

How to Be Fair to Psychopaths

Shaun Nichols

Department of Philosophy

University of Arizona

sbn@email.arizona.edu

and

Manuel Vargas

Department of Philosophy

University of San Francisco

mrvargas@usfca.edu

keywords: fairness, moral responsibility, psychopaths, revisionism

Shaun Nichols is Professor of Philosophy at the University of Arizona. His recent research concerns experimental philosophy, cultural evolution, free will, and cognitive theories of the imagination. He is the author of *Sentimental Rules: On the Natural Foundations of Moral Judgments* (Oxford, 2004) and (with Stephen Stich) *Mindreading* (Oxford, 2003).

Manuel Vargas is Associate Professor of Philosophy at the University of San Francisco. His research interests include moral psychology, moral responsibility, evil, and topics in Latin American philosophy. He is author (with John Fischer, Robert Kane, and Derk Pereboom) of *Four Views on Free Will* (Blackwell, 2007).

How to Be Fair to Psychopaths

Shaun Nichols and Manuel Vargas

Neil Levy's provocative paper raises a number of fascinating issues. Here we want to focus on just one of these – the role of principles concerning fairness in his basic argument that we shouldn't punish psychopaths. For present purposes, we will simply go along with Levy's claim that psychopaths lack moral knowledge (though see Vargas & Nichols 2007).

In the background of Levy's central argument is the assumption that a person isn't morally responsible for things which fall entirely outside his scope of control. This assumption shows up in the following claim: "If an agent comes to be bad through a process that entirely bypasses her ability to appreciate and to respond to reasons, including moral reasons, she is not a responsible agent at all" (MS, 13). Psychopathy is a wonderful example here, since there's reason to think it has a strong genetic component. But why should we accept his claim that we have to absolve those who are born irrevocably bad?

It would seem that Levy here appeals to a familiar and powerful intuition that it is *unfair* to blame and punish an agent for a characteristic (or behavior) that was caused by factors outside the agent's scope of control. For instance, Levy writes, "It would be *unfair* to blame me for my bad art if I lack talent, since there was nothing I could reasonably have been expected to do to make myself a good artist" (MS, 11, emphasis in original). Let's give this principle a label:

Unfairness: It's unfair to punish someone for a characteristic (or behavior) that was caused by factors outside the agent's scope of control.

That is a powerfully intuitive claim. Levy uses this principle to argue that, since psychopaths' deficiencies in moral judgment were caused by factors outside their scope of control, it's unfair to punish psychopaths for the behavior that issues from these deficiencies. But there is another powerfully intuitive claim in the immediate vicinity:

Fairness: It's (sometimes) fair to punish psychopaths.

As Watson (1987) illustrates beautifully with the case of Robert Harris, we find a claim like *Fairness* especially intuitive when provided with details about the crimes of the psychopath. If *Unfairness* together with the facts about psychopathy implies that it's unfair to punish psychopaths, then another option is to run a modus tollens. That is, we can fix on *Fairness* (along with other similar considerations) and use that as a basis for rejecting *Unfairness*. So here we have an interesting, and we think quite deep, question: Should we sustain *Unfairness* or *Fairness*?

Levy doesn't seem to consider the option of revising the principles that lead to his conclusion that psychopaths aren't responsible. He writes, "We are not here concerned with whether it is appropriate to act *as if* psychopaths were responsible. We are instead concerned with a factual question: whether they *are* responsible for their actions" (fn 2). Given that the question is whether psychopaths are *morally* responsible, it is not obvious that this is indeed a factual question. But the more important point is that, while we often use our principles to guide judgments about cases, we sometimes use cases to drive revisions of principles. Consider first a familiar example from scientific classification. In the 19th century, the category of mammal was largely structured around the principles that mammals bear live young and have milk glands. The discovery of the platypus raised a problem because the platypus has the latter trait but not the former. The result, of course, was to revise the principles of mammalhood to encompass the platypus. While most mammals bear live young, taxonomists decided that this is not a necessary condition.

In the normative domain, we also find case-driven revisions of principles. For centuries, it was a central assumption that an artwork had to be crafted by the artist. Duchamp's ready-mades violated this assumption with great aplomb. The artworld revised the principles to make room for Duchamp. Finally, and of most direct relevance, sometimes the principles governing attributions of moral responsibility themselves are subject to revision. In the 17th century, it was presumably commonly thought to be unfair (or at least inappropriate) to punish an individual that lacks a Cartesian soul. That principle has been widely given up in contemporary times. Most philosophers who deny that people have Cartesian souls still maintain that it is fair to punish people. Moreover,

virtually all contemporary compatibilists would maintain that we *should* reject the principle that it is unfair to punish those who lack Cartesian souls.

Thus there are clear precedents for revising our principles of classification in light of new information about cases. In the present context, then, it seems entirely possible that the correct response to the evidence on psychopathy is to withdraw our commitment to *Unfairness* and preserve our view that psychopaths are (sometimes) blameworthy. Levy owes us a substantial argument for why we should move from *Unfairness* to the rejection of *Fairness* rather than appeal to *Fairness (inter alia)*, as a basis for rejecting *Unfairness*.

One obvious consideration here is that we want to have some way to distinguish between factors that do undermine responsible agency, and thus, blameworthiness, and those that don't (MS 13-14).¹ Borrowing from Reznik's tumor thought experiment (see Levy MS 15), we can imagine cases of brain tumors that would count as completely undercutting an agent's blameworthiness for a behavior. For instance, if a person has a tumor that generates behavior in a way that bypasses reasons entirely, it seems wrong to punish that individual for that behavior. And we want to have principles that would allow us to excuse those behaviors. But *Unfairness* isn't the only resource we have for that. Over the last several decades, a number of alternative principles have emerged that might justify treating the behavior of psychopaths as blameworthy while also excusing the behavior of many other individuals. Levy mentions attribution theories (e.g. Scanlon), but one might also reject *Unfairness* and adopt a version of hierarchical compatibilism (Frankfurt 1988), "reason" compatibilism (Wolf 1990), or reason-responsiveness compatibilism (Fischer & Ravizza 1998). Now, we stress that we are not here endorsing the details of any of these theories. Moreover, we are not suggesting that all of our intuitions would be captured if we give up *Unfairness* and adopt one of these versions of compatibilism. On the contrary, we grant that *Unfairness* better captures our intuitions. The point is simply that giving up *Unfairness* does not leave one bereft of all resources for excusing certain tumor victims while blaming psychopaths.

Thus, Levy's argument might be usefully cast as a dilemma. Either (i) it is unfair to punish psychopaths or (ii) it is sometimes fair to punish people for behavior that was caused by factors

outside their scope of control. As we have noted, one might evade the dilemma by revising one's principles to accept (ii). Now, is there any reason, apart from the pressure from psychopathy, that would indicate that we should accept (ii)? This is a difficult matter to settle, but one consideration is that the principle *Unfairness* is precisely the kind of principle that leads incompatibilists to conclude that no one is fully morally responsible (e.g. Pereboom 2001, Smilansky 2000). If incompatibilists are right about this, then psychopaths are by no means an exceptional case. The *Unfairness* principle, on this view, leads to sweeping exemptions for everyone. However this would, we think, provide a powerful reason in favor of giving up on the *Unfairness* principle. Of course, the status of incompatibilism is enormously controversial. But in light of these considerations, we would very much like to know the kinds of reasons that move Levy to think that the proper way to resolve the dilemma is to conclude that it's unfair to punish psychopaths.

References:

Fischer, J. & Ravizza, M. 1998. *Responsibility and control : A theory of moral responsibility*.

Cambridge: Cambridge University Press.

Frankfurt, H. 1988. *The importance of what we care about: Philosophical essays*. Cambridge:

Cambridge University Press.

Pereboom, D. 2001. *Living without free will*. Cambridge: Cambridge University Press.

Smilansky, S. 2000. *Free will and illusion*. Oxford: Oxford University Press.

Vargas, M. 2006. The importance of history for responsible agency. *Philosophical Studies*, 127, 351-382.

Vargas, M. and S. Nichols 2007. Psychopaths and moral knowledge. *Philosophy, Psychiatry, and Psychology* (this volume): pp-pp.

Watson, G. 1987. Responsibility and the limits of evil: Variations on a Strawsonian theme. in *Responsibility, character, and the emotions: New essays in moral psychology*, ed. F. Schoeman, Cambridge University Press.

Wolf, S. 1990. *Freedom and reason*. Oxford: Oxford University Press.

ⁱ There is some debate in the literature about whether or not an agent's history is at all relevant to assessments of moral responsibility. We think it is, but perhaps less frequently so than one might suppose (cf. Vargas 2006). What every account needs, historical or not, is an account of excusing conditions (or, alternately, an account of the conditions for responsible agency).