

**Kane is Not Able:
A Reply to Vicens' "Self-Forming Actions
and Conflicts of Intention"**

Gregg D. Caruso

SUNY Corning

Robert Kane's event-causal libertarianism proposes a naturalized account of libertarian free will, one that attempts to fit ourselves as free agents into the natural world characterized by science. As Kane describes it, his theory is an attempt "to see how far one can go in making sense of libertarian freedom without appealing either to *sui generis* kinds of agency or causation" (2002, p. 416). If Kane were successful in providing an intelligible account of libertarian freedom, one that avoided such extra (or special) factors, such an account would need to be taken seriously by naturalists like myself. But as Leigh Vicens (2015) describes in her excellent paper, Kane's account faces a number of devastating criticisms. The two that Vicens focuses on has to do with (1) Kane's notion of dual efforts and (2) the intelligibility of talking about *efforts* to form particular *intentions*.

The first criticism maintains that Kane's notion of dual efforts—which is an essential component of his overall account—makes *Self-forming actions* (SFA) irrational since they involve an agent making efforts to do two incompatible things simultaneously. This criticism is a good one in my book. In fact, I developed a version of it myself in *Free Will and Consciousness* (2012). There I presented the following example:

[H]ow are we to understand an agent that is *actively* and *simultaneously* trying to bring about two inconsistent ends? When an agent is confronted with a difficult moral choice—like whether they should accept the sexual advances of a stranger or stay faithful to their spouse—it's easy to imagine the agent experiencing a conflict of desires. But to say that in such a situation the agent is actively willing *both* the moral choice to stay faithful to their wife *and* the choice to give in to their temptation makes the agent appear irrational. Such simultaneous but inconsistent efforts of will amount to the agent willing P and Not-P at the same time. Not only is this of dubious coherence, it is far from the model of rationally guided behavior we were promised [by Kane]. (Furthermore, it's worth reiterating just how counter-intuitive this is as an account

Gregg D. Caruso

of “overcoming temptation.” Try explaining to your wife that you *actively willed* (hence *tried*) sleeping with a stranger but was unsuccessful because the competing neural network “won” out—the result of an indeterminate event!) (2012, p. 51)

Given these difficulties, I agree with Vicens that this is a serious objection to Kane’s account.

Vicens’ second objection is that “it does not make sense to speak of *efforts* to form particular *intentions* at all” (2015, p. 93). Here too I agree. On Kane’s account, the dual efforts involved in SFAs are efforts to *make a choice*, which Kane defines as “the formation of an intention... to do something” (2007, p. 33). I agree with Vicens when she questions, “what exactly does it mean to make an effort to *intend* to do something?” (2015, p. 96). This component of Kane’s theory leads to conceptual confusion. As Vicens correctly notes, when we speak of an agent “making an effort” what we typically have in mind is the agent making an effort to do something *intentional*. “But if this is what ‘making an effort’ means,” Vicens argues, “then to say that someone is making an effort to form a particular intention would imply that she is *intending to form a particular intention*, which would seem nonsensical” (2015, p. 96).

Given that I agree with both of Vicens’ criticisms of Kane, I will focus the remainder of my comments on evaluating Vicens’ alternative account. Vicens maintains that both of these objections can be avoided if we instead place the indeterminism earlier in the sequence—that is, instead of viewing conflicting desires as leading to competing *efforts of will*, which is then only settled by an indeterminist event (i.e., one neural network temporarily winning out over another), Vicens suggests that we instead conceive of the conflict between our desires as being resolved *before* they lead to competing volitional efforts. In her own words:

The indeterminism in Kane’s model could still be preserved by supposing that it is undetermined which of the conflicting *desires* will “win”—or, to put the point in more agent-friendly terms, which of two incommensurable reasons will be selected by an agent as her motivation for acting, and so which of two possible but incompatible intentions (to act in a particular way *for a particular reason*) the agent will form. On such an alternative model, when one reason is selected, the agent thereby forms an intention to act in a particular way rather than another. (2015, p. 95)

A Reply to Vicens' "Self-Forming Actions and Conflicts of Intention"

Vicens maintains that this alternative model of SFAs preserves the core of what Kane is after: indeterminism, voluntariness, and rationality. While I'm sympathetic to Vicens attempt to avoid Kane's dual efforts, I'm not entirely convinced her alternative account is capable of making intelligible event-causal libertarianism.

Here are some of my worries. First, it's not entirely clear to me that Vicens' account preserves one of the *virtues* of Kane's account. On Kane's account, whatever the outcome turns out to be in the case of a SFA, the agent can be said to have wanted and willed that action. Can we say the same on Vicens' account? It's unclear to me that we can. If an agent forms an intention to act in a particular way rather than another—that is, they settle which of the conflicting desires they intend to act on *before* they make an effort of will—how can we then say that they *willed both outcomes*? Vicens has a couple of moves available to here. She could argue that once the agent forms an intention to act, they *could not have willed otherwise* since their intention determines the subsequent effort of will. The problem with this response, however, is that is doesn't allow for the ability to do otherwise after the point at which the intention has been settled on. It also doesn't preserve the idea that the agent willed *both* outcomes, hence deserves to be held morally responsible (in the basic desert sense) for either outcome.

Of course Vicens could argue that willing both options is exactly the problem with Kane's account, and that all that is needed to preserve *reasons-backing-ness* is that the agent desired both outcomes, had reasons for acting on either, and that indeterminism existed at the point of intention formation. But I'm not convinced that this is enough to preserve basic desert moral responsibility. I have plenty of desires that I do not act on, which do not reach the level of intention formation, and for which I think it would be wrong to judge me praiseworthy or blameworthy. To say, then, that the agent is morally responsible, *whatever the outcome*, would require that both desires reach the level of intention formation since it is at this level that we hold agents morally responsible.

Vicens could, I imagine, go counterfactual here. Suppose, to use Kane's example, a businesswoman on her way to an important meeting witnesses an assault taking place, and is faced with the choice of stopping to call for help, or continuing on her way. Vicens could argue that the businesswoman wants to do both, but cannot; and what she will end up forming an intention to do is undetermined. Suppose she ends up forming an intention to call for help. On Vicens' view, she won't have made an effort of will to form this intention but she presumably will subsequently will to *execute* her intention, or carry out the course of action that she has

Gregg D. Caruso

decided to do. Vicens could therefore say that if she ends up stopping to call for help, this is something she will have made an effort of will to do. On the other hand, if she ends up forming an intention to continue on to her meeting, then she will presumably will to execute *this* intention. So if she ends up continuing on to her meeting, then *this* will be something she will have made an effort of will to do. Thus on this possible reply, the agent is not determined to do what she does, but whatever she ends up doing it can be said to be something she willed.

While a more promising response than the first option, I think it effectively turns Vicens' account into a covert compatibilism (of the conditional analyses type) with an idle indeterminist premise dangling subjoined! What seems to be doing the real work here is not the indeterminism, but the counterfactual analysis of willing. What this reply is really saying is that, *if* the businesswoman had decided to continue on to her meeting (*contrary to fact*), then she would have been blameworthy since she would have made an effort of will to do so. If this is what Vicens has in mind, she needs to explain how this differs from a now widely rejected version of old-school compatibilism.

Another option open to Vicens is to argue that *even after* the agent has formed an intention to act in a particular way, it is still indeterminate which effort of will and subsequent choice will be made. But how would this reply respond to a case where an agent forms an intention to do X (say, to do the moral thing) but ends up doing Y (acting from ambition). In such a case, it would be impossible to say that the agent *willed* and *wanted* to do Y since the outcome runs contrary to his/her settled intention.

My second concern has to do with naturalism. Is Vicens smuggling in "sui generis kinds of agency and causation," something Kane works hard to avoid (and for good reason)? When presenting her alternative account, Vicens adopts the language of agent-causation. She says that it's the *agent* who selects her motivation for action—i.e., which of two possible but incompatible intentions she will form. I find this language confusing since the main appeal of Kane's account (at least for many) is that it avoids the scientifically dubious commitments of agent-causation. If Vicens is appealing to *sui generis* kinds of agency and causation here, then there are several additional objections she would need to address—which she does not take up in her paper. I would therefore like Vicens to clarify her metaphysical commitments further. Is her proposal a version of event-causal libertarianism? Or is she introducing additional metaphysical commitments?

Lastly, assuming for the moment that Vicens proposal is a version of event-causal libertarianism, then I think it faces the same *luck objection*, or

A Reply to Vicens' "Self-Forming Actions and Conflicts of Intention"

disappearing agent objection, typically raised against other event-causal theories. As Derk Pereboom describes the disappearing agent objection:

Consider a decision that occurs in a context in which the agent's moral motivations favor that decision, and her prudential motivations favor her refraining from making it, and the strengths of these motivations are in equipoise. On an event-causal libertarian picture, the relevant causal conditions antecedent to the decision, i.e., the occurrence of certain agent-involving events, do not settle whether the decision will occur, but only render the occurrence of the decision about 50% probable. In fact, because no occurrence of antecedent events settles whether the decision will occur, and only antecedent events are causally relevant, *nothing* settles whether the decision will occur. Thus it can't be that the agent or anything about the agent settles whether the decision will occur, and she therefore will lack the control required for basic desert moral responsibility for it. (2014, p. 32)

The concern is that, because event-causal libertarian agents will not have the power to *settle* whether the decision will occur, they cannot have the role in action that basic desert moral responsibility demands. I don't see how Vicens account fairs any better here. Without bringing agent-causal powers back into the picture, what does it mean to say that the agent "selects" one set of reasons (as her motivation for acting) over another? Presumably this "selection" is not within the active control of the agent since it is the result of an indeterminist "winning out" that the agent has *no ultimate control* over. [Put in terms of my earlier example (where one is torn between accepting the sexual advances of a stranger or staying faithful to their spouse), I don't see how we can accurately describe one desire temporarily winning out because of an indeterminist event (say, the desire to remain faithful) as an act of "overcoming temptation"—an act for which the agent is morally praiseworthy!]

I conclude, then, that while Kane is not able to preserve free will, neither can Vicens revised event-causal account.

Gregg D. Caruso

Works Cited

- Caruso, Gregg D. (2012) *Consciousness and Free Will: A Determinist Account of the Illusion of Free Will*. Lanham, MD: Lexington Books.
- Kane, Robert. (2002) Some Neglected Pathways in the Free Will Labyrinth. In R. Kane (ed.), *The Oxford Handbook of Free Will* (pp. 3-41). New York: Oxford University Press.
- . (2007) Libertarianism. In John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas, *Four Views on Free Will* (pp. 5-43). Malden, MA: Blackwell Publishing.
- Pereboom, Derk. (2014) *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.
- Vicens, Leigh. (2015) Self-Forming Actions and Conflicts of Intention. *Southwest Philosophical Review* 31(1): 93-100.