Recent philosophical discussion of free will has come to focus, in large part, on identification of the freedom condition for moral responsibility and this shift has raised further questions about the concept of moral responsibility itself. Some have stipulated that a successful analysis of the kind of freedom required for moral responsibility must account for our basic desert of moral praiseworthiness and blameworthiness (see Derk Pereboom, *Living Without Free Will*, Cambridge: University Press, 2001). However, little work has been done to explicate precisely what desert in this basic sense amounts to, and there is much disagreement about whether competing views can successfully deliver it. In *Rethinking Responsibility* K.E. Boxer tackles some of these issues head on, raising a challenge to incompatibilists to provide convincing arguments in support of their assumption that basic desert of moral praise and blame requires *ultimate responsibility* – that we be the ultimate originators or sources of our actions “all the way back.” Boxer argues that incompatibilists’ prospects for meeting this challenge look grim, and so one of the most prominent avenues to incompatibilism fails.

In what follows I will focus on Boxer’s treatment of two possible routes to incompatibilist arguments for the claim that moral responsibility requires ultimate responsibility. In the remainder of the book, Chapters 4 and 5, Boxer explores whether punishment on either a ‘communicative’ or ‘non-communicative’ understanding might plausibly require ultimate responsibility. She argues that, on the communicative view, desert of punishment does not plausibly require ultimate responsibility, but grants that on the non-communicative view it may. However, appeal to the claim that non-communicative desert of punishment requires ultimate responsibility is not a move that Boxer expects many incompatibilists will be willing to make, given that there is little to distinguish the non-communicative view from problematic retributivist views of punishment more generally. For those interested in issues having to do with ultimate responsibility and desert of punishment, Boxer’s discussion here is both provocative and illuminating.

In the heart of the book, Chapters 2 and 3 Boxer assesses the plausibility of two possible incompatibilist arguments for the claim that basic desert of moral blame requires ultimate responsibility. These arguments appeal to two distinct accounts of the nature of moral blame, and Boxer argues that agents need not be ultimately responsible in order to deserve moral blame on either.

On the first account, moral blame is to be understood as a cognitive judgment or assessment, the content of which is that the agent’s action reflects a
defect in her moral character. Such defects are to be understood as categorically different from other defects in that they are such that agents have an unconditional obligation to eliminate them *qua* their status as moral agents. Based on this account of moral blame, incompatibilists might argue that there can be no such unconditional obligations unless we are ultimately responsible for our moral character. As Boxer puts the point, incompatibilists might claim that “unconditional oughts imply unconditional cans” and basic desert of moral blame requires that agents are capable of “having the required character (i.e., of being morally good persons), no matter what their character up to that point” (p. 47). However, Boxer rejects the claim that the unconditional obligation to be a morally good person at issue requires ultimate responsibility for one’s character. She takes the most plausible incompatibilist line of argument to establish this claim to be one that appeals to a Kantian picture of agency. On this picture, an agent’s character is constituted by her subjective practical principles. The unconditional obligation to be a morally good person requires that she be capable of having as principles only those that meet morality’s requirements (those that do not fail to give proper weight to the moral standing of other rational beings or the agent herself). And, the proposed incompatibilist argument continues, this capability requires ultimate responsibility for one’s character because without it an agent’s circumstances and existing principles (for which they are not themselves responsible) might be systematically faulty to the extent that the agent either cannot identify the principles that meet morality’s requirements or their existing principles fail to provide a motive to act on them. But, Boxer argues, this argument depends on a failure to recognize the kind of agents at issue – those properly subject to moral obligations in the first place (or, in other words, normal adult human beings). The Kantian picture of agency appealed to in order to motivate this possible incompatibilist argument guarantees that such agents have the capabilities required to ground an unconditional obligation to be a morally good person simply in virtue of their nature as persons, “those for whom the recognition of moral requirements is a sufficient determining ground of the will” (50).

Or so says Boxer. One obvious worry for her dismissal of this potential incompatibilist argument is that it targets only those incompatibilists who presuppose a Kantian view of agency. It makes no mention of the prospects for incompatibilists who might appeal to alternative views. While Boxer offers persuasive arguments for thinking that this particular Kantian incompatibilist route fails, the arguments in Chapter 2 fall far short of establishing their intended conclusion: that, insofar as moral blame reflects a defect in the agent’s moral character, incompatibilist arguments for the claim that desert of moral blame requires ultimate responsibility fail.