# On Maurice Mandelbaum's "Determinism and Moral Responsibility"[1]
Forthcoming in *Ethics*.

## Manuel R. Vargas
Philosophy and Law, University of San Francisco

It is sometimes instructive to reflect on a problem as it appeared before our current philosophical presumptions became ingrained. In this context, Maurice Mandelbaum's "Determinism and Moral Responsibility" is of particular interest. Published in 1960, it appeared only a few years before the wave of work that gave us much of our contemporary understanding of moral responsibility, free will, and determinism.[2] Mandelbaum's account repays reconsideration.

Mandelbaum argues that (1) there is an underappreciated threat to "determinist" or compatibilist accounts of responsibility,[3] and (2) this threat can be met with a suitable compatibilist account of the etiology of action.[4]

Regarding the first claim—the threat to responsibility—the argument Mandelbaum

---

[1] A retrospective essay on Mandelbaum, Maurice. "Determinism and Moral Responsibility." *Ethics* 70, no. 3 (1960): 204–19. All references to page numbers are to this article, unless otherwise noted.

[2] Apart from the occasional nod to Kant and Hume, almost no contemporary work on moral responsibility, free will, and determinism is visibly indebted to work prior to the early 1960s.

[3] Mandelbaum understands moral responsibility in terms of an agent being answerable for her actions in a way connected to justifiable praise or blame (207). He never defines determinism, but he understands it primarily in terms of the causal explanation of action (205-6). Compatibilist theories of moral responsibility hold that moral responsibility is compatible with the thesis of determinism. Mandelbaum refers to compatibilists as "determinists" because in that era, compatibilists typically thought determinism was true. Contemporary compatibilists tend to be agnostic about determinism, and usually maintain that free and responsible action is compatible with both determinism and indeterminism.

[4] It is worth noting that Mandelbaum's alleged threat to compatibilist theories seems to have remained underappreciated. Mandelbaum's article has not been cited in the literatures on moral responsibility and free will in the past two decades. Nor was it ever anthologized.

presents proceeds from the assumption that any judgment of blameworthiness (or judgment that someone deserves punishment) is itself subject to determinism. Such judgments will thus be explicable in terms of, for example, that person's heredity, childhood experience, and even the events at her breakfast table that morning. If that is so, any instance of blaming and punishing will lack adequate justification. Why? If determinism is true, then the explanation of the person's judgment of blameworthiness should be regarded as a product of the distant past, as opposed to a product of the morally relevant features of the blameworthy act. Indeed, "any attempt to justify a moral judgment would simply be another case in which we were causally determined to talk one way rather than another" (210). In short, Mandelbaum worries that "our 'justification' of these acts is reduced to the acknowledgment that we could not help but praise or blame, reward or punish" (211).

From our current vantage point, it can seem somewhat opaque why Mandelbaum believes that a complete explanation in terms of antecedent causal forces precludes there simultaneously being a fully adequate account of the justification for a judge or blamer's judgments in terms of his or her reasons. We may be tempted to object that as long as blameworthiness judgments are properly responsive to features of the world, those judgments are not *merely* a case in which agents judged as they did because of forces in the distant past. Compatibilism's basic strategy has almost never been to deny that actions may have distant causes. Rather, the strategy has been to insist that the main question is always whether the relevant features of agency mediate the effects of the distant cause.

To appreciate Mandelbaum's contribution, it helps to be clear about the particular deterministic threat he envisions. He is especially concerned to overcome what he labels

"traditional psychological determinism" (212).[5] What he rejects in this (purportedly) traditional deterministic view of agency is the idea that (1) past experience has a direct role in determining present behavior and (2) determination is always by antecedent causes. Instead, he maintains (plausibly enough) that the effects of the past are mediated by one's physiological makeup, so that past experience "can only affect present behavior by virtue of the neurophysiological traces which it has left" (213). He then goes on to propose that ordinary decision-making involves the interaction of what amount to structuring and triggering causes.[6] This latter proposal is especially significant.

On Mandelbaum's account, actions are typically products of the interaction of structuring and occurrent triggering causes. That is, an agent's prior mental states (the structuring causes) interact with the qualities of the choice set (the triggering causes) to produce the agent's decision. Although such actions remain determined, their determination is located in the occurrent interaction of an agent's psychology with the context. As Mandelbaum puts it, the "conjoint effect *is* my decision" (212).

For Mandelbaum, the most important feature of this view is that judgments of blame are not products of the distant past. Rather, such judgments are determined by a composite of the present features of situation (including the blamed agent) and the psychology of the blamer. On this picture, it is the wrongful nature of the blameworthy agent's act that explains the judgment of blameworthiness, not the morally irrelevant features of the blamer's past (216).

One curious irony is that even though Mandelbaum is partly motivated by dissatisfaction with libertarianism, the basic picture to which he appeals is compatible with it. That is, either or

---

[5] Edwards, Hospers, and Ross are the apparent proponents of this view.

[6] This idea later became influential through the work of Fred Dretske. See, for example, his *Explaining Behavior* Cambridge, MA: MIT Press, 1988.

both triggering and structural causes could be indeterministic, and it would appear that the ensuing effects would be just as much the agent's decision as it is in the deterministic case. Perhaps this possibility is masked by Mandelbaum's tendency to slide between the idea of something being settled (made *determinate*), especially via deliberation, and its being a product of causal necessitation (i.e., causal *determinism*). Mandelbaum's interest is primarily in the former, but he assumes that it entails the latter.

I doubt many contemporary philosophers will be moved by Mandelbaum's worries about psychological determinism. Nevertheless, his solution is independently interesting, and its basic framework is one that many philosophers would now readily accept. It is but one of several intriguing ideas lurking in this forgotten essay.