

***FREE WILL:
PHILOSOPHERS AND NEUROSCIENTISTS IN CONVERSATION***

edited by Uri Maoz (Chapman University)
and Walter Sinnott-Armstrong (Duke University)

QUESTION FROM NEUROSCIENTISTS: When Is an Action Voluntary?

ANSWER BY A PHILOSOPHER: Pamela Hieronymi (UCLA)

To borrow from Aristotle, “voluntary” is said in many ways. I here survey different contrasts one might want to mark with the words “voluntary” and “involuntary.”

I start with a theory-driven usage that connects to my chapter in this volume on “What Is a Will?” In answer to that question I sketched two broad, contrasting pictures. On the first, the will is a capacity first to “step back” from all that would influence us and to determine for ourselves (perhaps “endogenously”) how we will proceed. On this first picture, the will is a capacity for independent, or spontaneous, choice.

With this first conception of the will in hand, we could give a quick answer to the current question: Action is voluntary whenever it is the product of *this* special capacity, the will.

On the second conception of the will, the will is not a special capacity for independent or spontaneous choice, but rather is that collection of ordinary, influenceable, interacting aspects of the person’s psychology that generates intentional, or voluntary, or responsible activity. On this picture, willed activity is “free,” not because it is independent or spontaneous, but rather because it is the unhindered self-movement of a person.

On this second picture of the will, we identify a person’s will as whatever it is that generates a distinctive or significant kind of activity—intentional, or voluntary, or responsible self-movement. Thus, we cannot, on pain of circularity, identify voluntary action simply as the product of the person’s will.

If we work in this second direction—from significant activity in, so to speak, rather than from special capacity out—the question of when an action is “voluntary” is much less straightforward. We must ask which human activities are intended by, or up to, or the responsibility of, the person. And these are not restatements of a single question, but rather several interacting questions. Nonetheless, these are questions on which philosophical reflection has made some progress.

Philosopher Harry Frankfurt identified *action* as purposive movement guided by an agent. Action thus contrasts with purposive movement guided by something *other than* an agent. And, indeed, this is one distinction we sometimes mark with the word “involuntary.” Sometimes when we say that a movement is “involuntary,” we mean that, like blinking, digestion, or homeostasis, it is a reflex or anatomical response. These are purposive movements, but they are not guided by the agent. Call this “involuntary₁” movement. It contrasts with what we can call voluntary₁ movement—or, simply, “action.”

Notice that, in marking this first distinction, we need not start with a complete understanding of who or what an “agent” is, and work from the inside out. We can, instead, make progress by considering our intuitive, pre-theoretical understanding of the difference between, on the one hand, activities such as digesting and, on the other, actions such as cooking—we can then work from our understanding of such activities, in, so to speak, to understanding of the agent. We might, for example, identify sufficiently flexible activities, those open to intelligent updating in novel situations, as those guided by the agent (as actions, voluntary₁ movements). We would thereby learn something about agents (they are capable of intelligent updating in light of new information).

Typically, in order to identify an event as a human action, rather than a mere happening or an involuntary₁ process, we identify it as something that happened because the person *meant* for something to happen. Further, it is not important, in identifying an event as an action (a voluntary₁ movement), that it was spontaneous or that the person meant it independently of external influence. You might be commanded to act, or forced to act by your circumstances, or coerced. You nonetheless act—your behavior is not thereby rendered mere involuntary₁ movement. In my opinion, an important building block in scientific understanding is identifying the neurological differences between voluntary₁ and involuntary₁ movement—between those movements that are guided by the agent, those that happened because someone meant for something to happen, and those that are not.

Moving on: In identifying an action as something that happened because someone meant for *something* to happen, we must remember that what happened may not be what was meant. You meant to send the email to your mother, but instead you sent it to your supervisor. Sending the email to your supervisor was an action—in fact, it was your action. It was not an involuntary₁ movement. Yet, we would not say that you sent the email to your supervisor voluntarily. Your action was not involuntary, yet also not voluntary. We need another distinction. Call this voluntary₂: An action is voluntary₂ if you meant to do *it*—voluntary₂ actions are done on purpose and successfully.

We could add, at this point, complication about unforeseen or foreseen but unintended consequences. Michael Bratman, a philosopher of action, has offered a theory to map this terrain. We would then be studying what Bratman calls “planning agency,” which one might think of as a kind of “executive capacity.” Planning agency is the ability to intelligently,

successfully bring about that which one (in some way) represents as to be brought about, the ability to form and execute intentions.

Again, if we were to study this, it would not be important that your plans are isolated from external influence, and thus it would not be important, in experimental design, to avoid commanded actions or actions done on cue. The fact that you were told to email your mother would not make your emailing her any more or less an exercise of your planning agency—no more or less voluntary₂.

However, we have not yet captured all we sometimes mean when classing an action as “voluntary.” Aristotle provides what many think of as a paradigm instance of involuntary action: sailors throwing cargo overboard during a storm.¹ The sailors throw the cargo overboard against their will: they are forced to do so, by the storm. But throwing the cargo overboard is not an involuntary₁ movement. In fact, it is a voluntary₂ action: they set and successfully achieve their aim. Yet, we want to say it was involuntary, because they did it only under duress.

Thus, we need yet another distinction. Let us say that the sailor’s action was involuntary₃. Cases of involuntary₃ action contrast with cases in which you are, we might say, happy with your choices. Duress is not the only source of unhappiness. We can add coercion, a sense of obligation, obedience, or need. In contrast, to act voluntarily₃ would be to act, so to speak, as a *volunteer*: to do what you do happily, without unwanted pressure—when you act, as we sometimes say “of your own free will.”

The problem encountered by the sailors, as well as by those subject to threats or coercion, is still *not* that they are subject to external influences or that they fail to act spontaneously. It is, rather, that all of their available options are contrary to their preferences. (If they were instead given an attractive offer for their cargo, they would then unload the cargo voluntarily, of their own free will—despite the fact that the offer provided a strong external influence.) Likewise, the problem with acting from obligation or need is that it does not align with your preferences. Lacking good options that align well with your preferences is a familiar and straightforward enough difficulty. But it is not a problem with your capacity for action—it is rather a problem with the circumstances in which that capacity must operate. It is, so to speak, a problem in life. It would, I think, be odd to think that the difference between acting as a volunteer, happily and without unwanted pressure, and acting under coercion, duress, obligation, command, etc., is to be studied by studying the human capacity for agency, rather than by considering the circumstances in which such a capacity operates.

But, we are not yet done. Not every case of failing to act “of your own free will” due to unhappy life circumstances are cases of failing to act as a volunteer: By use of deceit, indoctrination, propaganda, or clever advertising, people can be *manipulated* into acting as a volunteer. The denizens of Huxley’s *Brave New World* act voluntarily_{1, 2, and 3}—they successfully form and execute plans, while acting happily, as volunteers—and yet they lack another important

¹Aristotle, *Nicomachean Ethics*, trans. Terence Irwin, 2nd ed. (Oxford: Oxford University Press, 1999), Book 3.

form of freedom. If one wants to capture this form of freedom with the word “voluntary,” one will need yet another distinction: One might say that are acting involuntarily₄.

Once again, the problem with involuntary₄ action is not that those who are manipulated, deceived, indoctrinated, etc., are subject to influence, *per se*. Rather, the problem is that they are subject to a *problematic* sort of influence, a kind of influence to which people ought not subject one another. It is famously hard to know how to draw the line between, say, education and indoctrination, coercion and persuasion, or threats and offers. But drawing these lines will not be a matter of locating a particular capacity in the individual that typically operates free of influence (or free of some degree of influence) and identifying when it is influenced (or influenced to a greater degree). It will, instead, be a matter of determining which *ways* of influencing others are permissible and which are problematic. That is to say, drawing *these* lines will be a matter of ethical reflection, not scientific or metaphysical discovery.

My suggestion, then, is that neuroscience should study (or, more modestly, should *first* study) both the difference between voluntary₁ and involuntary₁ movement and our capacity for voluntary₂ action—for successful planning agency. The forces that render our activities involuntary₃ or involuntary₄ are problems encountered by humans acting in a difficult world—they are problems in life, not problems with our capacity for significant human action.

FOLLOW-UP QUESTIONS:

Liad Mudrik:

When the author contrasts agent-guided purposive actions and non-agent-guided purposive movements, like reflexes (following Frankfurt), does the difference rest on the agent’s conscious state? That is, what determines whether an action is guided by an agent? Is it the agent’s consciousness of the purpose? If so, is it accurate to say that consciousness of the purpose is the criterion for an action being voluntary? And if so, can any voluntary action be unconscious?

Walter Sinnott-Armstrong:

How are the four kinds of voluntariness that you distinguish related to moral responsibility? Are agents responsible for all of their voluntary acts? Are agents not responsible for any of their involuntary acts?

Uri Maoz:

You propose that neuroscience should focus on distinctions between in/voluntary₁ and in/voluntary₂. If this is because you think that more philosophical work needs to be done to better

understand the distinctions between in/voluntary₃ and in/voluntary₄, that is one thing. But, if you think that those latter distinctions are beyond neuroscience or that neuroscientific investigations are not useful there, please elaborate on your reasons for this.

In particular, one reason that we care about voluntariness in actions relates to moral responsibility. Supposedly agents whose acts are involuntary₁, involuntary₂, involuntary₃, or involuntary₄ are somehow less responsible (perhaps to different degrees?) than those who act voluntarily on all those accounts. So, one reason that we might be interested in a neuroscientific account of the voluntariness of these actions is to possess a more objective account of an agent's responsibility, if any. To that extent, it appears that neuroscience can potentially help with the distinctions between voluntary and involuntary 1, 2, 3, and 4. While it might be more straightforward to understand how to draw neuroscientific distinctions between in/voluntary₁ and in/voluntary₂, there is no reason to conclude that understanding whether a person is happy with their choice—as in the in/voluntary₃ distinction—is beyond neuroscience. For the in/voluntary₄ distinction the difference lies not within the agent but within the psychological or brain states of those influencing the agent. And those could be further studied by neuroscience. So, neuroscience could also be brought to bear on in/voluntary₄. Do you agree?

Gabriel Kreiman:

Imagine that you can measure every possible neuroscientific variable of interest: action potentials of every neuron, the concentration of every ion in every cell; you name the variables you want to measure, and you have them all! Is there any measurement that could tell you whether an action was voluntary or not? If you think that neuroscience is not sufficient, then feel free to add whatever variables you want, as long as they are really measurable things. You can add all the cells in the liver, the position of all the stars in the universe, or any kind of behavior, but it has to be a physically measurable variable and cannot be a “will” or an “intention”, unless you tell me how to measure those things. Is there any empirical measurement that can tell us whether an action is or is not voluntary (in any of the four senses that you distinguish)? Is that a problem?

REPLIES TO FOLLOW-UP QUESTIONS

Pamela Hieronymi (UCLA)

The distinctions I've drawn between different senses of “voluntary” concern the employment of a particular human capacity, the capacity for what might be called self-guided movement (voluntary₁ movement). The first distinction marks whether this capacity is in operation at all; the second marks the *skillful* and *successful* deployment of that capacity

(voluntary₂); and the third and fourth mark whether the operation of that capacity is in some way hindered, constrained, or problematically interfered with or manipulated (voluntary₃ and 4).

Reply to Mudrik

The difference between self-guided (voluntary₁) movement, such as cooking or opening a door, and movement that is guided in some other way (involuntary₁ movement), such as digesting or blinking, is not whether either the movement or its guidance is something of which the agent is conscious. I am sometimes conscious of my digestion, or of my blinking; the song going through my head is itself a conscious state. None are guided by me, in the relevant sense. On the other hand, I typically am not conscious of the ways in which my fingers move to grasp the doorknob, even as I grasp it purposively. Likewise for the ways I move my joints in order to catch the ball you threw. My conscious awareness of goings-on, or of their purposes, or of how they achieve their purpose, does not correlate well with whether or not they are self-guided, voluntary₁.

Reply to Sinnott-Armstrong

Questions of responsibility overlap with these distinctions in voluntariness, but do not map neatly onto them, because questions about responsibility concern, not just the deployment of our capacity for self-guided movement, but also the expectations or demands we can reasonably put on one another. I might be responsible for my digestion, or my allergies, because they are things I am rightly expected to manage. I am responsible for my short temper and my distrust, not only because they are things I can manage, but also because they reveal my take on who can be trusted and what is of value—despite the fact that these are not voluntary. If I am acting under the influence of a powerful drug, or suffering from a temporary delusion, then I might not be responsible for something I did voluntarily₁ and 2. Likewise, I might not be culpable for something I did voluntary_{1, 2, and 3}, because I had been systematically deceived.

Reply to Maoz and Kreiman

The first and second senses of voluntary pick out a particularly important capacity of individual human beings—activity that is self-guided, and often skillfully and successfully so. Given a theory, we can figure out how to measure these. In the third sense, the agent either acts as a volunteer or else is in some way unhappy with their situation. Insofar as unhappiness is itself a psychological state, perhaps unhappiness interacts in interesting ways with the capacity for self-guided movement. To study this interaction, we would need some way to identify the two.

Some neurological studies adopt what seems to me a peculiar interpretation of “voluntary.” They focus on activity that is “endogenous” in a very specific sense: The subject is

not told exactly what to do or exactly when to do it, but is instead asked to pick from a menu of options (left or right, this or that) at some unspecified point, after a certain point in time. Granted: we are able to do this. We are able to resolve uncertainty, to do this-rather-than-that, now-rather-than-then, and to do so for no reason other than the need, or the desire, to move on. This ability is an important component of our capacity for skillful, successful self-guided movement; without it, we would not be able to execute many of our plans. But this particular ability seems to me just one part or aspect of our larger ability to execute our plans skillfully and successfully.

Further, I would argue that those moments at which we employ this “endogenous” ability are not ones in which our activity is especially “voluntary” or “free.” The mere presence of “exogenous” factors which you might use to resolve the uncertainty (or which resolve the uncertainty for you, so to speak) does not itself show that your capacity for self-directed action has been in any way interfered with, hindered, or constrained. Nor does the presence of such factors itself show that what you do is any less self-directed or one’s own (unless one simply identifies the self with this capacity for endogenous action, adopting something like what I called the first conception of the will in my chapter in this volume on the will). The presence of exogenous factors may in fact be part of your own complex plan. Whether or not they are, responding to them, or taking them to be reason-giving, may be something you do happily, as a volunteer.

The fourth sort of “voluntariness” (or, better, “freedom” or “liberty”) concerns the ethical question of *which* ways of influencing others impinges on their freedom, understood as an ethical or political ideal—which ways of influencing others are illicit, ethically or politically. Some people think this ethical or political question should be answered by determining the conditions under which a person acts “freely” in some *other* sense; they would try to draw the line between offers and threats, or between persuasion and coercion, by first identifying the point at which the person’s capacity for self-directed movement has been overpowered. They would suggest that the ethical line is crossed when that capacity is overpowered.

I think this is mistaken—we do not draw these lines by considering when someone’s capacity as been overpowered (when I hand over my money to save my life, I am acting voluntarily^{1 and 2}; and, the fact that I am unhappy with the terms of our agreement, that I do not enter it voluntarily³, does not show that you made a threat rather than an offer). Rather, we first determine which way of influencing others are unethical—which ways are disrespectful, and so count as coercion or threats or manipulation, rather than persuasion, offers, or influence—and then facts about whether the person was free, in this final sense, follow from this.

If I am right about the order of explanation, then, while impingements and constraints generated by ethically illicit behavior will, of course, be both constraints *on* and constraints *designed by* the operation of human brains, the relevant variables are unlikely to be neurological. An attempt to motivate the thought: the brain of a Beta, in *Brave New World*, might, over some stretch of time, be neurologically indistinguishable from mine. Perhaps we have both, to this

point, lived indistinguishable, unremarkable lives. Yet the Beta's freedom is greatly diminished, due to her larger political situation—she has been systematically manipulated, while I have not. Of course, if we take a wide enough picture of space-time, we will find differences (there is “global supervenience” of the social on the physical), and some of them will be in human brains. But the fact that there will be such differences should not lead us to expect that the best way to study the phenomena we are after is by looking at those differences.)

REFERENCES

- Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. 2nd ed. Oxford: Oxford University Press, 1999.
- Bratman, Michael E. *Intention, Plans, and Practical Reason*. Cambridge: Cambridge University Press, 1987.
- Frankfurt, Harry. "The Problem of Action." In *The Importance of What We Care About*, 69–79. Cambridge: Cambridge University Press, 1988.