

Book Reviews:

The Significance of Free Will by Robert Kane. New York: Oxford University Press, 1996.

In skimming through the notes and references to Robert Kane's critical overview of the last few decades of literature on the free will debate, one cannot help but be awed by the sheer volume of work that has been done on the subject — as well as its persistence as one of the great problems of philosophy. (Millennialist readers no doubt nod approvingly as they read Kane's citation of 12th century Persian poet Jalalu'ddin Rumi, who opined that the free will "disputation" would continue to the Judgment Day.) More importantly, in following Kane as he develops his own challenging incompatibilist position during the course of this survey, one cannot help but admire the various methodological strengths of his approach, including especially the ability to see and elucidate a simple, yet powerful analytical structure on which to hang all of the intricate twists and turns of the fabric of the debate — although, as I hope to show shortly, it seems that he has not laid out that structure in quite the way that he has labeled it.

Kane is very sensitive to the need to come up with not only a coherent model of free will, but also one that corresponds to reality — in other words, a model that not only meshes in a non-contradictory way with other basic ideas and values (viz., the nature of causality and one's needs as a human being), but also refers to something real and intelligible. (Readers of this journal may be struck, as this reviewer was, by the similarity of Kane's approach to the "integration-reduction" model espoused by Leonard Peikoff in *Objectivism: the Philosophy of Ayn Rand*, New York: Dutton, 1991, the task of gaining valid knowledge being taken to be twofold: derive with logic from one's perceptually given data a system of abstract ideas that cohere with one another, and make sure those ideas adhere to reality by being able to logically trace their connection back to a foundation in perceptual data.) The task, Kane says, is like climbing a mountain (the "ascent" problem) and getting back down again (the "descent" problem):

[A]bstract arguments for incompatibilism that seem to get us to the top of the mountain are not good enough if we can't get down the other side by making intelligible the incompatibilist freedom these arguments require. The air is cold and thin up there on Incompatibilist Mountain, and if one stays up there for any length of time without getting down the other side, one's mind becomes clouded in mist and is visited by visions of noumenal selves, nonoccurrent causes, transempirical egos, and other fantasies. (14)

One also has to appreciate Kane's unwillingness to rest with the traditional impasses and the attendant dismissal of further discussion, his pioneering urge to "dig more deeply into the conflicting intuitions that lie beneath the disagreements," to delve into "whole passages in the labyrinth of free will as yet unexplored" (5, 16). Kane gratefully accepts the compatibilist acknowledgment of various kinds of free will "worth wanting," which amount to *acting* freely, i.e., being "unhindered in the pursuit of your purposes." He pushes on, however, to consider the kind of freedom that is not compatible with even soft determinism, namely, "the power of agents to be the ultimate creators (or originators) and sustainers of their own ends or purposes" (4). Is this kind of free will compatible with determinism? Why do we, or should we, want it? Does such a freedom make sense? Does it actually exist? In dealing with these four questions — the questions of Compatibility and Significance (grouped together by Kane as the "ascent problem") and the questions of Intelligibility and Existence (treated under the "descent problem"), respectively — he seeks to refute the compatibilist claims that free will as "ultimate creation of purposes" should be dismissed as obscure and unintelligible and as not fitting "the modern image of human beings in the natural and social sciences" (5).

As said above, Kane seems to have erred in grouping and/or labeling these problems. It appears to this reviewer that Compatibility and Significance, while naturally (as Kane notes) being considered together, comprise an ascent and descent movement, as do Intelligibility and Existence. In the first movement, we consider the *plausibility* and *real value* of incompatibilist free will, the possibility of a kind of non-determinist free will that is "worth wanting." In the second movement, we consider the *intelligibility* and *real existence* of incompatibilist free will, the possibility of a kind of non-determinist free will that has a place "in the natural order where we exist and exercise our freedom" (184). Plausibility and intelligibility both (not just the former) pertain to ascent, to focusing on the coherence and non-contradictory status of ideas, while real value and real existence pertain to descent, to focusing on the adherence of ideas to reality. In support of this claim, I note that Kane states *all* of his "theses" — sentences or paragraphs elaborating his theoretical view of incompatibilist free will — in the sections on Compatibility and Intelligibility, *ninety* of them in all, and none in the sections on Significance and Existence. Secondly, I note that the sections on Compatibility and Intelligibility contain a total of 122 pages of material, while the sections on Significance (for which the book is named!) and Existence together amount to a scant 34 pages.

So, not once but twice, it appears, Kane takes us up "Incompatibilist Mountain" in a noticeably lengthier and more laborious ascent, then back down in a breathtakingly swift, brief descent. There is nothing wrong with climbing the mountain twice, so long as one is clear about what one is doing; and it would probably not be too difficult to arrive at a suitable relabelling of the sections comprising those movements. It may well be, however, that Kane's Significance chapter — and thus the very need and justification of his book! — is fatally compromised by his having hung its conclusion on the (as I will argue) at least somewhat suspect notion of "objective worth," which in turn may be a result of his failure to realize that Significance is an issue of descent, viz., of correspondence to reality of one's values. Despite this methodological problem,

though, does Kane nonetheless succeed in meeting the compatibilist challenge(s)? Not entirely, as I hope to show in the remainder of this review, which will briefly consider and evaluate Kane's interesting and sometimes innovative answers to these problems and questions.

In addressing the Compatibility Question, Kane says that while freedom from constraint, coercion, and compulsion are all worthwhile as well as compatible with determinism, another important and more basic kind of freedom, the "power of agents to be ultimate creators and sustainers of their own ends and purposes," is not (32). Although two frequently asserted requirements for the latter freedom are often regarded as equivalent, Kane believes that clearly distinguishing them is a necessary condition for moving beyond the current impasse between compatibilists and incompatibilists. One of these requirements is that an agent "could have done otherwise," that the agent has "Alternate Possibilities" (AP) — the other that the agent of an action is the source or explanation of its action, that the agent has "Ultimate Responsibility" (UR). Unlike many other incompatibilists, Kane holds that the case for free will requires not only AP but also UR, and that AP is only seen to be necessary for free will by invoking UR. Indeed, Kane says, the basic reason for the stalemate on the question of whether someone "could have done otherwise" is that compatibilists do not take the UR condition seriously. Thus, the two sides unwittingly argue at cross purposes. Were they to explicitly acknowledge the necessity of UR, it would be clearly seen that free will, "freedom to do otherwise," cannot be accounted for in compatibilist terms.

Kane unpacks the two facets of UR as follows:

[I]n order to be ultimately responsible for being what you are (or for having the character and motives you do have) there must have been something you could have voluntarily done (or omitted) at some time or other that would have made a difference in what you are (or in the character and motives you now have) [I]f antecedent conditions and laws of nature (or prior character and motives) provide a sufficient reason or explanation for an action, then . . . the agent must be responsible for at least some of those explaining conditions. Something the agent willingly did or omitted must have made a difference in whether or not these explaining conditions were the case. (72, 73)

By "sufficient reason," Kane means either a logically sufficient condition, a sufficient cause (i.e., an antecedent condition governed by a law of nature), or a sufficient motive. By the latter, Kane means to include the requirement that ultimately responsible agents be the source of not only their actions but also their *will* to engage in those actions. In order to avoid an infinite regress of voluntary actions for which one is personally responsible, he argues, there must

be some such voluntary actions (“self-forming actions”) that were not determined (i.e., had no sufficient causes). What one must have been able to voluntarily do that would make a difference in whether or not voluntary actions occurred “is simply *doing otherwise*, rather than doing something else that would have causally contributed to their not occurring” (75).

Thus, Kane says, “could have done otherwise” has a legitimate meaning beyond the thinner sense attached to it by compatibilists. Moreover, one can also be ultimately responsible for actions when one could *not* have done otherwise, so long as those actions were willed actions (issuing from one’s character and motives) and one was responsible by earlier self-forming actions for the character and motives from which one’s actions followed. Acts “of one’s own free will” can be *determined* by one’s will and still be the *product* of one’s free will. In regard to the *forming* of that will, however, determinism cannot be the sole factor, if one is to be held responsible for its formation — not to mention the actions flowing from its exercise. The problem I see with this position is that it does not allow for the possibility that the attainment or formation of one’s free will is not a voluntary achievement but instead an *emergent capability* that arises as a matter of course in the (relatively) healthy development of a (relatively) normal human being. One at some point may simply be *provided* with this capability by the unfolding of one’s genetic endowment, much as one is provided — after sufficient development and absent excessive constraint, coercion, compulsion, or oppression — with one’s sexuality or one’s capacity to realize that objects continue to exist when unseen. What one *then* does with this capacity, is what is “up to one,” that for which one can be held accountable; but even how one’s will becomes “set one way,” a way in which one most wants to act or is *intending* or *trying* to act, may be determined before one has anything to say about it voluntarily (114). I hasten to add that I am not asserting that this is true and that indeterministic free will is false. Indeed, I don’t think we yet know what is the case. More information is needed: the jury is still out as to whether self-forming willings are voluntary acts or pre-programmed developmental events.

Kane, however, on the strength of his confidence that he has demonstrated the plausibility of free will cum ultimate responsibility, next considers the Significance Question: “What is so important about ultimate responsibility that should make a freedom requiring it ‘worth wanting’?” Many of us, after all, believe we want “sole authorship” or “underived origination” over our actions. Are we realistic in wanting this? Traditionally, it is held to be necessary for a number of other desired and worthwhile things such as creativity, autonomy or self-creation, desert, moral responsibility, suitability of being an object of reactive attitudes, dignity or self-worth, a sense of individuality or uniqueness, life-hopes, freely given love and friendship, and acting of one’s own free will. In “the dialectic of origination,” compatibilists try to “deconstruct” or “demythologize” these things, “to provide plausible explanations of them that do not require underived origination and incompatibilist free will,” while incompatibilists respond by trying to argue that compatibilist versions “fall short of what we *really* want” (80).

To move beyond this impasse, Kane leads us through “the dialectic of self-hood,” where he considers the importance of free will in relation to “our place

or importance in the scheme of things" (92). Far from "delv[ing] rather deeply into the metaphysical depths," however, Kane instead looks at a most non-esoteric topic in developmental psychology: how a human being arrives at and expands his sense of self. At each stage of the unfolding process of understanding the relation of our selves to the world, we come into contact/conflict with people or things or ideas — including, at one point, the doctrines of determinism — that threaten our experienced status as "independent sources of activity or action in the world," rather than "mere products of forces coming wholly from the world — forces that are not the products of our own wills" (95). Since we actually desire and believe in this independent selfhood, which Kane says is "a precondition for moral agency in the fullest sense," we tend to resist anything such as deterministic arguments against free will that could undermine that belief and value (97). Thus, free will is strongly implicated with the idea of independent selfhood, which seems to be a natural, inevitable part of healthy human development; if either is valued, both will be. It seems unclear, however, that anything beyond the compatibilist desires for freedom from constraint, coercion, compulsion, and oppression is required by the desire to be and view oneself as an independent self and actor and willer in the world — which is a relative independence, after all, as Kane himself concedes in his conclusion (213-4). The dialectic of selfhood seems more a continuation and deepening of the dialectic of origination than a new point.

Where Kane really begins to break new ground is in his consideration of "objective worth," which he suggests is intimately related to ultimate responsibility and free will. When we dig down to the roots of the desire for incompatibilist free will, he says, we find two basic, correlative desires:

- (i) the desire to be independent sources of activity in the world, which is connected . . . from the earliest stages of childhood to the sense we have of our uniqueness and importance as individuals; and (ii) the desire that some of our deeds and accomplishments . . . have objective worth— worth not just from one's own subjective point of view, but true (i.e., nondeceptive) worth from the point of view of the world. (98)

Kane is saying here that there is more to value than "a person's subjectively felt happiness" about the results of his acts; and if there is, then incompatibilist freedom with ultimate responsibility *could* have a value beyond that of the compatibilist freedoms: "Such freedoms would be enough, if we did not care about more than what pleases us — namely, if we did not care in addition . . . about our 'worthiness' or 'deservingness' to be pleased" (97-8). Thus, Kane seems to characterize the debate over a kind of free will "worth wanting," by depicting the compatibilists as being enmired in a worldly, hedonistic, pleasure-seeking paradigm, in which what they want is most important, in contrast to incompatibilists who want to rise above this shallowness to a more universal, spiritual, "objective worth" paradigm, in which what is *worthy* is most

important. If the proper ethical perspective is from *inside* our own life and awareness, a perspective in which all that matters is supposedly “subjective happiness,” then “objective worth” will seem pointless; but if we “stand back and take an objective view of the universe and our place in it,” Kane says, we will come to believe that our acts have worth “from the point of view of the universe,” which is what he appears to mean by an “objective point of view.”

Kane attempts to persuade the reader on this point by using an “alternative world” scenario, in which a person experiences the same rewards and satisfactions in two different worlds, but one of the experiences is based on reality while the other is a deception. Wouldn’t the former world be “better” to live in, even if no difference was *experienced* between them? Of course it would, but only by virtue of one’s (or someone’s) *projecting* himself into the position of knowing the *truth* about the two worlds; absent such knowledge, all one can go on is the evidence one has. We are not infallible; we may in fact be deceived in one instance or another. But deceit must exist in relation to the truth. If one is deceived and finds out, one can try to adjust. If one never finds out the truth and has a full life of success and happiness in a fool’s paradise, what is the difference — and to whom? Significance is not absolute and cosmic; it is *contextual* and pertains to people living *in this world*. Desiring that there be a cosmic tallyboard toting up one’s true virtues and achievements so that one’s legacy will be accurate, even if humanity’s awareness of it was not, is at best a lapse in self-confidence, a failure to trust one’s own best judgment about whether or not what one has done and made of one’s life and character is good. That some people feel that life is not fully meaningful without this kind of assurance that the universe is “looking out for them” is not a valid argument for incompatibilist free will.

In all fairness, it must be granted that Kane does not attempt in this volume a rigorous justification of his notion of “objective worth” (deferring instead to his *Through the Moral Maze*, Arnouk, NY: M.E. Sharpe, 1994). Even so, it is difficult to see how his proposal is more than just another strategy— and not an altogether necessary one — people might use to help themselves to “resist the idea that the activity we direct back *upon* the world . . . has merely illusory and not real significance or worth *for* the world” (100). In this reviewer’s opinion, however, that idea should *not* be resisted, but instead *embraced*. Kane’s wording reveals rather clearly that he has accepted a traditional false dichotomy between subjective-as-personal and objective-as-impersonal — to wit, that personal values must be subjective and objective values must be impersonal. While personal values often are subjective, they are not always so; more importantly, there is no such thing as “impersonal value,” only the things valued by some *person* (or other living being) for some purpose, life-serving or otherwise.

What this defective dualism thus leaves out is a third view, only recently coming into elucidation: the idea that values can be objective and personal, i.e., factually based and “agent-relative.” (One of the best presentations of this theory can be found in Douglas Den Uyl’s and Douglas Rasmussen’s *Liberty and Nature*, LaSalle, IL: Open Court, 1991. An equivalent perspective can be found in the aforementioned volume by Peikoff, although his relabelling the traditional objective view as “intrinsic” tends to muddy rather than clarify the debate.) Regarding value and significance as not necessarily subjective, but

often instead as agent-relative and factual, allows us to expand the sense of our own personal values beyond the feeling-based, without wandering onto the untenable ground of the point of view of an impersonal universe (into which are often smuggled, without acknowledgment, the *personal values* of someone else and/or the values that someone else who wants to control us wants us to adopt). It is for us and, by extension, those who matter to us (or to whom we matter) — not for “the world” — that our activities have “real significance or worth,” if they have such at all. It is not our “worthiness” to the world that determines whether we deserve happiness, but our *self-worth*. Self-esteem does go beyond the “subjectively felt happiness” (i.e., pride and satisfaction) about the results of our acts, but it is merely a more global, holistic sense of well-being of *ourselves, for ourselves*, not something for which we need the endorsement of an impersonal world. It is Kane’s apparent failure to realize this that leads him to regard the significance of free will as a “metaphysical” issue and part of “the problem of ascent;” and it is this same failure that undercuts his attempt to move beyond the impasse between compatibilists and incompatibilists. In the judgment of this reviewer, Kane has not made his case that there is a noncompatibilist form of free will that is worth wanting.

If there is a case to be made for a form of human freedom that goes beyond the compatibilist freedoms that are subsumable by the traditional event/state, cause-effect model of causality, yet does not lapse completely into the indeterminism that Kane favors for human action, it must pass the test he describes as the Intelligibility Question. This means that such a theory must amount to more than “a perfunctory treatment that consists of putting labels on mysteries” (105), though Kane himself bows to the apparent need to accept such mysteries even for his own view, which includes such notions as “indeterminate efforts” (151) Before fleshing out and defending his own model of “self-forming willings” (124-183), Kane examines a number of traditional and contemporary libertarian strategies to reconcile indeterminism with free will, the requirement being that an agent must be able to willingly do and do otherwise, “all past circumstances and laws of nature remaining the same.” The problem is to explain how a person’s free choice of either of two alternatives satisfies this condition, how one is “able to choose either option rationally, voluntarily, and under [one’s] voluntary control, given the same past and laws of nature.” Proposed solutions to this problem almost always involve some additional factor beyond past circumstances and laws of nature, a strategy that Kane regards as “dangerous,” in that it tempts libertarians to make recourse to “mysterious sources outside of the nature order or to postulate unusual forms of agency or causation whose manner of influencing events is at best obscure.”

In seeking to avoid such approaches involving special forms of agency or causation and to try instead to engage modern science in a “more meaningful dialogue” (115), Kane invokes the Free Agency Principle, which eschews the appeal to entities or causes that are not also needed by nonlibertarian theories: “In the attempt to formulate an incompatibilist or libertarian account of free agency . . . we shall not appeal to categories that are *not also needed by non-libertarian* (compatibilist or determinist) *accounts of free agency* . . .” (116). In the process, Kane weighs and finds wanting a plethora of ideas, including Kant’s noumenal self, Eccles’ “transempirical power center,” the “Will” as a

homunculus or “agent within the agent,” the mind-body dualism of the Cartesian ego, antecedent-cause reasons, and special non-antecedently-caused “acts of will” or “acts of attention.” (Readers of this journal who are familiar with how Leonard Peikoff, Nathaniel Branden, and other proponents of Ayn Rand’s Objectivist philosophy, place the locus of non-determinist free will in the act of focusing one’s awareness or attention, may want to ponder Kane’s assertion that armchair speculation or a priori assertion is not enough to support the claim that such acts are undetermined in principle.)

While he regards such accounts of free agency as risky, however, Kane does not completely reject them out of hand, simply as unnecessary for an explanation of incompatibilist free will; he has his explanation in hand, so these other strategies are at best superfluous. However, I think Kane has missed something vital in at least one of the strategies he mentions: the idea of agent or “nonoccurrent” causation. Kane concedes that his Free Agency Principle only rules out the version of nonoccurrent causation that is designed to account specifically for undetermined, incompatibilistic free actions; but he has little to say about the view held by “only a minority,” viz., the idea that “non-occurrent causation is required to account for action in general, whether determined or undetermined” (122). Edward Pols (see my review of his *Mind Regained* elsewhere in this volume) puts a special twist on this latter idea. His hierarchical model of causality, which puts entities in the causal driver’s seat, agrees that it makes no sense to say that agents engage in immanent or nonoccurrent (“atemporal”) causation in regard to their *actions*, a relation that is fully temporal or occurrent, though not determined by antecedent conditions. He would agree emphatically with Kane’s statement that “[T]he time has come in the history of free will debates when [the] pernicious assumption [that all control must be antecedent determining control] must be subjected to greater scrutiny” (186-7). Pols adds, however, this crucial insight: entities *do* exist in an immanent, atemporal, nonoccurrent causal relation to the parts of their own physical infrastructure that support their temporal actions. This whole-part, downward-upward, pyramidal aspect of causality is, Pols says, what is missing from the microreductionism of modern science’s account of action in general; and invoking the complex arrangement of temporal and atemporal relations in hierarchical causality to help explain human action — whether or not such action is at all indeterministic — would fall well within the strictures imposed by Kane’s Free Agency Principle.

In his further consideration of agent determinism in the brief section dealing with the Existence Question, Kane concedes that “To reject agent-causation is therefore not to deny that there are agents and that they cause things in ordinary senses of the term . . . rather to deny that we need a special relation of nonoccurrent causation to explain all of this” (189). Pols’s way of including nonoccurrent causation in human (and other) action avoids this problem, while showing how it is necessary for fully understanding the hierarchical nature of causality. Kane states that something along the lines of a “recurrent brain network” has to be going on inside an agent’s brain in order to allow an agent’s actions to “outflow” from itself; and similarly to the way I characterized my proposed model of how the brain supports introspection in my review of Pols’s book, Kane says that:

[Such a] “self-network” [need not have] a specific location rather than being a complex network distributed widely throughout the brain. Its unity would lie in the dynamical properties of neural circuits and connections that make possible synchronized and causally interacting oscillations or patterns of firings throughout the entire network, like those described by Crick, Koch, Llinas and others, for conscious awareness and wakefulness. (193)

After the dust settles from Kane’s careful trek through the various controversies about free will and determinism, we are left with a form of indeterminist or libertarian free will that is amazingly like the most tenable versions of soft determinism. Once appeals to “obscure or mysterious forms of agency or causation” are set aside, and a rigorous effort made to formulate an intelligible model of free will and to reconcile it with the scientific view of the world, it becomes clear that the exercise of free will is not “entirely above and beyond the influences of natural causes and conditioning,” and that free will and moral responsibility instead “are matters of degree, and our possession of them can be very much influenced by circumstances” (212-3). Kane thus identifies and rejects for the myth that it is “the idea that we might attain complete autonomy or perfect freedom” (214). Looking to the future, he says that we need to recognize the many ways in which one’s free will and responsibility can be limited by circumstances of birth and upbringing, without being caught in the pitfall of thinking that all of us are inevitably the victims of those circumstances. By rejecting the naive all-or-nothing views of free will, we can excuse or mitigate the guilt of those with diminished moral responsibility due, for instance, to severe childhood abuse, while (with few exceptions for extreme cases) not letting such people completely off the hook. As for the future of the continuing intellectual debate, it is safe to predict that the framework provided in Kane’s book will be adopted by many as the best guide for delving into the intricacies and nuances he has outlined.

Roger E. Bissell