

# **DELIBERATION, CHOICE, AND THE EMERGENCE OF FREE WILL**

**Jeffery L. Johnson**

**Professor of Philosophy  
Philosophy, Politics, and Economics  
Eastern Oregon University  
One University Blvd.  
La Grande, Oregon 97850  
[jjohnson@eou.edu](mailto:jjohnson@eou.edu)**

## **DELIBERATION, CHOICE, AND THE EMERGENCE OF FREE WILL**

### **[Abstract]**

I argue that a plausible compatibilist position of free will can be developed by treating concepts like choice and decision as emergent neurophysiological processes. The paper focuses on the concept of deliberation as a sophisticated cognitive ability that clearly emerges from our evolutionary history. I argue that deliberation is an informationally driven process where agents "talk" to themselves. The end state of successful deliberation is not so much a "jump" from a state of uncertainty to one of decision, but rather a process of discovering what one really wants and values.

## **DELIBERATION, CHOICE, AND THE EMERGENCE OF FREE WILL**

### **I.**

I have settled down to the task of writing these lectures and have drawn up my chairs to my two tables . . . My scientific table is mostly emptiness. Sparsely scattered in that emptiness are numerous electric charges rushing about with great speed; but their combined bulk amounts to less than a billionth of the bulk of the table itself. . . . There is nothing substantial about my second table. It is nearly all empty space.<sup>1</sup>

Here is a little philosophical fable. Once upon a time in a very possible philosophical world, generations of intellectuals talked themselves into a serious philosophical problem. They first intuited, and then in later generations accumulated great amounts of data that supported, the modern microphysical account of middle-size objects like tables and chairs. From this perspective, “there is nothing substantial about” tables and chairs; they are “nearly all empty space.” This spawned the notorious problem of solidity versus microphysics. How could desktops and the like really be solid, when they are mostly empty space?

In this philosophical world three quite different responses dominated the discussions in the finest journals and introductory textbooks. The first group of thinkers – we might call them “soliditarians” -- took it as obvious that tables were solid, and given what physicists had to say, that meant that solidity must be a very unique property. This tradition began in a time much friendlier to religion, and suggested that solidity was actually a completely non-physical property, perhaps the gift of a benevolent creator.<sup>2</sup> Later generations with more secular inclinations, speculated that it was an imperfectly understood product of quantum mechanics, chaos, and complexity.<sup>3</sup>

The polar opposite position – call them the “hard-microphysicists” -- enthusiastically endorsed the best science. We now understand that physical reality is one of mostly empty space, and this clearly implies that there really is no such thing as solidity. It is of course true that tables seem to be solid, but just as previous generations had abandon things as “obvious” as witches, phlogiston, and luminous ether, modern science forces us to rid ourselves of the illusion of solidity.<sup>4</sup> These thinkers were quite audacious, for they readily conceded that this illusion had all sorts of implications in other areas of the intellectual and cultural landscape. But they argued that radically altering widely-held moral and conceptual connections was a necessary price of scientific and intellectual progress.

The compromise camp – to coin a phrase, the compatibilists – were equally struck with the reality of solidity and microphysics. To these philosophers it was obvious that the perspectives must be logically compatible, since they were both manifestly exemplified in the real world. Ancient versions of the position were clearly identifiable,<sup>5</sup> but it was given new life with the advent of ordinary language philosophy.<sup>6</sup> The concepts of solid and non-solid performed recognizable linguistic, perceptual, and metaphysical jobs – the distinction was real, important, and clearly useful. If the solidity/microphysics problem seemed to imply the unreality of solidity, then the philosopher’s job was to offer an analysis of solidity that removed any logical tension between the perspectives.

“Soft-microphysicism” offered a redefinition of solidity in terms of dispositional properties. It proposed a subjunctive analysis of the meaning of solidity in terms of how solid objects would behave in certain circumstances. Most agreed that this was

insightful, and captured an unnoticed part of the concept of solidity, but many still worried that the analysis was subject to counter-examples, and ducked the really hard questions about the problem.

Later generations took the bull more squarely by the horns. They noted that there were clearly different levels of explanation, and that satisfactory accounts at one level did not perfectly fit into the idioms, indeed the conceptual schemes, of other levels.<sup>7</sup> For centuries theoreticians had offered reductive accounts where more complicated concepts and phenomena were accounted for in terms of more basic properties and components. Theories of supervenient properties simply inverted these accounts. In the same way that the biologist could propose that life supervened on biochemistry, or the cognitive scientist could speculate that consciousness supervened on neurophysiology, the compatibilist hypothesized that the solidity of the tabletop was a supervenient property of its underlying microphysical components.

I think that all would agree that the compatibilists had the best of this debate.

## **II.**

[W]e have most of the resources for a general, albeit rough, understanding of the free will problem. . . . Real progress has been made in our philosophical understanding of the free will problem, and the outline of a solution seems to be coming into view.<sup>8</sup>

I defend a compatibilist approach to free will and determinism. As a secular naturalist it is obvious to me that human beings are simply biological organisms, with complicated evolutionary histories, and that every part of us, including of course our psychology, is to be understood in biological and neurological terms. This doesn't exactly entail determinism, but it strongly suggests it, and I willingly endorse a prima

facie connection. At the same time, both free will and moral responsibility are manifestly also parts of our most basic understanding of ourselves. I can't see how these concepts could be "illusions" or vestiges of some superstitious cultural past. Thus, I am stuck with both determinism and free will. They must be compatible.

I suggest that determinism is most explanatorily powerful at the level of our underlying neurophysiology. Just as the microphysics of tables tells us that they are mostly empty space, contemporary neuroscience tells us that all psychological phenomena can be understood in terms of neural events with causal histories. Freedom, choice, and deliberation, however, are "higher" level phenomena that supervene on these neural events. All of this will wring hollow, of course, without some account of how this supervenience might work. Only in the past fifty years or so have cognitive scientists begun to understand the relevant story. I will leave it to the experts to articulate the neurophysiological details, and content myself with starting some of the necessary conceptual housekeeping – particularly some reflections on the connection between deliberation and free action.

### **III.**

The concept of an intentional system is a relatively uncluttered and unmetaphysical notion, abstracted as it is from questions of the composition, constitution, consciousness, morality, or divinity of the entities falling under it. Thus, for example, it is much easier to decide whether a machine can be an intentional system than it is to decide whether a machine can *really* think, or be conscious, or morally responsible.<sup>9</sup>

In a certain cell in my spreadsheet program I insert a relatively simple function. I ask the program to find values in nine other cells, and sum them. I also ask it to survey those nine values and find the minimum. I then ask the program to subtract the

minimum from the sum. Finally I ask it to divide the resulting value by eight. I do all of this because I have told the students in my Critical Thinking course that their grade will be determined on the basis of nine weekly quizzes, and that they may throw out their lowest score. My function calculates their grade. For student Sarah Smith that grade was 3.4, a B+.

What determined that value of 3.4 grade points? Many answers are possible. The underlying physical, chemical, and electrical reality guaranteed, if my computer hardware and software are working properly, that the value 3.4 would appear in the cell in question. Pages of incomprehensible (to me) computer code, also guaranteed (again, assuming all works well) that with the particular inputs I made for Sarah's file, 3.4 would appear. Finally, the whole purpose of my function was to get my spreadsheet to do the dirty work for me, so *ceteris paribus*, the rules of minimization, addition, subtraction, and division, dictate the 3.4.

Dennett defines intentional systems as follows:

[A] system whose behavior can be—at least sometimes—explained and predicted by relying on ascriptions to the system of beliefs and desires . . . [A] particular thing is an intentional system only in relation to the strategies of someone who is trying to explain and predict its behavior.<sup>10</sup>

My computer running the spreadsheet is an intentional system. It "believes" that Sarah's grades are entered in the nine cells, and it "desires" to do the math correctly. Adopting what Dennett now call the intentional strategy,<sup>11</sup> is crucial for me if I want to understand or predict Sarah's grade. The details of the underlying physical, electrical and chemical reality, even if I somehow magically could come to possess them, would be useless to me in predicting the number that would appear in the cell. And the

details of the computer code would be little better. I do know the simple mathematic calculations, though, so give me a pencil and paper, and a little time and I can easily predict the grade Sarah is going to get.

Dennett sometimes sounds as though the prediction of behavior is the only thing important about intentional systems and strategies. I intend to help myself to a somewhat richer notion of intentionality whereby we can assert with confidence that the system *really is* behaving intentionally. None of us believe for a second that my computer running its program has beliefs about students and grades, nor desires about mathematical accuracy or a job well done. I want to claim, however, that it is not just a manner of speaking, nor a predictive strategy, to say the system *really does* perform the arithmetic. It adds, subtracts, and divides just fine – more reliably, indeed, than you or I. Its ability to do all this is a supervenient ability that emerges out of its underlying physical and informational underpinnings. It would be absurd to claim that life was an illusion, because it was *really* just bio-chemical processes. In like fashion, doing addition, running a program, and a microchip being in a certain physical state, are all logically compatible, and empirically accurate, descriptions of the goings on in my computer.

#### IV.

Peter van Inwagen has argued on these grounds that a belief in determinism, if taken seriously, would make it impossible for us to deliberate. Van Inwagen argues that deliberation “manifests” a belief that it is possible *tout court* for the agent to perform her various alternatives, since one cannot deliberate about whether or not to perform some course of action unless one believes it to be possible *tout court*. Determinism implies that only one such action is possible *tout court*, and therefore a determinist cannot

believe that all of her supposed alternatives are in fact possible *tout court*.<sup>12</sup>

Philosophical thought-experiments and arguments can lead to insights and even new ways of thinking. We should always be suspicious, though, when the philosopher suggests, overtly or covertly, that we should adopt a new way of talking. Semantic legislation is a futile, and ultimately misguided, enterprise. Although I very much doubt it, I concede that we may discover that free will is an illusion, but I would argue that it is damn near impossible that we would discover that deliberation is an illusion. The difference, of course, is that the notion of free will is a term of art in philosophical discourse, while deliberation is most naturally at home in colloquial discourse. Deliberation is a kind of mental phenomenon or action that all normal human beings are intimately familiar with. Philosophers, psychologists, and cognitive scientists may teach us surprising and interesting things about deliberation, but their theories could never establish that it is not really there.

Let us reflect, therefore, on what happens when we deliberate about things that really matter, what Robert Kane calls self-forming actions.

[S]elf forming actions or SFAs occur at those difficult times in our lives when we are torn between competing visions of what we should do or become. Perhaps we are torn between doing the moral thing or acting from ambition, or between powerful present desires and long term goals, or we are faced with a (sic) difficult tasks for which we have aversions. In all such cases, we are faced with competing motivations and have to make an effort to overcome temptation to do something else we also strongly want. There is a tension and uncertainty in our minds about what to do at such times . . .<sup>13</sup>

Kane provides a highly speculative model of the underlying neurophysiology that occurs as we resolve this mental "tension and uncertainty." The cognitive scientist is better

equipped to assess his account. I want to offer a more phenomenological account of the process of deliberation involved in SFAs.

## **V.**

Colleagues on campus have recently started the process of forming a faculty union. I have always proclaimed a pro-union stance, and I sincerely believe that faculty in the state are seriously underpaid. At the same time, I have serious reservations about the composition, strength, and agenda of this proposed union. Should I remain silent and neutral, work for, or work against, this union? I am faced with deliberating on a personal and professional course of action that I consider to be "self forming" for the remainder of my career.

Serious deliberation, at least for people of my ilk, is a matter of talking to myself. I consciously recall things that I have said – "I would support unionization even if I were only a week away from retirement." I speculate about the motives of colleagues on either side of the debate. I have discussions with both administrators and co-workers, and reflect to myself on these conversations. I imagine future scenarios involving the legislature and strike votes. I try to foresee the campus culture in a union environment. And, of course, a host other things. The point of the very real and personal example, is that deliberation appears to be a linguistic, or at least, an informational process. When it strikes me that it would be institutional suicide to consider striking, or that I cannot shake my reservations regarding the motives of certain union organizers, it is the representational content of these thoughts that leads to the next consideration, or eventually my decision to oppose the union.

If some form of mind/body materialism is true, and my money says that it is, then underlying all of this are complicated neuro-chemical events in my central nervous system. Presumably all these events have causal histories. Does any of this imply that this is only the illusion of deliberation? On one plausible view it seems to. If the informational content of my thoughts are merely epiphenomena of the goings on in my brain, and those goings on are what really matter, then the fact that they are the product of earlier neuro-chemical events, does suggest a form of determinism that calls into question, not only the reality of free will, but genuine deliberation as well.

I want to suggest, however, that the causal arrow between phenomena and epiphenomena points in the opposite direction. It is my memory of the casual comment about unions years ago that explains my current brain states. Realizing that I hold suspect the motives of certain colleagues is similarly responsible for other brain states. In the same way it was the informational content of what was entered in my spreadsheet that explained what was going on at the level of electron and microchips, it is the semantic content of conscious deliberation that accounts for what is currently happening in my brain. From this perspective deliberation is very real; it occupies the driver's seat in the story that is being told, and plausibly results in a "freely chosen" decision to oppose the union.

## **VI.**

All of this, of course, is far from a complete theory of free will. For one thing, SFAs are not the only instances of personal behavior for which we hold agents accountable. And there still remains the nagging worry about how Kane's "tension and

uncertainty" is resolved through the process of deliberation. But hope for a robust compatibilist position survives.

Daniel Dennett has come the closest to articulating such a theory.<sup>14</sup> Very real cognitive states, including deliberation, can plausibly be taken to supervene on an individual's underlying neuro-physiology. Descriptions of these states and processes at the intentional level are just as legitimate, and just as real, as those at the neurological level. Thus the beginnings of an emergent theory of free will are already on the table.

And perhaps most importantly in the present context, we see the basic error in the accepted view of freedom and deliberation alluded to in the above quote about van Inwagen's, and Kane's, models. Kane talks of the "tension and uncertainty in our minds about what to do at such times [deliberating about SFAs]." It is far from clear that we perform an action, make a choice, free or caused, that resolves that tension. I didn't *decide* to work against the new union, I discovered that I opposed it. In a sense, my previous worries, judgments, and professional values, caused me to oppose the union. But rather than seeing myself as a mindless puppet, totally at the mercy of previous causes and events in my life, I see my opposition as emerging from the core of who I am as a person and a professional. I take full responsibility for my opposition. What more could we desire in a meaningful theory of personal freedom?

## ENDNOTES

1. Sir Arthur S. Edington, *The Nature of the Physical World* (Folcroft Library Editions, 1935), p. 3.
2. See, Richard Swinburne, *The Evolution of the Soul* (New York: Oxford University Press, 1986).
3. See, Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996).
4. See Saul Smilansky, *Free Will and Illusion* (New York: Oxford University Press, 2000).
5. See, John Locke (1690), *An Essay Concerning Human Understanding* (New York Oxford: Clarendon Press, 1975).
6. See, G. E. Moore, *Ethics* (1912).
7. See, Larry Wright, "Rival Explanations" *Mind* (October 1973), 82(328):497-515.
8. Smilansky, *op. cit.*, p.1.
9. Daniel C. Dennett, *Brainstorms* (Montgomery, VT: Bradford Books, 1978), p. 16.
10. *Ibid*, pp. 3-4.
11. See, *ibid*, and Daniel C. Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).
12. Hilary Bok, *Freedom and Responsibility* (Princeton: Princeton University Press, 1998). Chapter "Freedom and Practical Reason," reprinted in Gary Watson, editor, *Free Will* (New York: Oxford University Press, 2003), p. 152.
13. Robert Kane, "Reflections on Free Will, Determinism and Indeterminism," on-line: <http://www.ucl.ac.uk/~uctytho/dfwVariousKane.html>
14. Daniel Dennett, *Elbow Room* (Cambridge, MA: The MIT Press, 1966), and *Freedom Evolves* (New York: Viking, 2003).

---

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14