological Review*, *108*, 814–834.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, England: Cambridge University Press.
- James, W. (1950). *The principles of psychology*. New York: Dover. (Original work published 1890)
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871–888.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*, 239–337.
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167–184.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pizarro, D. (2000). Nothing more than feelings? The role of emotions in moral judgment. *Journal for the Theory of Social Behaviour*, *30*, 355–375.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). “Unlearning” automatic biases: The malleability of implicit stereotypes and prejudice. *Journal of Personality and Social Psychology*, *81*, 856–868.
- Saletan, W. (2001, April 5). Shag the dog. *Slate.com*. Retrieved April 5, 2001, from http://slate.msn.com/FrameGame/entries/01-04-04_103801.asp
- Schelling, T. C. (1984). *Choice and consequence*. Cambridge, MA: Harvard University Press.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.
- Shaw, L. L., Batson, C. D., & Todd, R. M. (1994). Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology*, *67*, 879–887.
- Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. New York: Farrar, Straus & Giroux.
- Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–37.
- Turiel, E., & Neff, K. (2000). Religion, culture, and beliefs about reality in moral reasoning. In K. S. Rosengren, C. N. Johnson, & P. L. Harris (Eds.), *Imagining the impossible: Magical, scientific, and religious thinking in children* (pp. 269–304). New York: Cambridge University Press.
- Wright, R. (1994). *The moral animal: The new science of evolutionary psychology*. New York: Pantheon Books.

Received September 10, 2001

Revision received January 31, 2002

Accepted February 12, 2002 ■

Wanted: Your Old Issues!

As APA continues its efforts to digitize journal issues for the PsycARTICLES database, we are finding that older issues are increasingly unavailable in our inventory. We are turning to our long-time subscribers for assistance. If you would like to donate any back issues toward this effort (preceding 1982), please get in touch with us at journals@apa.org and specify the journal titles, volumes, and issue numbers that you would like us to take off your hands.