

*Aversions, sentiments, moral judgments, and taboos*

**Richard Joyce**

Comment on Debra Lieberman's "Moral sentiments relating to incest: Discerning adaptations from by-products" Penultimate draft of paper appearing in W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 1: The Evolution of Morality: Adaptations and Innateness* (MIT Press, 2008): 195-204.

Nobody knows to what extent humans have innate mechanisms pertaining to incest regulation. Debra Lieberman is certainly interested in advocating the case that we do have evolved adaptations dedicated to governing this domain of behavior, but ultimately her position is that the positive hypothesis is "plausible" and that much more empirical work remains to be done. Jesse Prinz (this volume) advocates the anti-nativist position—in relation to both morality in general and incest in particular—but he, too, settles for the conclusion that the nativist case is "incomplete" and that there is "an exciting research program waiting to be explored." Arguments over various general and specific forms of human psychological nativism seem to generate a peculiar amount of entrenched intellectual acerbity (not in evidence, I hasten to add, in the two papers just mentioned), and it is a useful palliative to frequently remind ourselves of the common ground shared by all reasonable advocates: that we really don't know yet.

Empirical work of the kind undertaken Lieberman (and colleagues) is worth much more than any amount of armchair speculation. No matter how inventive and valuable such work is, however, there inevitably will be—as there should be—a body of opposition that will offer alternative models to explain the empirical data. Thus perhaps the most useful critical response to Lieberman's paper would be to take her data and suggest a non-nativist interpretation of them. But that is not what I intend to do in this brief commentary; rather, I will undertake a less ambitious, perhaps more pedestrian, but nevertheless equally important task: that of clarifying and paying careful critical attention to the words and phrases Lieberman uses to describe her results and conclusions. I will in particular criticize her use of the words "moral" and "sentiment." Such an enterprise will, I suppose, be rejected as being "merely a semantic argument," but semantic arguments *matter*: If one presents one's conclusions as casting light on the natures of Xs, but one is using "X" in an eccentric or restricted manner, then it may well turn out that one's conclusions do not concern Xs at all.

Lieberman chooses to frame her discussion using a rather quaint term: "moral sentiments." This phrase—seemingly more at home in an eighteenth-century treatise—appears to be undergoing a minor revival (see Slote 2003, Nichols 2004, Gintis et. al 2005), though I must confess that it is not clear to me what has been gained by substituting "sentiment" for the word it is so obviously designed to replace: "emotion."<sup>1</sup> Even Edward Westermarck—some

---

<sup>1</sup> Certain philosophers and psychologists have offered nuanced distinctions between emotions and sentiments (and passions, affects, etc.), but it seems fair to say that these are efforts at

of whose views Lieberman takes her findings to corroborate—eschewed the term “sentiment” in favor of “emotion.” It is true that arguments over what “emotion” denotes have become so entangled that in serious discussion one uses the word unqualified at one’s own risk, yet there seems little profit in substituting a fresh term when the inevitable outcome is that the intellectual battle lines over sentiments will be drawn in exactly the same places as they were when the argument was over emotions.

The phrase “moral sentiment” (just like “moral emotion”) is open to several interpretations; let me draw attention to three. Sometimes it is used as a rough synonym for “prosocial sentiment,” in which case love and sympathy may be thought of as moral sentiments. Clearly one can love someone without judging that it is morally right to love him, just as one can feel sympathy for someone without judging that one is duty-bound to care. Indeed, a creature constitutionally incapable of forming a moral concept might feel love or sympathy. Thus, if such sentiments count as “moral,” it may be only in the sense that they *merit* moral praise. Alternatively, some sentiments deserve to be called “moral” in virtue of the fact that they usually (or always) *involve* the making of a moral judgment. Many theorists argue that emotions necessarily involve a cognitive element, and, it would seem, in certain cases this cognitive element involves a normative judgment. Guilt, for example, is an emotion that necessarily involves the thought that the subject has in some manner transgressed.<sup>2</sup> Disgust can of course have nothing to do with any moral judgment (e.g., disgust at standing in dog feces), but seems also to come in a “moralized” form (e.g., disgust at concentration camps) (see Haidt et al. 1997; Rozin et al. 2000).

Moral sentiments may also be characterized in a third way: by reference to their subject matter, or the domain of their prototypical elicitors. This appears to be Lieberman’s route, for she uses the term “moral sentiment” to denote reactions prompted by or aimed at third party behaviors.<sup>3</sup> This interpretation certainly promises to make it easier to operationalize moral sentiments, but it is nevertheless problematic. For a start, there surely exist sentiments that are directed at third parties (just as frequently as they pertain to the self) that we would not ordinarily think of as particularly *moral*: surprise,

---

stipulating terms of art, rather than reflections of any vernacular distinction. In any case, someone who intends to use the word “sentiment” *in contrast to* “emotion” owes the reader an explanation.

<sup>2</sup> Cases of “survivor guilt” might be supposed to be counter-examples to this claim, but although such sufferers may know very well at a rational level that they are not responsible for the harm that befell others, they cannot shake the feeling that they have “done something wrong,” and it is precisely because their experience has this phenomenology that we are inclined to call their distress “guilt.” Saying that guilt necessarily involves an application of the concept *transgression* is not to say that the subject agrees *all-things-considered* that she has transgressed. For further discussion, see Sinnott-Armstrong 2005; Joyce 2006, chapter 3.

<sup>3</sup> In her opening paragraph she asks “What is the origin of our sentiments regarding 3rd party behaviors, *that is*, what is the origin of our moral sentiments?” She makes the same identification later, referring to “3rd party behavior – the moral domain” (where one must assume that the dash indicates an “i.e.”). The same connection is made in her 2003 paper with John Tooby and Leda Cosmides.

horror, and pity, for example. Conversely, there are certain sentiments usually classified as “moral” which seem very self-regarding—guilt and shame being the obvious ones. Other sentiments often classified as moral—contempt, punitive anger, moral disgust—seem to come equally in self-regarding and third-party-regarding forms, with neither having an obvious claim to being prototypical. (See Haidt 2003 for good discussion of moral emotions.) Lieberman’s restriction of the moral realm implies the unfortunate result that if a person is disgusted at his own incestuous thoughts this is not a moral sentiment, and if he is disgusted at his sister’s sexual advances this is not a moral sentiment, but if he hears tell of an incestuous union between another brother and sister, only then will his subsequent disgust count as moral.

It is important to bear this restriction in mind while reading Lieberman’s paper, or else matters become quickly confusing. For example, one might have been tempted to see Lieberman’s contrast between aversion to incest and incest-denouncing moral sentiments as a contrast between different possible mechanisms of self-regulation. There are, on the one hand, various possible self-regulatory mechanisms with *non-moral* outputs: (i) an absence of any sexual desire towards kin (i.e., sexual indifference), (ii) a positive desire to refrain from such activities (i.e., sexual aversion), (iii) a negative emotional response at the thought of committing incest (self-disgust, repugnance, etc.).<sup>4</sup> On the other hand, one might have a *moral* resistance to one’s own potential incest: (iv) judging it to be forbidden, the expression of a vice, as deserving reprimand or punishment. Moral sentiments, in the second sense identified above, blur the boundary between these (iii) and (iv), because they are to be distinguished both from non-moral emotional opposition to incest (e.g., non-moralized versions of anger, repugnance, fear) and from non-emotional moral condemnation of incest (e.g., agreeing that it is morally prohibited, but feeling emotionally unruffled when contemplating it). One might have expected from the title of Lieberman’s paper that the discussion would focus on whether any such “moralized” emotions elicited by incest are adaptations or by-products of adaptations. But this is not her project, and her distinction between aversion and moral sentiment is not the one just canvassed. Rather, the term “aversion” seems to do the work of denoting the output of any psychological mechanism that decreases an individual’s motivation to engage in incest, while “moral sentiment” is reserved for the output of any mechanism that prompts the individual to interfere with other parties’ (potential) incestuous activity.

This manner of classifying matters, in my opinion, masks much that may be of interest. For a start, even when confining ourselves just to *self*-regulation with respect to incest, it would be useful to distinguish moral from non-moral mechanisms. There is, after all, a world of difference between simply *not wanting* to have sex with someone and judging such sex to be a disgusting

---

<sup>4</sup> The distinction between (ii) and (iii) assumes that desires and emotions are distinct phenomena, which seems a safe assumption, widely upheld in both vernacular and theoretical discussions. Certainly many or perhaps all emotions implicate desires, but we usually allow that one can have a desire (e.g., to one day read *A la recherche du temps perdu*) without any *emotions* being engaged in favor of the desired outcome.

evil. Natural selection may have had a hand in designing either or both of these motivation-engaging mechanisms, and it would be interesting to see the matter investigated. Darwin himself considered the moral sense to be a faculty of *self-regarding* moral appraisal (i.e., the moral conscience), and claimed that it is “of all the differences between man and the lower animals ... by far the most important” ([1879] 2004: 120). Westermarck, in contrast, argued that self-directed moral evaluations are circuitously reached only “through a prior critique upon our fellow-men” (1906: 123).

By the same token, it would be useful to distinguish possible moral from non-moral negative responses to others’ incestuous behavior. There is, again, a big difference between *not wanting* others to commit incest and judging any such union to be a moral abomination deserving of punishment. Even if one’s opposition to others’ incest manifests itself as directed anger, and even if this anger produces a retaliative reaction, there need be nothing especially moralistic about the anger or the retaliation. It is only if the anger incorporates or is accompanied by the thought that the third party has *transgressed against a norm*, and the retaliative response motivated or justified by the thought that the punishment is *deserved*, that the opposition counts as moral. It is possible that humans have a dedicated innate mechanism that generates some form or other of non-moral opposition to third party incest, while the widespread human trait of moral opposition is a by-product or a cultural artifact. It is also possible that certain forms of moral opposition are also the product of another (complementary) innate mechanism. Again, it would be extremely interesting to see more empirical work aimed at resolving this matter.

It is not my intention to criticize Lieberman for failing to pursue these research questions; rather, I am just trying to tease apart these distinct research projects and get clear on the misleading terminology involved. To put my point strongly (perhaps overly so): most of Lieberman’s discussion might have proceeded without the word “moral” even appearing. What she is really interested in is mechanisms that motivate humans to act against (potential) incestuous relations among third parties (are they adaptations or by-products?), and the relation these mechanisms have to any other mechanisms that motivate humans to themselves avoid incest. Whether either of these mechanisms affect motivation by generating desires, or engaging emotions, or prompting moral judgments, or some combination of the above (“moral sentiments”), is not a matter that her experiments are designed to discriminate.

That Lieberman’s conceptual space is coarse-grained in this respect is evidenced by the way that very different forms of “psychological opposition” (to incest) are treated as being on all fours. It is claimed that behaviors negatively affecting one’s fitness are ones we tend to categorize as “morally repugnant,” whereas ones that enhance fitness are considered “morally virtuous and praiseworthy.” Already there are difficulties, for “moral repugnance” sounds very much like a kind of occurrent moralized emotion, whereas classifying something as “morally praiseworthy” need involve no emotion at all. In the next sentence Lieberman refers to “moral disapproval and approval”—terms that are (perhaps usefully) indeterminate. On the one

hand, they are sometimes used to stand for certain moral emotions (Westermarck, for example, insists on this usage<sup>5</sup>); on the other hand, someone who holds that moral judgments may be entirely cognitive and non-emotional (i.e., a Psychological Rationalist<sup>6</sup>) can, without straining language in the least, agree that moral evaluations are ways of “approving” or “disapproving.” Later in the paper Lieberman reports her 2003 study (with Tooby and Cosmides) in which subjects were asked to rank actions in terms of their “moral wrongness,” and yet the experiment is taken to corroborate something about “moral sentiments.” It is, again, not clear to me what is gained by introducing this loaded and problematic term; if subjects are asked questions framed explicitly in terms of “morally wrongness,” then why not report the findings as concerning just their “moral *opinions*” or “moral *assessments*”? Contrary to what is claimed in the earlier article, the experiment does not “directly measure moral sentiments” (Lieberman et al. 2003: 825), unless “sentiment” is misleadingly being used in an extremely watered-down manner to mean nothing more than “moral opinion.”<sup>7</sup>

Let me close by saying something about another word that often appears in these discussions, and, indeed, crops up in Lieberman’s paper: “incest *taboo*.” Westermarck’s position is that the taboo against incest is a by-product of an innate aversion; he claims that humans’ natural aversion to incestuous relations “displays itself in custom and law as a prohibition” (1921: 193). Lieberman’s position is sympathetic to Westermarck’s, but she also speculates (on some evidential ground) that “moral sentiments” relating to family members’ incest (i.e., motivational systems targeting family members’ incest) are the output of a discrete adaptation. By contrast, Prinz (this volume) pushes the anti-nativist agenda, arguing that the evidence in support of an innate incest taboo is “less secure” than it is widely assumed to be. These seem to be three competing positions, but I would like to suggest that on a closer reading the three authors are to some extent talking past each other.

Note, to begin with, that a taboo is most naturally interpreted as a trait instantiated by *a group*; there’s something very peculiar about the idea of a person having her own personal taboos at odds with those of the population of which she is a member.<sup>8</sup> If this is correct, then a taboo can be innate only if we are willing to appeal to some model of multi-level selection whereby adaptations may be properties instantiated by groups rather than individuals. I hazard to suggest that those people (like Prinz) who frame the debate explicitly in terms of whether there is an “innate taboo” against incest

---

<sup>5</sup> See Westermarck 1932: 63 ff.

<sup>6</sup> See my contribution to this series (vol.3), for a characterization of different forms of moral rationalism.

<sup>7</sup> Even here, of course, the experiment doesn’t *directly* measure moral opinion, but rather measures subjects’ self-reports of their moral opinion.

<sup>8</sup> This is not to say that we cannot make any sense at all of the thought of a person having her own personal taboos. We *can* make some sense of the thought, just as we can make *some* sense of the thought of a person having her own private language. It suffices to make the point that whatever comprehension we seem to have of personal taboos is surely heavily derivative on our grasp of the idea a group-held taboo, to such an extent that it seems a semi-metaphorical usage.

probably do not intend to invoke any such models, and thus should probably rephrase their terms.

What Prinz presumably has in mind is the question of whether it is a human adaptation to think of incest as morally wrong, or forbidden, or sinful (or some other morally-loaded form of opposition). If everyone in a population had such an adaptation, then it is natural to think that this would manifest itself as an incest taboo (but it would still be a mistake to speak of the taboo itself being an adaptation). There are two things to note about Prinz's target question in relation to Lieberman's concerns. First, Prinz is very clear that he is discussing a *moral* form of opposition. Such opposition is related to, but nevertheless is to be distinguished from, mere *dislike* (or even hatred) of incest, and distinguished from any un-moralized emotional response (e.g., repugnance) towards incest. Second, Prinz makes no distinction between moral opposition that is self-directed and that which is other-directed. Thus Prinz is making a distinction that Lieberman does not make, and Lieberman is making a distinction that Prinz does not make.

Westermarck is interested in the relation between the human individual's aversion to incest (which he thinks is innate) and the social taboo (which he thinks is a by-product of the former). He doesn't "moralize" the self-directed aversion; he doesn't argue that we are designed to think of sex with our own family members as *evil* or even *prohibited*; rather, we are designed just to find the prospect repellant. But, he holds, when a group of like-minded humans form a society, then this aversion that each has to the thought of having sex with his or her own family members will manifest itself as a moral prohibition within the group, as a taboo. (Quite why this would occur is something of a mystery, which Arthur Wolf, in his extremely useful discussion, calls "the representation problem": "The fact that early association inhibits sexual attraction explains why most people avoid sexual relations with their parents and siblings, but it does not explain why they condemn other people for having sexual relations with *their* parent or *their* sibling" (2005: 11), nor does it explain why this condemnation should have "a strongly felt moral content" (12).<sup>9</sup>) Lieberman seeks to corroborate Westermarck's "by-product

---

<sup>9</sup> Note that this problem is not the same as that identified by Prinz (this volume): that "If we naturally avoid something, we don't need a moral rule against it." Prinz's objection to the hypothesis that humans are innately predisposed to morally condemn incest is closely related to an objection that has been leveled many times at Westermarck's claim that the incest taboo is a by-product of an innate aversion—first by Sir James Frazer in 1910. Wolf notes that the "argument has been repeated, mantralike, by Westermarck's many critics" ever since (2005: 5). But in fact Westermarck offered a perfectly cogent rejoinder in *The History of Human Marriage* (1921, vol.2: 203ff), which, it seems to me, serves just as convincingly against Prinz as it did against Frazer. Of course there are many instances of natural aversions for which no moral condemnation is necessary (Prinz lists some), but we can equally well think of cases where a natural aversion engages motivation only imperfectly, and where an internalized moral imperative may usefully supplement the aversion. One advantage of moral opposition is that it can be used to *justify* punishment in a way that dislike, or even emotional abhorrence, cannot. See Joyce 2006, chapter 4, for further discussion of this point.

hypothesis”<sup>10</sup> by showing that one’s attitudes to third-party incest are a function of one’s levels of aversion towards one’s own potential incestuous relations. But this seems to be investigating a different relation from the one Westermarck proposed. His claim was that an incest taboo as a societal trait would arise from individuals’ collective self-directed aversions, not that an individual’s attitudes to third-party incest would arise from an individual’s self-directed aversions. The latter relation could be empirically confirmed—we could, for example, discover that the stronger an individual’s aversion to her own potential incest, the stronger her dislike of others’ incestuous relations—while nothing is revealed about the origin of the society’s moral proscription.

The extent of disorder among these three apparently competing views is brought home when we realize that they all could be true. Suppose humans have an innate aversion to committing incest, and this aversion, when individuals join together to form a society, manifests itself (somehow) as a moral taboo. In other words, suppose Westermarck is correct. Since the innate aversion is not “moralized,” and since what is moralized (the taboo) is not innate, Prinz’s anti-nativist view of morality (and the morality of incest in particular) is consistent with Westermarck’s view. Now suppose also that humans have a discrete innate mechanism designed to motivate action in response to perceived third-party incest. In other words, suppose that Lieberman is correct. This clashes with no part of Westermarck’s view. Lieberman calls any such motivation-engaging responses “moral sentiments,” but I have suggested that the word “moral” here is at best optional, and at worst misguided. In any case, it is not the same sense of “moral” that Prinz employs when he denies that humans have any innate moral attitudes towards incest. Thus, it turns out that—a terminological discrepancy aside—Lieberman’s view and Prinz’s view could both be correct.

Lieberman writes: “Morality, traditionally the province of theologians and philosophers, has been invaded by scientists.” The invasion is wholly to be welcomed. But if the assorted thinkers and researchers are to avoid talking past each other, it is as well that we all pay careful critical attention to the subtleties of the language we use in attempting to describe various phenomena.

## References

- Darwin, C. (1879) 2004. *The Descent of Man, and Selection in Relation to Sex*. London: Penguin Books.
- Gintis, H., Bowles, S., Boyd, R.T. & Fehr, E. (eds.) 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press.

---

<sup>10</sup> This is to be distinguished from what Lieberman calls “the Westermarck hypothesis,” which is the claim that childhood association is a trigger for sexual aversion. Nothing I say in this critical response bears on the Westermarck hypothesis.

- Haidt, J. 2003. "The moral emotions." In R.J. Davidson, K.R. Scherer, & H.H. Goldsmith (eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press: 852-870.
- Haidt, J., Rozin, P., McCauley, C., & Imada, S. 1997. "Body, psyche, and culture: The relationship of disgust to morality." *Psychology and Developing Societies* 9: 107-131.
- Joyce, R. 2006. *The Evolution of Morality*. MIT Press.
- Rozin, P., Haidt, J., & McCauley, C.R. 2000. "Disgust." In M. Lewis & J.M. Haviland-Jones (eds.), *Handbook of Emotions*, 2nd edition. New York: Guilford Press: 637-653.
- Sinnott-Armstrong, W. 2005. "You ought to be ashamed of yourself (when you violate an imperfect moral obligation)." *Philosophical Issues* 15: 193-208.
- Slote, M. 2003. *Morals from Motives*. Oxford University Press.
- Westermarck, E. 1906. *The Origin and Development of the Moral Ideas*, vol. 2. London: MacMillan and Co., Ltd.
- Westermarck, E. 1921. *The History of Human Marriage*, vol. 2. London: MacMillan and Co., Ltd.
- Westermarck, E. 1932. *Ethical Relativity*. London: Kegan Paul, Trench, Trubner & Co., Ltd.
- Wolf, A.P. 2005. "Introduction." In A.P. Wolf & W.H. Durham (eds.), *Inbreeding, Incest, and the Incest Taboo*. Stanford, CA.: Stanford University Press: 1-23.