

Discussion

Reply to William Dwyer

Compatibilism and Evolution

George Lyons

Regarding William Dwyer's review of *Initiative* by Tibor Machan, some limitations in the arguments relied upon ought to be noted (Dwyer 2001; Machan 2000). Only one type of compatibilism is represented in the review, a classical kind associated with Hobbes and ancient Greek attitudes. A different position is represented by scientifically oriented compatibilism and in the views of such theorists as Daniel Dennett (1978–1996). Without this broader context, the Objectivist ideas of Machan remain portrayed as falling within a simple dichotomy between determinism and freedom, lacking the full extent of a dialectical advance over this intuitive, but long criticized, distinction.

Compatibilism in General

Compatibilism as a term superseded soft determinism to emphasize how different concepts of freedom have been applied to the psychology of conscious choice, and how some theories of freedom defined its essence in ways compatible with predetermination of events. "Softness" as a relaxation of causal necessity is really a different idea. In this respect, compatibilism defines a role for freedom in explaining knowledge, ethics, and liberty—subjects of Machan's book—in some ways familiar, while finding the relevant features of conscious choice, the essence of all these phenomena, to be different from traditional freedom from predetermination. There are any number of things to be free from and to do in different senses. The issue is which, if any, of these is relevant to moral psychology and consistently definable. Yet Dwyer does not explicitly

refer to freedom of will in compatibilism at all, preferring to emphasize motivation or goal-direction as an alternative to concepts of freedom altogether.

Dwyer's de-emphasis of mental freedom appears to tacitly accept a view of compatibilism Machan and others express, though not in the book reviewed, which speaks as though compatibilism never existed, despite its prominence in modern philosophy (Kane 1998). This view holds genuine freedom cannot be conceived as anything other than escape from predetermination, so that if there is anything wrong with that concept of freedom, freedom just would not exist. Compatibilism is then treated as some zen-like contradiction, one of the extreme deviations from reason in modern philosophy, or some arbitrary linguistic construction. Objectively, Machan allows only a dichotomy between freedom and predetermination. Knowledge, ethics, and liberty might all have justifications nonetheless, but more from a determinist rather than compatibilist analysis.

This incompatibilist view is understandable from a common sense perspective on choices made at a moment in time: if choice is determined, you cannot stand back and try to improve what happens in that choice; you cannot stand outside of time—except through the magic of indeterminate potentials coexisting in one event. While this requirement is true, however questionable is physically accomplishing it, compatibilism points to something else about what is determined: that it changes over time. There are differences between motion or changing freely over time, and predictable, unvarying paths through time which cannot be consciously altered—which do not, in effect, include consciousness within them. When the alteration we make is realized over time, our contribution is part of the patterns of change in events without contradicting their predetermination—provided change is described with sufficient complexity to encompass meaningful mental attributes (elaborated below in connection with Dennett). True, we cannot alter what is determined in a single event, but choice should not be conceived in that way. Choices are revisited and re-evaluated over time and deliberations for every choice extend over time advancing knowledge. The issue then is whether our decisions evolve freely over time or are determined in arbitrary ways.

Such analysis is typical dialectics-transcending-dichotomies, a practice for which Rand is lauded, but in this case, Objectivism clearly prefers the dichotomy.

In contrast to Dwyer's emphasis, other advocates of the motivational approach have identified this view with a concept of freedom—it basically came from Hobbes—so Dwyer is correct in referring to it as compatibilist, even though he does not emphasize the full meaning of compatibility. The motivational kind of freedom is one to accomplish motivating ends free of impediments from the causes of mind beyond ourselves—a freedom to carry out our will, instead of a freedom to formulate it, and a freedom to choose means to ends, instead of choosing the ends. The motivational concept of freedom is related to a classical one of freedom as power, of having more choices in virtue of abilities and faculties supporting pursuit of alternatives unavailable to less endowed individuals (Adler 1958). The psychological analog to material power of wealth, etc., is the scope afforded by knowledge: before choices can be motivated by ends, you need to know what will accomplish those ends, and have the skill to solve problems encountered in accomplishing aims. One then becomes free to make right choices from having the requisite wisdom—as would Plato's class of philosopher-kings be more endowed with powers of choice than the masses. At least superficially, one can infer such problems in basing liberty on this concept of freedom.

Problems of Incompatibilism

The Objectivist concept of freedom is completely different from the classical mentioned above, and harks back to a much older idea embodied in religion as well as law, in which people choose between good and evil with a knowledge of right and wrong available to everyone. This is not a choice about what to believe, but about whether to act on what we believe. Rand secularizes this idea by combining it with the classical virtue of rationality, so that we choose between an inherently good way of functioning mentally or a wrong way. Her emphasis on functioning as opposed to just beliefs of

rational or irrational kinds, gives a connection to introspectively observable features of how we think instead of what we think. However, the real issue is what causes how we think, which in principle encompasses everything you do not directly observe.

Choices of good and evil in principle idealize a familiar, even prominent experience known introspectively, of choosing between emotional and intellectual motivations. While this common experience serves as an intuitive motivation for describing choices as indeterminate, subtleties in idealizing it are not so easily noticed. Both the religious and secular views have a quality of psychologizing, of attributing to choices a selection between right and wrong in principle, as ultimate objectives. When we assume a complete knowledge of good and evil, then choices we make have no evolution over time of the kind described above as central to compatibilism. There is a finality to choices in the traditional view that can then have freedom only through some indeterminism associated with selection made at a moment in time. The very fact subconscious processes, much less physical neural activities, are not knowable directly by introspection gives an impression we choose whether or not to override such processes. Yet this impression is driven completely by the sophistication with which we are able to theoretically describe what subconscious activity we cannot directly observe. Sophistication is not just a matter of how detailed a description of mind we have, for there are also involved very general features of how change takes place over time, such as evolution (see below).

In the Randian concept of choosing between function and dysfunction as such, whether to think or not, the same logic arises, in that once having chosen dysfunction, there is no freedom in the further evolution of such a state; you are dead in the water, so to speak. Interestingly, the ancient origins of this idealization of choice in Western religion comes from a time of traditional ethics, in which no thought is involved in knowing right from wrong; religious authorities have already provided all the answers. Choices then become acts of faith, and the magical quality of faith as a means of knowing, naturally spills over into a similar quality of choice as a process. So again, the evolution of thought over time is dissociated

from choice.

Rand is similar here in that the rationality, mental effort, or focus that is chosen, contains within itself an inherent knowledge of how to function, a methodology if not specific facts. Yet, in modern cognitive science, including computational logic theory, the distinction between method and fact is not so clear; even methods have to be learned. After choosing to focus, how do you know what to focus on? That problem, Objectivists might say, is not a moral issue; we learn what to focus on as circumstances dictate; learning is just a mechanical function not our responsibility at least in its content; our responsibility is fulfilled in just choosing to learn in the first place. There is a learning function we perform, but not one exercising any freedom within it. So the question is which view is really fatalistic, the compatibilism of inevitably learning, subject to risks along the way, or incompatibilism of indeterminately choosing whether to learn? In some respects, the incompatibilist view serves as an escape from the real responsibilities of learning by shifting the burden to an imaginary choice whether to engage in learning.

The problems above with Hobbes should not be taken as support for Machan's concept of freedom from predetermination. The paradoxical nature of such freedom has long been noted. While Objectivism can refer to an agent being causally responsible for choices as an entity causing its actions, evoking a sense of power over affairs, when what is caused that way is only an indeterminism in choices made, there is no real ability, faculty, or process free to function and thus be worth blaming or crediting for its results. Nor is there any way we can make any use of self-criticism. There's no point in blaming a roulette wheel. The idea may survive for lack of clearly defined and simple alternatives. Hobbes's and Dwyer's motivationalism supplied one such alternative, but there is a strong sense in which this alternative begs fundamental questions. What determines the goals people have, and is this a fatalistic situation or not? What is the point in holding ourselves responsible for goals we just happen to have? There is no clear distinction in this view between a very mechanical pursuit of goals by a machine, which we would not hold responsible for its program, and a genuinely responsi-

ble agent. We hold agents responsible because they change their goals, and we expect blame and credit to help bring about such changes—especially when we apply such evaluations to ourselves introspectively, however intransigent other people may be. Change, as noted above, is not limited to the traditional indeterminism of different results at the same point in time; real change takes place over time without any such magic. The difficult problem, however, is conceiving how consciousness consists of such real change instead of a spiritual power divorced from physical events.

Scientific Compatibilism

Besides the general goal-direction conceived by Hobbes and others, another place to look for relevant kinds of freedom is in the means by which we find ways to accomplish goals, in how we come to know which choices are right and which wrong in order to choose with goal-direction. The idea of goal-direction as such obscures a complex process of discovering what in fact accomplishes goals, how we solve the problem presented by a goal to be achieved, like the problem of winning in a game of chess. This process can exhibit the combinations of seemingly contradictory elements making choice paradoxical. We may know intellectually right from wrong, and thus superficially appear to choose between them, but the subconscious processes that appear to select between rationality and irrationality as such, can themselves be learning and discovery processes. We do not really know how to act in a whole sense until the subconscious is as much informed as intellect.

Intelligence involves expanding information applied to choices, a process of change, and thus freedom of movement. In this view, the basis for assigning blame or credit to an agent is whether it is engaged in accumulating information, learning, investigating, exploring, inquiring, innovating—and making mistakes along the way, producing a need for our ideas of blame and credit. Both Objectivism and Dwyer seem to view expanding knowledge in a rationalistic way, as though just exerting mental effort or being goal-directed somehow presents a solution to the problems of accomplishing ends.

In practice, it actually requires a risky and unpredictable creativity.

Yet predictability is what Dwyer claims, within the Hobbesian view, as the real basis for blame, credit, or responsibility, a completely different theory than proposed here. Dwyer says “moral judgment is a means of assessing someone’s character, so that we can predict his future conduct” (Dwyer 2001, 101). A different view is that we need to blame or credit what changes in someone’s character, because it is from that evaluation improvement in what changes is accomplished. When character is predictable, we need no moral judgment at all, only what Machan associated with determinism—social engineering of people’s behavior. Because a computer is predictable, we can program it as needed. An unpredictable, evolving computer requires judgment, within itself as well as without, to accomplish its evolution; it is not programmable in advance (to most people, a computer *means* what is predictable or programmable; technically, however, computers are just combinations of computing elements, which may have unpredictable and changing organizations, despite the simplicity of the individual components).

Dwyer’s concept of judgment refers to something we do need to know—the goals that adaptive goal-seeking things have—in order to respond to them effectively. This is blame and credit in the sense of identifying causes of action. Yet a cause of action can be innovation in goals just as much as goals themselves. There is a different response needed then to innovating minds and to programmable minds, though both are goal-directed. Dwyer comments upon the role of judgment in social relationships, which can involve dealing with people who lack free will. But a different concept of responsibility may be required when applying judgment introspectively, to ourselves. The whole purpose of identifying goals we may have, without realizing it consciously, is to change them. When we apply this introspective attitude to other people, we may find a need for different strategies than a strictly Hobbesian approach promotes. There is a kind of liberty involved in the Hobbesian approach, for example, in which we have predicted everyone’s needs and satisfied them through state planning. The reason we might say the state cannot predict everyone’s needs is that they keep changing. Just as

freedom is a relative term, so is liberty. The kind of liberty relevant to human needs, the essence of all the associated phenomena, must be identified.

From Dennett's theories, we can view the expansion of knowledge accomplished by intelligence technically as an evolutionary process, a general form of change, or dynamics, of the kind Darwin applied to biological evolution in natural history. Darwin's discovery is revolutionary in Dennett's view in that it explains not just natural history, but also how complex design and solutions to problems in general, can be created by physical means, which themselves lack order, design, and purpose. This is fundamentally a resolution of basic mind/body problems in philosophy, including paradoxes of will. Evolution adds something to the dichotomy of events being either predetermined or indeterminately random: a relationship and interaction between causal necessities and random chance events that changes the effects chance otherwise has, just as relations between the parts of any system give the whole properties beyond those of the individual parts in isolation. In disordered physical systems, randomness is a decay process, the entropic dissolution of things in thermodynamics. Throwing dice can hit a good result by chance, but revert to a bad result in the next iteration. Natural selection adds a latching effect that preserves the good results and builds upon them, converting chance from a source of decay into a means of exploring opportunities for advancing order and design—in effect, the creativity of intelligence.

What we do in making choices, then, is explore for new opportunities, not just compare known alternatives on a menu with objectives we pursue. Such exploration is risky and fallible in the short-run, but has a long-term probability of success that justifies attributing blame and credit to it, or as Dennett would say, is a variety of free will worth wanting. Here we have a process that is responsible for both wrong and right choices, a seeming paradox, but which remains a definable faculty with a definite end result because we separate the conflicting aspects in time: uncertainty in the short-run and certainty in the long-run.

At the same time, evolutionary processes are consistent, or

compatible, with physical determination of mind, without imposing any special demands on physics to explain mental qualities. Computers have been programmed to do it in a limited way already. In being so programmed, their programmers cannot be blamed for mistakes the computers make, because the designer does not predict what the machine will do, not even hypothetically in prescribed rules. Evolution completely fits the modern paradigm of functionalism in psychology, which superseded the crudely positivist Behaviorism of the early twentieth century. Neither unusual uncertainties of Quantum Mechanics nor any mysterious kind of "emergence," in which predetermination magically generates some indeterminate result, is required for evolutionary explanations. As a concept in mathematical complexity theory, evolution is a precursor, a century and a half old. In philosophy, it is still relatively new, coming as it does after the Enlightenment in the Industrial Age, and from science rather than philosophy.

References

- Adler, M. J. 1958. *The Idea of Freedom*. Garden City, New York: Doubleday.
- Dennett, D. C. 1978. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, Massachusetts: MIT Press.
- . 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. New York: Oxford University Press.
- . 1991. *Consciousness Explained*. Boston: Little Brown.
- . 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- . 1996. *Kinds of Minds*. New York: Basic Books.
- . 2003 (forthcoming). *Freedom Evolves*. New York: Penguin Putnam.
- Dwyer, William. 2001. Do knowledge, ethics, and liberty require free will? *The Journal of Ayn Rand Studies* 3, no. 1 (Fall): 83–108.
- Kane, Robert. 1998. *The Significance of Free Will*. New York: Oxford University Press.
- Machan, Tibor. 2000. *Initiative: Human Agency and Society*. Stanford: Hoover Institution Press.