

Session 15 Cushman, "Crime and Punishment: Distinguishing the roles of causal and intentional analyses in moral judgment"

I. On Moral Luck:

Control Principle: People cannot be morally assessed for what is due to factors beyond their control.

Corollary to the CP: Two people ought not to be morally assessed differently if the only other differences between them are due to factors beyond their control.

Moral Luck: Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called *moral luck*.

Some initial examples:

- We think manslaughter is worse than reckless driving that does not result in a death. At least, we wouldn't want to let the killer off with a fine, or punish the simple reckless driver with a lengthy jail term. We've probably all done things (nodded off for a second at the wheel, not checked our blind-spot, etc.) that could have led to death and a manslaughter charge.
- We may think murder is worse than attempted murder.
- Someone who was an officer in a concentration camp might have led a quiet and harmless life if the Nazis had never come to power in Germany (assume, plausibly, that this was something he could not control), but we still judge him harshly.

There are lots of ways morality might seem to depend on luck: luck in how our actions and projects turn out; luck in the kind of person we turn out to be (our inclinations, capacities, and temperament); luck in the circumstances we're born into or happen to face (*situationism, Milgram experiments*)...

Problem: Because of the many ways in which the shape of our lives and intentions and effects of our actions are beyond our control, if the condition of control is consistently applied, it threatens to erode most of the moral assessments we find it natural to make.

So is this just evidence that the CP is false?

At the very least, we should recognize that it has implications far beyond drunk driving and attempted murder cases. It may also, for example, tell us to withhold blame from the SS officer. If we want to resist that, we'll have to:

- find some way of drawing a principled distinction between the drunk driving case and the SS officer case, to explain why we can appropriately judge the latter more harshly than we judge ourselves, but can't appropriately judge the drunk driver more harshly if he hits someone.

- This distinction can't rely just on the CP.

Some Questions:

- *One big question here, especially relevant to assessing Cushman's article, is how much we can learn about what, e.g., blameworthiness, desert of punishment, wrongness, impermissibility require by looking at the judgments people actually make about those categories.*
- *Why do we accept that bad moral luck might make someone morally worse, but not that bad luck can turn an otherwise innocent action into a morally bad one? (E.g. If I have the bad luck of hitting a child who suddenly runs in front of my car to chase a ball, my bad luck doesn't make my action more blameworthy (of course, it makes it more unfortunate). But if I'm already behaving negligently – say, driving drunk – and I have the bad luck of a child happening to be on the sidewalk at just the place where I drive up onto it, that seems to make my already blameworthy action more blameworthy. Why this difference?)*

Some Replies to the Problem of Moral Luck:

- *Rejecting the Control Principle.*
- *Partial defenses of moral luck:* we might offer targeted defenses of some kinds of moral luck – e.g., we might say it's appropriate to judge people for what they do, if it results from who they are (and if they have control over what they do), even though they cannot control who they are.
- *Epistemic arguments:* we're justified in treating cases of bad resultant luck and bad circumstantial luck differently from cases of good luck because we can't know, e.g., of the drunk driver who doesn't hit the pedestrian whether he would have noticed the person and stopped had she been there; can't know how we might have acted in, e.g., Milgram's experiment; can't know whether the murder failed because the attacker wasn't fully committed to carrying through.
- *Arguments rejecting Moral Luck (bad luck in the moral status of an action) but defending differential treatment of unlucky actions (e.g., through blame, punishment):*
 - *Pragmatic arguments:* perhaps we have pragmatic reasons for, e.g., treating such cases differently in the law (because we can't lock up every drunk driver, but we can threaten every drunk driver with locking them up should they be unlucky – consider also Lewis' "The Punishment that Leaves Something to Chance", which we'll read next).
 - *Arguments based on the purpose of punishment:* Satisfying the requirements of desert may not be the only factor determining what level of punishment is appropriate; indeed, perhaps it *can't* be the only factor (since considerations of desert may provide no non-arbitrary way of fixing the level of punishment). Other purposes of punishment (e.g., deterrence, compensation for the victim, prevention of further crimes by the same person, etc.) *may* provide reasons for treating unlucky agents differently from lucky ones (this seems particularly clear in the case of luck in who we turn out to be, and also with regard to punishment to secure victim compensation, in the case of luck in how our actions turn out).

II. On Cushman

Cushman's general findings (p. 375):

- Judgments of wrongness and permissibility are overwhelmingly determined by an analysis of culpable mental states, while judgments of deserved punishment and blame show relatively enhanced sensitivity to an analysis of causal responsibility.
 - This finding contradicts the commonsense notion that acts are punished if and only if they are wrongful.
- Second, judgments of deserved punishment for failed attempts to harm are mitigated when the intended harm coincidentally occurs by some independent mechanism, relative to when the harm does not occur at all.
 - This result contradicts the commonsense notion that moral judgments of an agent depend only on whether the agent causes a harm, and not on whether the harm happens to occur independently.

Cushman makes a very strong claim: he purports to have undermined the “commonsense notion” the punishment is warranted if and only if an agent has behaved wrongfully. *Do you think his experiments show this? Why or why not?*

Four Experiments:

Experiment 1:

- (Jenny the welder) Tested the relative contribution of facts about what agents *believed* would follow from their action, facts about what agents *desired* would follow from their action, and facts about what *actually* followed from their action (i.e., the action's *consequences*) to subjects' assessment of the action's (a) *wrongness* and (b) *blameworthiness*.
- *Results:*
 - Both kinds of judgment were more sensitive to *agents' beliefs* than to *agents' desires*. (*Question: Does this reflect the fact that the scenarios concerned harms as opposed to benefits? Some studies suggest we're more willing to assign agents responsibility for un-desired but foreseen harms than for un-desired but foreseen benefits.*)
 - Judgments of *wrongness* did not depend significantly on the *consequences* of the action.
 - Judgments of *blameworthiness* did depend significantly on the *consequences* of the action.

Experiment 2:

- Tested whether the results of Experiment 1 would be mirrored by our judgments concerning (*im*)*permissibility* and *appropriate punishment*.
- *Results:*
 - Judgments of *impermissibility* closely tracked judgments of *wrongness*, and were not sensitive to consequences.

- Judgments of *appropriateness of punishment* tracked judgments of *blameworthiness*, and were significantly sensitive to consequences.
- Shows the differential judgments concerning the relevance of consequences to wrongness and blameworthiness don't merely reflect a non-normative (purely causal) understanding of being "to blame."

Cushman's Explanation of the results of Experiments 1 & 2:

- (i) The causal process of moral judgment is only fully engaged in the presences of a harmful consequence and
- (ii) when it fails to be activated, the mental-state process competitively dominates resultant judgments of deserved punishment and blame.

Some Questions about these experiments, and Cushman's conclusions:

- How are we to understand the *belief*-factor? What counts as having the relevant belief? For example, is the drunk driving case which Cushman describes at the outset one in which the agents have or lack the relevant belief about the harmfulness of their actions?
 - They probably don't believe their driving drunk will result in an accident when they decide to drive drunk, but they may well have believed, in advance of drinking, that driving drunk is dangerous.
 - If they didn't have this belief, they *should* have had it.
 - Cushman doesn't test for the effect of facts about what we *should* believe on judgments of *wrongness, blameworthiness, or appropriateness of punishment*. But facts about what we *should* believe may be just as important to such judgments as facts about what we *do* believe (as the drunk driving case helps bring out).
- How are we to understand the *desire*-factor? What counts as having the relevant desire? For example, does the desire have to be one we'd be willing to *act* on?
 - We might all (sometimes) desire other people's harm, even when these are desires we ourselves disapprove of, try to suppress, ignore in our decision-making, and would never act on.
 - Contrast that with a desire that shapes our decisions (even in the absence of a belief that we'll cause harm). These two different senses of desire might play very different roles in shaping our judgments about wrongness, blame, and punishment.
- How might we expect our judgments about such matters to react to deviant causal chains?
 - E.g., mountain-climber... Also, consider a similar story about Jenny...
- Cushman asked the subjects to consider how much blame the agents in Exp 1 deserved, and how much the agents in Exp 2 should be punished. But there seems to me to be an ambiguity built in to questions (like these) concerning the appropriateness of punishment/blame that is absent from questions concerning wrongness or impermissibility:
 - Calling an action/agent blameworthy /deserving of punishment may amount to making a moral judgment about that action/agent, but it also might be

taken as a moral judgment about a different action: the action of *blaming or punishing* the first action.

- This action may be justifiable or appropriate even when the first action is not, in the first sense, blameworthy or deserving of punishment.
 - Consider, e.g., utilitarian arguments about the appropriateness of blame or punishment, which can come apart from considerations of desert.
- Might this (partly) explain why subjects' judgments concerning blame/punishment come apart from their judgments concerning wrongness in cases where bad consequences followed?
- Relatedly, subjects may be treating questions of blameworthiness/appropriate punishment differently from wrongness and with more sensitivity to consequences because they are considering not just the perpetrator, but also what it seems reasonable for the *victim* to demand:
 - So, e.g., they might judge that, whether or not the agent acting *wrongly* (or more wrongly) in cases where harmful consequences resulted, in such cases there is a *victim* with a right to compensation or retribution. (Remember *Vincent v. Lake Erie*)
 - In harmless cases, by contrast, there is no victim with a right to compensation/retribution.
 - Might this (partly) explain why subjects' judgments concerning blame/punishment come apart from their judgments concerning wrongness in cases where bad consequences followed?

Experiment 3:

- Tested whether subjects' judgments of wrongness and appropriateness of blame or punishment varied according whether the actions caused harm, the harm occurred but was caused by something other than the action, or no harm occurred.
- *Results:*
 - Judgments of *wrongness* were not significantly affected by whether the agents' actions had a harmful consequence, the harm occurred but had another cause, or no harm occurred.
 - Judgments of *blameworthiness* and *appropriate punishment* were sensitive to whether the action caused the harm, the harm had another cause, or no harm occurred. In particular, subjects appeared far more willing to exonerate the agent from blame when another agent caused a harmful consequence, as opposed to when no harmful consequence occurred.
 - Judgments of *blameworthiness* and *appropriate punishment* were much more sensitive to the agent's beliefs when the agent *caused harm* than when the agent did *not* cause harm.

Experiment 4:

- (Smith, Brown, and the allergies) Further examined subjects' tendency to assign less punishment to attempted crimes when harm befalls the intended victim but has an external cause than to attempted crimes when no harm befalls the victim at all.

- *Results:*
 - Subjects were almost twice as likely to assign *no punishment* to agents in the cases where harm had an external source than to cases where no harm occurred.
 - Subjects who *did* assign punishment assigned (on average) nearly the *same* punishment in cases where harm had an external source as in cases where no harm occurred.

Cushman's Explanation of Experiments 3 & 4:

- (i) The "blame-blocking" hypothesis: the evaluation of causal blame for a harmful event can block the consideration of an attempted harm-doer's malicious intent. The occurrence of harmful consequences in experiments 3 and 4 can trigger an assessment of causal responsibility (whether or not they are caused by the agent's actions).
- (ii) In the absence of any harmful consequences to trigger the assessment of causal responsibility, an assessment of the agent's culpable mental state dominates punishment judgments.

Some Questions about these experiments and Cushman's conclusions:

- Might the results of Exp 4 be muddied by the possibility that some subjects blamed *Smith* for the harm that came to him, because he ate the salad with hazelnuts despite having a nut allergy?
 - If so, this would not (from their perspective) be a case where no one is responsible for the harm.
 - Thus it would not rule out the possibility (which Cushman was attempting to exclude) that the subjects were assigning attempted harm-doers less blame when harm occurred but had a different cause than when no harm occurred because they see blame/punishment as a kind of "limited resource."
- Perhaps some subjects took themselves to be answering a *different question* in the case where Smith is harmed by an external cause than in the case where no harm befalls Smith.
 - Perhaps in the harmful consequence cases, subjects (sometimes) took themselves to be addressing *not* how much punishment Smith deserves for his attempt to harm Brown, but rather how much punishment Smith deserves for the fact that Brown ended up suffering from an allergic reaction to nuts.
 - If they are indeed answering this different question in the harmful consequence cases, it's not surprising they come up with a different answer (no punishment).
- Generally, Cushman's blame-blocking hypothesis suggests we're comparing apples to apples: that subjects judging the appropriate level of punishment in external-harm cases and no-harm cases are making different judgments *about the same thing* – how much punishment the perpetrator (e.g., Brown) deserves *for what he did*. But that seems questionable:

- Firstly, subjects might take their task to be judging what an appropriate reaction on the part of the victim would be when a harm occurs, but not when it does not occur (e.g, what kind of compensation does the victim have a right to claim, and from whom?).
- Secondly, subjects might take their task to be judging culpability for *the actual harm* in cases where the harm occurs (but has an external cause), but not in cases where no harm occurs (in those cases, they must be take their task to be assessing culpability for the attempt).

III. On the Relevance of Empirical Moral Psychology to Ethics

What can we learn from Cushman's experiments and others like them?

- Can we learn anything from such experiments about how we *should* assess blameworthiness and assign punishment, as opposed to how we *do* do these things?
- Can we, that is, learn anything about which such judgments are *true*? This may depend on our meta-ethical views:
 - Is morality more like mathematics or linguistics?
 - We'd be disinclined to pursue mathematical knowledge by means of such experiments (they'd only tell us, maybe, what kinds of confusions and mistakes people are prone to when doing math!).
 - But we do pursue linguistic knowledge by such means (because we think that, to a certain extent, use facts shape facts about meaning and grammar)
 - Relatedly, are moral truths *response-dependent*?
- Even if morality is response-dependent, there are further reasons to be skeptical about the possibility of learning moral truths from such experiments:
 - It's much harder in the moral case than in the linguistic case to figure out what constitutes "normal" or "ideal" conditions and "competent" users.
 - Perhaps unlike the linguistics case, identifying the conditions under which asking people their opinions might teach us about moral truths might *itself be a moral question* – one we can't answer empirically.
 - There's a wide-open question about which "responses" should interest us when we're trying to uncover moral truths by means of such experiments:
 - Are we interested in what *judgments* people make (e.g. Cushman)?
 - But note that as Cushman's discussion makes clear, people often make *conflicting* judgments – e.g., that the appropriateness of blame/punishment tracks wrongfulness, and that consequences matter to blame/punishment *but not* to wrongfulness. Which judgments get at the truth?
 - Are we interested in what *actions* people would take (e.g., how they actually would punish people)?
 - If we're interested in people's actions, are we interested in them in *real* or *ideal* conditions?

MIT OpenCourseWare
<http://ocw.mit.edu>

24.235J / 17.021J Philosophy of Law
Spring 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.