Evolutionary Psychology

www.epjournal.net – 2007. 5(2): 306-312

Book Review

Peering Underneath the Hood of Morality

A review of *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong* by Marc D. Hauser, HarperCollins, New York, 2006, xx + 489 pages, Hardcover, $27.95. ISBN: 0060780703

Peter DeScioli, Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104, USA. Email: descioli@psych.upenn.edu.

On July 2, 1938, Pope Pius XI issued an urgent encyclical endorsing a "holy crusade against the abuses of motion pictures" (Cortesi, 1936). The Pope enumerated several harmful consequences caused by bad movies:

> "Every one knows what damage is done to the soul by bad motion pictures. They are occasions of sin, they seduce young people along the ways of evil by glorifying the passions, they show life under a false light, they cloud ideals, they destroy pure love, respect for marriage and affection for the family. They are capable also of creating prejudices among individuals, and misunderstanding among nations, among social classes and among entire races."

The "war on 'evil' films" had been waged several years prior by an organization known as the Legion of Decency ("Cardinal declares war," 1934). A leader of the organization, Cardinal Hayes, condemned "lustful and depraved pictures" that featured "the glorification of crime, lust, and, in general, the serious violation of the law of God." Hayes pointed to the dire consequences that such movies would provoke: "History records the inevitable ruin of nation after nation when moral laws are violated and spiritual ideals vanish. Evil motion pictures in their baneful influence undermine the moral foundation of the State."

The causal claims made by the Legion of Decency were, to say the least, immodest. The organization hypothesized that eating popcorn at a showing of *The Outlaw*, a condemned film about Billy the Kid, could, by a diffuse but inexorable causal chain, result in the destruction of the soul, love, marriage, family, social harmony, and the nation itself.

The movies that raised eyebrows in the first half of the twentieth century are quite tame by today's film standards. Lengthy kissing, racially mixed romance, childbirth, and other innocuous scenes were forbidden by the Hays Production Code, which regulated

film-making from the 1930s to the 1950s. So was any portrayal of crime that might rouse sympathy with the criminal, such as *The Outlaw*. In today's theaters, *nearly all films* transgress taboos that were obsessively enforced by the Legion of Decency. Yet the sky never fell. In other words, the data are in, and the pontiff's pontifications have been roundly falsified.

The present aim is not to question the perils of Hollywood, to challenge Christian morality, or to shoot any other fish in a barrel. Rather, the Legion of Decency provides a convenient case study of moral condemnation – a subject that has inspired a flurry of psychological research in the last decade. The fascinating results of this research are the subject of *Moral Minds* by Marc Hauser.

In particular, a key starting point of Hauser's *Moral Minds* is the question of the psychological source of moral judgment. Namely, do moral judgments arise from cool reason, hot emotion, unconscious computations, or some combination of these? What makes the problem tricky is that people everywhere are quick to provide justifications of their moral judgments, usually by appealing to someone's welfare (self, family, friends, the group, etc.) or to a moral authority (divine, political, etc.). The tight correspondence of condemnation and rationale renders unclear the causal relation between these phenomena. Is condemnation caused by reasoning about consequences? Or does condemnation occur first, only afterwards motivating post hoc reasoning to justify one's moral position?

The Legion of Decency illustrates the puzzle. Did Pope Pius XI dispassionately assess the evidence about the likely effects of bad movies on the soul, love, family, etc., condemning such films only after their harmful consequences seemed well-substantiated? Or did the pontiff condemn first and ask questions later? That is, were the Pope's doomsday claims generated post hoc to justify his moral conclusions?

Moral Minds reviews research by Hauser and other moral psychologists (e.g., Haidt, 2001) that answers this question. Human moral judgment does not generally depend on reasoning; instead, people engage in moral reasoning post hoc to justify their condemnations. This fact opens a vista of possibilities regarding how human minds actually make moral judgments. Hauser insightfully identifies shortcomings of previous theories hinging on the emotions. He further offers a linguistic analogy to guide research into the underlying logic of morality and reviews his own original research motivated by this framework. So far, his experiments, which employ moral dilemmas, have produced tantalizing results. Along the way, *Moral Minds* points to the relevant evolutionary debates.

Not Reason

Consider this thought experiment. A psychologist asks Pope Pius XI to consider *The Moon is Blue*, a movie condemned by the Legion of Decency for, among other offenses, using the provocative words "virgin," "seduce," and "mistress." As expected, Pius condemns the film, and when asked to explain his judgment, he claims that the film damages the soul and society. Now, the psychologist asks Pius to use his imagination to consider, hypothetically, that the movie does not, in fact, cause harm to souls or societies. The Pope is asked whether the film is still morally wrong. He maintains his resolve, citing other harms caused by the bad movie. The psychologist repeats the procedure until all

harms have been hypothetically ruled out to the Pope's satisfaction. Is the movie still wrong? Exasperated, the Pope insists that the movie is still wrong, regardless of whether it causes harm or not.

This is not merely a thought experiment. Hauser explains that this clever procedure has been employed by psychologists to uncover an intriguing phenomenon termed *moral dumbfounding* (e.g., Haidt, 2001; Tetlock, 1999). Most participants in these experiments cite various harms as their rationale for condemnation. But when the worrisome harms have been ruled out, most people do not change their moral judgments. Instead, they insist on the wrongness of the action, while appearing dumbfounded by their inability to justify condemnation.

Hauser writes that moral dumbfounding betrays "a fundamental illusion in our psychology: conscious moral reasoning often plays no role in our moral judgments, and in many cases reflects a post-hoc justification or rationalization of previously held biases or beliefs" (p. 25). This discovery reveals a new landscape of possibilities for the operation of moral judgment. If condemnation does not hinge on conscious reasoning about harm, exactly how does it operate?

Not Emotion

Moral psychologists are still coming to terms with the discovery of moral dumbfounding and subsequent experimental confirmations of the associated insights (reviewed by Hauser). The initial interpretation was that the findings support the view that emotion rather than reason causes condemnation (e.g., Haidt, 2001). Hauser astutely points out that this explanation is incomplete.

Hauser asks, "How can we be sure that feeling caused our judgment as opposed to following from it?" (p. 7). He further clarifies that, "Neither we nor any other feeling creature can just *have* an emotion. Something in the brain must recognize – quickly or slowly – that this is an emotion worthy situation" (italics original, p. 8). Emotions figure prominently in our folk explanations of others' behavior, e.g., one might say, "she yelled because she was angry." But in the science of behavior, emotion has little explanatory depth. This is because among human sensori-motor processes, emotions lie far downstream towards action. A stipulated emotional cause of behavior invites further questions about the computations that triggered the emotion. Hauser writes, "The kind of emotion experienced follows from an unconscious analysis of the causes and consequences of action. This analysis, I argue, is the province of our moral faculty" (p.8).

Hauser discusses other limitations of accounts hinging on emotion. For example, he argues that these accounts can't explain the development of moral reasoning. He writes, "Our emotions can't explain how we judge what is right or wrong, and, in particular, can't explain how the child navigates the path between social norms in general and moral norms in particular" (p. 30). Here, Hauser is alluding to decades of research by developmental psychologists demonstrating that very young children have a nuanced understanding of the differences between morality and convention (e.g., Turiel, 1998). This understanding, Hauser argues, cannot be explained by children's experiences, emotional or otherwise.

Like Language?

Rather than reason or emotion, *Moral Minds* suggests that the source of moral judgment might be comparable to language grammar (following Rawls, 1971). That is, perhaps moral judgment arises from a suite of universal, innate, and unconscious computational mechanisms – a moral faculty.

For example, Hauser employs the linguistic analogy to suggest that in moral matters, like in language, "We know more than our actions reveal" (p. 37). It is well known that humans know much more about the grammar of their language than they (or even professional linguists) can consciously articulate. Specialized cognitive mechanisms quickly and automatically evaluate the grammar of utterances by employing sophisticated and intricate grammatical principles. These principles are a universal feature of *Homo sapiens* brains. After elaborate computations are performed, only a simple output is experienced by "user" consciousness: grammatical or ungrammatical. Though such judgments are beautifully nuanced and systematic across individuals and cultures, the operative principles behind the analysis remain entirely opaque to consciousness.

The same is very likely true of moral judgments. But it is also true of nearly all cognitive systems. Conscious awareness is the exception in mental life, not the rule. Breathing, catching a frisbee, assessing facial attractiveness, and falling in love are all enabled by sophisticated backstage computations that rarely enter consciousness. Why would morality be any different?

But along the dimension of function (the defining property of biological adaptations), morality is not like language. No more than a guillotine is like an alphabet or a raven is like a writing desk. Language is a computational machine for sharing mental representations with other humans. Morality is a computational machine for ganging up with others to attack, and sometimes execute, troublesome individuals. Morality is as much like language as it is like color perception or landscape aesthetics (both the subject of previous analogies), that is, not at all. Still, in the current state of perplexity about morality, broad-brush comparisons with other computational systems, such as the language faculty, might hold genuine didactic value, if not accuracy.

On the Horns of a Dilemma

But Hauser's *Moral Minds* goes beyond loose discussion of what morality is "like." The book lays out a strategy for mapping the universal logic of morality, and it describes the initial results produced by the approach. Linguists use grammaticality judgments of utterances to map universal grammar. Psychologists use "hot or not" judgments of facial attractiveness to map universal mate preferences. Analogously, Hauser suggests, "Students of moral behavior might begin by using ethicality judgments to uncover some of the principles underlying our judgments" (p. 43).

In particular, Hauser advocates the use of moral dilemmas as experimental stimuli. Philosophers have long employed this approach, using moral dilemmas to explore their own intuitions. Hauser uses moral dilemmas to great expository effect. He opens several sections by presenting the reader with pairs of moral dilemmas, predicting the reader's

judgments, calling attention to the reader's inability to explain their differing judgments, and then providing a plausible principle that might account for condemnation patterns. This tactic engages the reader and serves as a live demonstration of the principles at work.

In Part I "Universal Declarations," Hauser fleshes out his strategy of assaying moral intuitions with dilemma scenarios. In Chapter 2 "Justice For All," he analyzes decisions about helping those in need, suggesting over a dozen dimensions that might systematically influence judgments. He also examines ten different conceptions of fairness, wondering whether the particular conceptions active in a given culture might reflect different parameter settings on an underlying universal system (like Chomsky and others' views of language).

From this point on, Chapter 2 moves into material not obviously relevant to the agenda. Results from the Ultimatum Game are reviewed, but this game involves second-party revenge, not third-party moral condemnation. Revenge is probably underlain by its own computations, which are interesting, but not relevant to morality. Additionally, a puzzling claim is made that moralistic punishment is not costly among tribal societies. A single example is given (Wiessner, 2005), but a reasonable read of this paper does not support the conclusion that moralistic punishment is costless. Instead, the findings in this paper, as well as numerous other observations, establish the high costs of moralistic punishment (e.g., Knauft, 1987).

In Chapter 3 "Grammars of Violence," Hauser reports tantalizing data from his initial experiments using moral dilemmas ("The Moral Sense Test"). These experiments use variations of the Trolley Problem, in which a trolley will kill people on the tracks unless an actor intervenes. Typically, the actor is in a position to save the people but can do so only by killing someone else. Hauser turns up evidence that in moral judgments, actions are handled differently than omissions, scenarios with physical contact induce different evaluations than more distant acts, and that violations resulting indirectly from other actions are perceived as morally different from direct violations. All of these effects are remarkably consistent across thousands of individuals from different cultures and religious backgrounds. Further, many of them are not known by the participants making the decisions, just as the operating principles of grammar are often unavailable to introspection.

While Chapter 3 is nominally about grammars of "violence," it is not clear that the examined moral principles are specific to violence. The same principles might operate over other types of moral infractions such as theft or lying. This raises questions about the relevance of the end of the chapter, which discusses cultural variation in violent acts such as honor killing. Is violent behavior relevant to moral judgment of violent behavior? Given the many factors outside of morality that influence violent action, this link might be weak, particularly if moral judgment is governed by principles that are not specific to violence but cut across moral domains.

Seeming Tangents

Much of Parts II and III seem to drift from the strong arguments and framework constructed in the first third of *Moral Minds*. For instance, there is a review of what is known about perception of action and agency. But these systems probably work similarly

in all mammals, while no other species exhibits moral condemnation. Perhaps the claim is that these systems constrain moral cognition in interesting ways, but an argument along these lines is not articulated. There are also discussions of the self, theory of mind, and delay of gratification; it's not obvious that these systems are any more (or less) relevant to morality than they are to assessments of grammar or facial attractiveness.

Parts II and III also deal with many topics familiar to moral psychologists, such as reciprocal altruism, cheater detection, truth-telling, cooperation, and compassion. But according to the argument developed in Part I, these topics might not be immediately relevant to morality. In language, the rules of grammar don't differ across conversation topics, such as discussions of theft, sex, and fighting. Thus, understanding why people steal is not relevant to understanding the grammaticality of the sentences people utter about theft. If morality, like language, is underlain by a universal set of principles, these principles might operate uniformly across moral domains. Given moral wrong X, the system might assess various dimensions of the situation to determine whether a given action, neighboring around X, should be condemned. If so, particular details about how the mind handles sex, trade, honesty, etc. outside of moral judgment would be no more relevant to moral psychology than they are to linguistics.

Interestingly, this is what is suggested by the opening of Chapter 5, which discusses mother-fetus conflict. Those reducing morality to cooperation or bargaining might view this as relevant, but is it? Consider again these two phenomena: 1) a fetus employs a placenta to extract nutrients from its mother, who resists, and 2) the Legion of Decency morally condemns *The Moon is Blue* while claiming that evil movies can destroy souls and nations. Will understanding mother-fetus conflict inform our understanding of moral condemnation?

Evolved Function

Moral Minds is subtitled "How nature designed our universal sense of right and wrong." This subtitle suggests that the book will deal with the question of the evolutionary function of the moral faculty, including both its overarching function and how subcomponents operate in concert to subserve the overarching function. Yet treatment of this question is missing.

In a book similarly subtitled, "How the Aztecs designed their spears," one would expect to find an account of the function of Aztec spears (e.g., fishing, warfare, ritual) and a description of how the features of spears contribute to the function (e.g., a throwing spear might be specially balanced, a fishing spear might include a hook to retain the fish after stabbing). Less attention would be paid to general techniques of wood carving, even if these played a supporting role in the production of spears. Further, one would not want to "subtract out" the obsidian spearhead as a peripheral element on the grounds that Aztec knives also have obsidian blades (cf. Hauser's "subtraction method," e.g., pp. 49-54, 358, 411).

Moral Minds describes substantial advances in our understanding of the features of morality. It would be strange if increased knowledge about how morality operates *did not* cause revised theories of the function of morality. In fact, it would be worrisome. New data

showing that most Aztec spears are light and thin, contrary to previous views that they were heavy and thick, should cause revised theories of spear function, now favoring a projectile function over stabbing. If, on the other hand, functional accounts of spears are entirely insensitive to their empirically observed features, then theory is no longer informed by data.

It is surprising, then, that *Moral Minds* does not elaborate on the implications of its insights and data for theories of the evolved function of morality. There are several hints of Hauser's view on the subject, such as the passage: "What has allowed us to live in large groups of unrelated individuals that often come and go is an evolved faculty of the mind that generates universal and unconscious judgments concerning justice and harm" (p. 60). This echoes the popular view that morality is designed to stabilize cooperation in groups. What predictions does this theory make about the expected features of moral psychology? Would such a device handle actions differently than omissions, and if so, why? What about the other principles of moral logic described in *Moral Minds*? Do they provide evidence for or against various theories of the evolutionary function of right and wrong?

As details of moral cognition are uncovered, it should be possible to achieve greater resolution on the biological function of morality. *Moral Minds* has made great strides in this direction.

References

Cardinal declares war on 'evil' films. (1934, July 13). *The New York Times*, pp. 1, 14.

Cortesi, A. (1936, July 3) Pope orders world drive to raise film standards. *The New York Times*, pp. 1, 8.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834.

Knauft, B. M. (1987). Reconsidering violence in simple human societies: Homicide among the Gebusi of New Guinea. *Current Anthropology, 28*, 457-500.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Tetlock, P.E. (1999). Coping with trade-offs: Psychological constraints and political implications. In S. Lupia, M. McCubbins, & S. Popkin (Eds.), *Political reasoning and choice*. Berkeley: University of California Press

Turiel, E. (1998). The development of morality. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology* (pp. 95-130). New York: Academic Press.

Wiessner, P. (2005). Norm enforcement among the Ju/'hoansi Bushmen: A case of strong reciprocity? *Human Nature, 16*, 115-145.