

## Conscious Willing and the Emerging Sciences of Brain and Behavior

Timothy O'Connor

Department of Philosophy  
Indiana University  
Sycamore 026  
Bloomington, IN 47405  
toconnor@indiana.edu

**Summary.** Recent studies within neuroscience and cognitive psychology have explored the place of conscious willing in the generation of purposive action. Some have argued that certain findings indicate that the commonsensical view that we control many of our actions through conscious willing is largely or wholly illusory. I rebut such arguments, contending that they typically rest on a conflation of distinct phenomena. Nevertheless, I also suggest that traditional philosophical accounts of the will need to be revised: a raft of studies indicate that control over one's own will among human beings is limited, fragile, and – insofar as control depends to an extent on conscious knowledge – admitting of degrees. I briefly sketch several dimensions along which freedom of the will may vary over time and across agents.

**Keywords:** Free will, choice, consciousness, self-awareness, intention, willing, reasons, automaticity, Libet.

There is a wide panoply of philosophical views concerning the role of the will in human freedom of choice and action. Philosophers aim to give schematic accounts of the conditions that a person's willing and carrying out a course of action must satisfy to be appropriately accounted as free and subject to moral praise or blame. For my purpose here, it is enough to note only the main families of views, and these only briefly. Since at least the seventeenth century, it has been common to divide philosophical accounts of freedom in terms of whether or not they regard freedom as compatible with causal determinism in the natural order. On "compatibilist" views, choice is a *mechanism* by which an agent's predominant desires and intentions, along with corresponding beliefs concerning how they may be realized,

ineluctably initiate appropriately matching actions. On most compatibilist views, what distinguishes will or choice from other natural mechanisms is that it is, in John Fischer's term, *reasons-responsive*, a mechanism that is activated by the agent's own goals and desires. Some will add further requirements to this unadorned picture. For example, Harry Frankfurt and others maintain that a will that is free must have a hierarchical structure, such that the first-order desires that cause one's choices are ratified by a higher-order willing that one be moved by the first-order desire – a willing that one be the sort of person who is regularly moved to act by that sort of motivation. (Addicts might be people who typically do *not* will to be moved by the desires that in fact move them towards satisfying their addiction, and thus are judged not to be free in these willings.) On all compatibilist views, simple or adorned, the capacity to form choices or acts of will is embedded in a seamless causal flow of events constituting the universe's history. (Any causal indeterminacy in the decision-making process is at best irrelevant to what makes for freedom and at worst potentially undermining.)

The development and refinement of compatibilist views continues apace. But the last couple of decades has also seen a resurgence of theorizing about the will along "incompatibilist" lines. All incompatibilists who deem free will to be possible maintain that choice is some sort or other of indeterministic capacity by which I "directly" determine my choice while guided by reasons. This power of direct determination, or choice, is then elucidated in different ways: first, as a nondeterministic variant of the compatibilist account on which it is the power of one's beliefs, desires, and prior intentions to cause one's choices; second, as a *sui generis* noncausal variety of active power; or third, as a *sui generis* agent-causal power (a primitive capacity of a person to form an executive intention to act in a specific way). The way reasons are thought to guide the exercise of such a power of choice is either efficiently causal (albeit nondeterministic) in nature, or it is purely teleological, or it is some hybrid of these.

While we see a wide range of views, we should not lose sight of the fact that most of them are trying to articulate the very same features of freedom that seem compelling from a subjective, pretheoretical point of view. We may mark this underlying commonality by observing that, on all these views, the freedom and responsibility of human beings requires our having and regularly exercising three very general capacities: (1) an awareness of and sensitivity to reasons (including especially moral reasons) for actions; (2) the ability to weigh and even to critically probe our own desires and intentions and to reevaluate on occasion our overarching goals; and (3) the ability to choose, based on reasons, which action we shall undertake on a given occasion.

Perhaps the most striking feature in philosophical discussion of free will in very recent years is the rise of free will *skeptics*, who wish to challenge all the accounts indicated above. This skeptical view strikes at the core datum of traditional accounts of the will by maintaining that the experience of conscious will is illusory: we do not freely control *any* of our actions via the conscious exertion of will, however this may be characterized. (Their more measured cousins, the *revisionists*, recommend that we drastically revise the ordinary understanding of free will so as

to make it suited to the likes of you and me, who [according to the psalmist] were created a little – say the revisionists, a *lot* – lower than the angels.) An older skeptical view (one that continues to this day) was rooted in a judgment that the very idea of freedom of the will resists coherent articulation. According to the new-wave skeptics, however, the problem is thought to be empirical: the efficacy of conscious will indicated by our subjective experience of agency runs counter to mounting evidence from the sciences that in one way or another touch upon volition and the genesis of human action. Such a conclusion has been endorsed by some of the relevant scientists themselves and thereby has begun to enter the popular imagination.

In what follows, I will offer reasons for resisting this empirical argument for the skeptical view. I will first try to show that there are conceptual confusions underlying some of the planks in the skeptic's case. I will then make a start at confronting the challenge that remains, once the confusions are identified and set aside.

## 1 The Sciences of Volition and Agency: Empirical Challenges to Free Will on Three Fronts

I will begin by summarizing some of the main empirical findings of relevance to the existence and nature of free will in human beings. These come from three broad research areas: neuroscience, the study of clinical mental disorders, and social psychology.

### Neuroscience

Over the past three decades, neuroscience has delivered a variety of fascinating and sometimes surprising results concerning human action. Here I will consider (1) cases where actions are artificially produced but give rise to *unwitting confabulations by agents concerning their own agency*, (2) cases where *unperceived environmental stimulation of the brain can significantly influence seemingly free choices*, and lastly (3) the much-discussed studies concerning the *timing of preparatory brain activity vis-à-vis the experience of agency*.

In the confabulation cases, neurosurgeons operating on the brains of conscious patients were able to induce behaviors such as the raising of a hand by electrically stimulating motor-control areas of the brain (Delgado 1969; Gazzaniga 1994). We have very good reason to believe that the patients did not form a conscious choice to move their arm. Yet when asked immediately afterwards why they moved their arms, they tended to confabulate rational explanations – apparently sincerely – such as “I was trying to get your attention.”

In the external brain stimulation studies, subjects were asked to freely choose between moving either the left or right index finger when signaled. While they awaited the signal, a large magnet was secretly moved across the motor area of either the left or right side of the brain. It turns out that subjects showed an increased tendency to move the finger contralateral to the side that was stimulated,

while believing that that they were choosing voluntarily and without any discernible external influence (Brasil-Neto et al. 1992).

Finally, and most famously, there are Benjamin Libet's (1985) studies on the timing of the experience of willing (studies that have since been refined by Patrick Haggard, H.C. Lau, and others). Libet devised a study in which people are asked to wiggle their finger within a short interval of time (thirty seconds or so). The experimenter instructs them to do so whenever they wish – though spontaneously, not by deciding the moment in advance. Throughout, they are to watch a special clock with a very fast-moving dial (a beam of light) and note its location at the precise moment at which they felt the “urge” or “wish” to move the finger. During the experiment, a device measures electrical activity on the agent's scalp. Libet discovered that a steady increase in this activity (dubbed the “readiness potential,” or RP) consistently preceded the time the agents cited as when they experienced the will to move. By averaging results over hundreds of experiments, Libet determined that the RP preceded the “experience of will” by an average of some 400 milliseconds, a significant interval in the context of neural activity. Libet and others concluded from this result that “conscious will” is not the initiator of voluntary acting but instead a *consequence* of an unconscious physical process that also (and according to some hypotheses, independently) triggers the action.

### **Clinical Mental Disorders**

I turn next to clinical medical disorders. There are quite a number of abnormal clinical phenomena, but two deserve special attention in considering the relation of the will to purposive action. In anarchic (or “alien”) hand syndrome, complex movements of the person's hand are carried out in a smooth way towards the execution of an apparent goal (taking someone else's glass of water, say), yet the person claims not to have intended them, and they are generally unwanted, causing embarrassment. (Many of us think here of Peter Sellers' title character in Stanley Kubrick's film *Dr. Strangelove*.) And some schizophrenic patients report the perception that other agents are controlling their actions, even though there is good reason to believe that the actions in fact issue from intentions of the agent, albeit unconscious ones. In both kinds of cases, agents are acting purposively and under no discernible control by an external agent or source, yet they lack, in some sense, the feeling of being the authors of their actions. Some free will skeptics draw attention to these cases because the “sense of freedom,” or the sense of being in control of what one does, is arguably a fundamental basis of our belief that we are in fact free to choose, in some measure. And their thought is that these and other cases make plausible the idea that the sense of agency is not in fact anything like a direct perception of our own agency, even in normal cases: it may have a distinct physiological source from the causal pathway that leads to purposive choice and action.

### Social Psychology

Finally, we will consider the more low-tech studies in social psychology. In his provocative and entertaining book, *The Illusion of Conscious Will* (2002), psychologist Daniel Wegner elaborates a number of findings that show that people in general are profoundly susceptible to the *inducement of false beliefs regarding one's agency*. In one study, participants are invited to move a computer mouse in concert with someone who is a secret confederate of the experimenter. They are told to choose freely where on the screen to move the cursor, although in fact the confederate is gently forcing the selection. When, just before the mouse is moved, subjects hear in their headphones a recording of a word (e.g., "swan") corresponding to the confederate's chosen target, they have an increased tendency to report that *they* acted intentionally in making the selection. Other studies have shown that prompting subjects to have hostile thoughts about someone prior to the person's appearing to have suffered negative consequences makes the subjects more inclined to view themselves as responsible for the outcome. (For example, you tell the subject to think negative thoughts about a person before performing a voodoo curse on a voodoo doll and have the object of the curse feign a headache.) Likewise, envisioning a positive outcome before a favorite team's televised sporting event leads to increased feelings that one has somehow influenced the outcome. Running in the opposite direction, persistent testimony of others can lead people to believe falsely that they previously performed an action. And yet other studies have shown malleability in one's illusory sense of direct control over *another* person's actions.

These are all examples of *induced false beliefs regarding agency*. A second type of finding in social psychology is the surprising degree to which often *unrecognized situational factors influence individual moral choices*. A variety of studies indicate that the percentage of subjects willing to help someone in need will vary significantly or even dramatically depending on such factors as how many other persons were perceived to be in a position to help, how busy the subjects were at the time, and even whether ambient noise or odors differed from normal levels (Doris 2002; Doris & Stich 2006). It appears that these situational factors may in some cases be better predictors of behavior than the subject's general character traits, as self-reported. The challenge that these findings may pose to human freedom, in the view of some (Nahmias 2007 and forthcoming), is twofold: (1) our actions may be heavily influenced by arbitrary situational factors that we do not control and whose influence we are often unaware of; and (2) ever since Aristotle, it is common to ground human responsibility in our making choices that over time help to form a character out of which much subsequent behavior flows. The skeptical claim is that general moral character is not as significant in explaining our behavior as this presupposes.

## 2 Some Conceptual Tools in Aid of Deflating Some of the Challenges

We now have before us some of the empirical results thought to threaten the view that human beings are freely directing their actions via conscious willings. In thinking through the several cases, the distinctness of a number of agency-related concepts needs to be born in mind. Consider, then, each of the following:

First, three items in the category of *will or desire*:

- *Minimally voluntary action*: an action that unfolds “automatically,” rather than being consciously willed, while coinciding with one or more of one’s desires or intentions, conscious or unconscious, and lacking signs of external or internal compulsion.
- *Willing, or conscious forming of an intention to act*: a purposive and executive, or action-initiating, event. Note that these come in two sorts, one that is present-directed (deciding to act here and now) and one that is future-directed (deciding now to act at some particular future time or upon recognition of an appropriate stimulus).
- *Urge or want*: a felt desire to perform an action that one may or may not satisfy.

Next, two sorts of *belief*:

- *Belief concerning one’s action*: We often have beliefs concerning what we are about to do (and why), beliefs concerning what we are now doing (and why), and beliefs concerning what we have recently done (and why). It is an empirical question the extent to which these types of belief are aligned.
- *Belief concerning the wider causal import of one’s basic actions*: we also have beliefs concerning the more or less immediate effects of our basic actions (those mental actions or bodily movements that we directly control). These are generally inferred from observational clues, though rarely is the inference consciously made.

Finally, two sorts of *experience*:

- *Experience of willing or intention-formation*: when we consciously decide a course of action, there is an “actish phenomenal quality” (Ginet 1990) – we experience ourselves as willing or intending our basic actions. (*Occasionally*, as when we struggle to reach a difficult decision, this involves the further quality of *effort*.) It is both an empirical and a theoretical question as to the relationship of this experience to the willing/intention itself.
- *General “sense of authorship”*: It seems correct to say that there is a comparatively persisting yet phenomenally less distinct experience of being the author of what we do, an experience that coincides with any sort of activity of which one is minimally consciously aware. This has no evident conceptual or even

evidential relationship to the spontaneous, natural belief that our actions are metaphysically free, though it may help causally to sustain that belief.

It should be fairly uncontroversial that each of these categories pick out real and distinct phenomena (with the possible exception of the controverted category of conscious willing itself, which is in any case *conceptually* distinct from each of the others). Yet the attentive philosopher who reads recent studies in the psychology and neuroscience of volition will frequently be hard-pressed to map references he encounters to “experience of will” and “conscious decision,” terms that generally go undefined or underdefined, onto one or another of these categories with any confidence. In any case, when we do bear these distinctions firmly in mind, we can readily see that *some* of the empirical findings I canvassed do not pose any obvious threat to the belief in human freedom.

### 2.1 Ex Post Facto Confabulation

Gazzaniga’s neurosurgical patients who gave reasons for the arm movements that were artificially induced were clearly confabulating after the fact, perhaps due to a natural psychological mechanism that attempts to impose coherence between one’s beliefs and observations concerning one’s own movements. And social psychologists have shown that one can cause someone to confabulate the very occurrence of an action in order to produce coherence with the seemingly sincere testimony of others claiming to have observed such an action. But the scientist who notes these cases of unwitting, ex post facto formation or revision of actional beliefs should not be tempted to be partly complicit in them by agreeing that there was, after all, an *experience* of willing that fits the fabricated version of events! What we have here are unremarkable instances of our occasional penchant for forming false memories, rather than cases of illusory experiences of will, which is the skeptic’s purported target.

### 2.2 Erroneous Beliefs Concerning the Wider Effects of One’s Actions

Consider next the studies indicating that what is termed a “sense of agency” involving environmental outcomes one is not in fact controlling, such as another person’s falling ill or a sports team’s performance, can be induced *to some degree* by having the subject perform actions as simple as intentionally having certain thoughts. Whether in these cases there is the same “sense of agency” as generally accompanies ordinary intentional action is very doubtful. But what clearly is occurring is that subjects are being caused to form false beliefs concerning the *effects* of their basic actions. That we can easily be led to increased inclinations towards such beliefs absent good evidence is disconcerting, but again it seems besides the point when the question is our control over our own basic actions via the will.

### 2.3 *Automatisms: Minimally Voluntary Actions Unaccompanied by Any Feeling of Agency*

These strange cases include alien hand syndrome and the schizophrenic belief that someone else is controlling one's behavior from within. We might also include nonclinical, episodic instances among perfectly normal subjects, such as the table-turning phenomena among nineteenth-century spiritualists. (People seat themselves around a table with fingertips lightly touching the edge. As they await a message from a dead person, sensitive instruments reveal that, one by one, they begin ever so slightly to rotate the table – while quite sincerely believing that they are entirely passive observers of this event.) Such cases strike closer to the skeptical target insofar as they show that a feeling of agency is not necessary to *minimally voluntary* agency. However, such cases should be seen against the backdrop of the more pervasive fact of *automaticity*: for a great deal of our behavior, there is a feeling of authorship, yet we do not directly will or intend the actions in a conscious way. (Routine or learned skilled behaviors are the clearest instances, and they also underscore the usefulness of automaticity. As William James observed, we walk best along a beam the less we attend to the position of our feet upon it.) This pervasive automaticity has long been absorbed into our pretheoretical understanding of ourselves as agents, and, accordingly, it is not these sorts of actions that we take directly to involve conscious will. The automatisms of anarchic hand and severe forms of schizophrenia are unusual and striking only because the unconsciously generated action is at odds with the agent's standing intentions and is unaccompanied by the usual "sense of authorship." That subconscious processes can generate intentional behavior with these characteristics says nothing about the truth or falsity of our pretheoretic conception of ourselves as consciously and freely willing what we do on *other* occasions, just when we take ourselves to so will. They are not evidence in favor of some skeptical model of human agency over against a model on which we enjoy significant freedom because their data are not contrary to what our ordinary pretheoretic understanding would expect.

Instead of making the case in terms of a supposed *disconfirmation* of the ordinary view, on which the commonsense "hypothesis" is shown to predict consequences that are contrary to fact, a skeptic might try to argue alternatively that they are simply *better explained* by an alternative model, such as the one Wegner proposes, on which all actions are generated by subconscious mechanisms into which "conscious will" does not enter. I will not try to develop and respond to this form of the skeptical argument in any detail. But I note that it involves the difficult and contentious matter of whether, and how deeply, we are rationally entitled to start empirical inquiry with a strong presumption that our conviction that we (sometimes) direct our own actions via conscious willing is correct. The unavoidability of unargued yet essential starting assumptions is a familiar fact of life to philosophers, whose training, if not everyday practice, involves the contemplation of all manner of radically skeptical hypotheses about the world and our knowledge of it. (What evidence can show that my lifelong perceptual experience is not a continuous dream, or the product of some nonveridical source such as the *Matrix*?) It is a



different matter for scientists, of course, whose research is entirely unperturbed and unguided by such fanciful philosophical queries. This is a problem only insofar as it may induce in some scientists a false understanding of the scientific enterprise as resting on no foundational assumptions whatsoever. It is commonly recognized that science must presuppose that experience is not wholly illusory – there is a world external to our minds, roughly of the sort experience suggests; that our foundational forms of reasoning, deductive and inductive, are sound; and that laws that hold in the observed parts of the universe can reasonably be generalized to the unobserved. It is less often recognized that scientists must presume the fundamental reliability (not perfect, but to a high degree) of their *methods of intervention* in testing theories. That is to say, if we are to accept the data that are collectively adduced in a given domain, we must presume that it resulted from individual processes that, *inter alia*, reflected the actual intentions of those who collected it. Were it not the case that in the conduct of an experiment, scientists are reliably aware of what they are doing and why they are doing it, we would have little reason to accept the significance of their reports.

#### 2.4 Libet Cases

I now turn to Libet's findings (replicated by others) that in his experimental set-up there is a smooth build-up of electrical activity in the motor cortex a few hundred milliseconds prior to the time the subjects indicated as the onset of the conscious urge or wish to perform the movement. Let's begin by noting that in the experiment a subject agrees at the outset to perform a specific action within a short interval of time. All that is left to be determined by the agent is the precise time of its occurrence. Though there is more that needs remarking on, we should not ignore a very obvious fact here: in agreeing to cooperate with the experiment as described, *the agent has already decided to perform a specific action*. In our terms, she forms the future-directed intention to  $\emptyset$ , for a specific action-type  $\emptyset$  (say, flicking her wrist or wiggling her finger). No evidence is adduced that there is a slow build-up towards a readiness potential before *this* sort of decision. Libet, however, will say that this is inconsequential. For the agent goes on to form another, ostensibly free, intention to perform the action here and now, and his choice is shown to have a neural antecedent associated with the movement itself.

But is this really so – is this choice of timing a prototypical instance of a willing that is experienced as wholly spontaneous? As Al Mele (1997) points out, in describing the instructions that the experimenter gives to the subject, Libet (1985) uses several terms interchangeably: “urge,” “desire,” “wish,” and “intention.” The subject is asked to note the precise timing of such an urge/desire/wish/intention for each of several movements that will be performed in a single setting. This experimental set up and set of instructions, it seems to me, invites an interpretation differing sharply from Libet's own. First, the action is hardly a “spontaneous” one. Even though the subjects are instructed not to pre-plan the timing of their actions, and instead wait for the urge/desire to do so, the action *type* itself is pre-planned, and even its timing is to a significant degree, as they are instructed to make the

movement within a thirty-second or so interval of time. Secondly, by asking the subjects to carefully introspect to pinpoint the timing of the impulse to move, the experimenter is inviting the subject to adopt the role of observer in relation to his own conscious experience, and specifically to wait for an unplanned *urge* to occur. This certainly encourages a passive posture. Having decided that one will move, one looks for the *urge* to do so in order to act upon it. I suggest that such a pre-formed intention to act upon the right internal “cue” initiates an unconscious process that promotes the occurrence (or perhaps *evolution*) of a conscious state of desire or intention that is not actively formed. In context, the state’s default is to trigger the pre-planned activity, absent a last-second “veto” by the agent. In another study, Libet confirms the possibility of such vetoes subsequent to the experience of such an *urge*.

Given the plausibility of this alternative interpretation – on which the subject’s actions are not *prototypical* spontaneous conscious willings – Libet’s pioneering studies do not have anything like clearly negative implications for human freedom. To be sure, they have evidently opened up a fruitful field of neuroscientific inquiry and are helpful in pointing to the need for greater attention to the precise phenomenology of different instances of willing or desiring to act. One wonders, for example, whether some or all of the subjects experience a growing anticipation that they are about to act, one that is not experienced vividly at first. Clearly, conscious intentional states generally (as with our perceptual states) have phenomenal aspects that we cannot readily articulate and, importantly, our awareness of them comes in degrees, a point I shall now emphasize.

### 3 Philosophical and Scientific Models of the Will: Towards an Interface

To this point I have mostly thrown cold water on debunking-free-will claims that are based on the interesting and varied recent studies on the will. Nonetheless, I think that such findings do point to the need for fine-tuning of philosophical models of the will. Philosophers, especially those of us who are incompatibilists, are given to simple and idealized conceptions of human freedom. But the ever-accumulating empirical data and emerging partial theories present a messy picture that underscores the fragility of our freedom, and some elements of that picture cannot be readily mapped onto those idealized philosophical theories. One way in which philosophical models tend to over-idealize has already been remarked upon: given the pervasiveness of automaticity, the freedom and responsibility of much of what we do must be thought of as “inherited” from the comparatively few *directly free* choices that we make. While some recent philosophers have incorporated that fact into their thinking about the will, it is still not widely enough appreciated, so that philosophers often write as if we are constantly making explicit and considered choices.

In this final section, I want to suggest one other needed accommodation in our philosophical theorizing: a greater place for *conscious knowledge* in our account

of freedom of will. Making that accommodation leads pretty directly to a striking consequence: though freedom of the will requires a baseline capacity of choice, the nature of which is the subject of much traditional controversy I will not adjudicate here, the *freedom* this capacity makes possible is, nevertheless, a property that comes in *degrees* and can vary over time within an individual.<sup>1</sup>

Some of the studies summarized above highlight ways in which we can be significantly influenced in our decision-making by circumstantial factors without being aware that this is so. And it is a commonplace that our own motivations at times can be largely or entirely hidden to us. In such cases, I suggest, our freedom is diminished – not evaporated entirely, but diminished, as it is not possible for us to subject our unconscious motivations or guiding beliefs to critical scrutiny and decide what to do in a more reflective way.

But awareness of such factors comes in degrees, along more than one dimension. Here are three such dimensions, all of which seem freedom-relevant:

- degree of awareness of *y*, for each influencing factor *y* (desire, intention, belief, or circumstance);
- the relative “portion” of one’s total motivational structure (the totality of influencing factors) of which one is aware;
- the degree of likelihood that an unconscious influence *y* is a factor that one would reflectively endorse, were one to become aware of *y*’s influence. (It is one thing to be partly moved to act by a presently unconscious desire or intention that is well integrated into my character. It is another to be motivated by a factor – such as the level of ambient noise – whose relevance I would repudiate if asked.)<sup>2</sup>

There is at least one more degreed dimension to our motivations that I think bears importantly on our freedom. The medieval philosopher Robert Grosseteste in one place invites us to imagine God’s creating an angel that exists for a single instant only.<sup>3</sup> (Here I must ask my readers more at home in science than in such

---

<sup>1</sup> The philosophical account of freedom into which I myself wish to absorb these elements is an agent causal theory. Many philosophers think that the challenges of squaring free will with the emerging sciences of brain and behavior are more severe for the agent-causal view than for the other views, even other indeterminist views. But this is doubtful. All such views seem to require a strongly antireductionist assumption with respect to mental states and the capacity of choice vis-à-vis the complex physical states that subserve them. Once the other views’ robustly emergent ontological commitments are made explicit, the difference in strength of assumptions made by the models is diminished, as it concerns only what sort, not whether, a top-down causal factor must be introduced.

<sup>2</sup> A corollary to the freedom-relevance of these conditions is that one can increase one’s freedom by becoming more aware of typical unconscious influences on human decision-making. This point is noted by Nahmias (forthcoming), who also emphasizes the relevance of conscious knowledge to freedom of will.

<sup>3</sup> See his *On the Freedom of the Will* (translated into English by Lewis (1991). Grosseteste’s views influenced views influenced (via Henry of Ghent) the arch-champion of freedom of the will, John Duns Scotus.

philosophical flights of fancy to bear patiently with me.) In that instant, we are to imagine, the angel exercises his freedom by an instantaneous act of the will. In order to circumvent quite sensible worries about the coherence of Grosseteste's thought experiment, let us stretch out the angel's life to a few seconds. We will also suppose that the angel springs into existence with a fully developed psychology (of the typical angelic sort), complete with a bunch of pseudo-memories (false apparent memories) of a long, character-shaping history. And, as in Grosseteste's telling, he comes with a disposition to decide some matter straightaway. Finally, let us imagine that he is deciding between a plurality of alternatives, each having some attraction to him, is *fully* aware of his own motivations, and has the capacity to determine himself to do any of them. (Philosophers may insert here their favored account of this "self-determining" capacity.) In short, his will appears to be significantly free. Here is my question: in making his one and only choice, is our angel (whom we may call "Angelo") just as free as an intrinsically identical counterpart ("Angela") who at the time in question is intrinsically identical to Angelo, while differing dramatically *historically*: unlike Angelo, Angela really does have a history, one filled with many prior choices that have partly shaped her present inclinations and intentions? At the time at which they both chose, where they happen to coincide perfectly in their inclinations and capacities, are they equally *free*?

It seems not. The reason is that for Angela, unlike Angelo, the very factors that shape her choice were to some extent of her own making. Like Angelo, she began with a set of psychological and behavioral dispositions that were merely "given." But over time, as she habitually made certain choices, her psychological makeup reflected less and less this "givenness" and more and more something that is her own free creation.

And so, too, for ourselves. We come into the world with powerful tendencies that are refined by the particular circumstances in which we develop. All of these facts are for us merely "given." They determine which choices we have to make and which options we will consider (and how seriously) as we arrive at a more reflective age. Now, most of us are fortunate enough not to be impacted by traumatic events that will forever limit what is psychologically possible for us, and, on the positive side, are exposed to a suitably rich form of horizon-expanding opportunities. Where this is so, the framework of considerations that structures our choices increasingly reflects our own prior choices. And, in this way, our freedom grows over time.

For a further reason to think this is right, consider a scenario involving an agent much as we take ourselves to be – except that his psychology is regularly manipulated, altering some of his preferences and the strengths of others. Owing to the marvelous, wireless neural-intervention technology of the late twenty-first century, all this occurs while he remains wholly oblivious. Even if his capacity to choose remains robust, it seems clear that we must judge his freedom, his autonomy, to be diminished.<sup>4</sup> The integrity of the self-formation process is a component of freedom,

---

<sup>4</sup> Al Mele (1995), ch. 9, makes this point, and Clarke (2003, pp. 16n.4, 77) concurs. For a dissenting view, see Daniel Dennett (2003) pp. 281–87.

or of freedom of the most valuable sort. This conclusion is reflected in corresponding judgments about moral responsibility. If Angelo and Angela each contemplate a morally significant matter, are inclined to a degree to both a virtuous and a vicious action, and choose the virtuous one, Angela seems the more praiseworthy. The action is *hers* to a greater degree.

Thus, a fourth neglected dimension of freedom is historical: the degree to which motivation *y* is a product of the agent's own past free choices.<sup>5</sup> I suspect that there is a corresponding condition, harder to state, with respect to *beliefs* that also shape one's choices: I am freer to the extent to which I am not being influenced by beliefs that have been formed by defective mechanisms, or even by importantly false beliefs that have arisen through nonculpable misuse of normally functioning mechanisms. (The hallucination-generated beliefs in schizophrenics are an extreme case of what I have in mind here.)

It is an open empirical question the extent to which any given individual, or human beings in general, realize these conditions – just as it is an open question whether and when the basic capacity to choose, to which philosophers give most of their attention, is present and regularly exercised in a way necessary for true freedom of choice. And so empirical sciences of brain and behavior potentially have a lot to teach us about the *extent* and *scope* of human freedom, once we recognize the mistake of thinking that it is an all-or-nothing matter. In truth, human freedom is always limited, fragile, and variable over time and across agents. It is the sort of thing that comes in degrees – a fact that should inform not only our philosophical-cum-scientific theorizing, but also our moral understanding and assessment of one another.

## Acknowledgments

I have benefited greatly from discussion of ideas advanced in this chapter with several members of the “Top-Down Causation and Volition” symposium organized by the John Templeton Foundation in Yosemite National Park in April 2007, including Daniel Wegner, Cristoph Koch, Itzak Fried, Chris Frith, and Sean Spence. A version of this chapter was given as a plenary lecture at the 2007 meetings of the American Catholic Philosophical Association. I received helpful feedback and criticisms from Alex Pruss, Timothy Noone, and Robert Kane. In 2008, I read a revised version to a cognitive science seminar of the Institut Jean Nicod (École Normale Supérieure) in Paris; the 2008 German Congress of Philosophy in Essen; a 2008 conference at the University of Valencia, and colloquia at Denison and Purdue. I thank all the audiences on these occasions, especially the (insistently skeptical) Joëlle Proust, Max Kistler, Neil Roughley,

---

<sup>5</sup> As Neil Roughley pointed out to me, the specific effects of past decisions are relevant to their bearing on the freedom of future choices. We can imagine scenarios, for example, in which early choices cause one to fixate on certain goals, without one's having foreseen or intended this consequence, and this would clearly diminish rather than enhance one's future freedom. I'll not attempt here to develop an account of the historical dimension of freedom that captures this nuance.

Michael Bergmann, and Dana Nelkin. Some lines in the text were taken from my 2005 article and adapted for the present chapter.

## References

- Brasil-Neto, J.P., Pascual-Leone, A., Valls-Sole, J., Cohen, L.G., Hallett, M.: Focal transcranial magnetic stimulation and response bias in a forced choice task. *Journal of Neurology, Neurosurgery, and Psychiatry* 55, 964–966 (1992)
- Clarke, R.: *Libertarian accounts of free will*. Oxford University Press, New York (2003)
- Delgado, J.M.R.: *Physical control of the mind: Toward a psychocivilized society*. Harper & Row, New York (1969)
- Dennett, D.: *Freedom evolves*. Viking, New York (2003)
- Doris, J.: *Lack of character: Personality and moral behavior*. Cambridge University Press, Cambridge (2002)
- Doris, J., Stich, S. (2006), *Moral psychology: Empirical approaches*. In: *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/moral-psych-emp/>
- Fischer, J.M.: *The metaphysics of free will*. Blackwell, Oxford (1994)
- Frankfurt, H.: Alternative possibilities and moral responsibility. *Journal of Philosophy* 66, 829–839 (1969)
- Gazzaniga, M.S.: Consciousness and the cerebral hemispheres. In: Gazzaniga, M.S. (ed.) *The cognitive neurosciences*, pp. 1391–1399. MIT Press, Cambridge (1994)
- Ginet, C.: *On action*. Cambridge University Press, Cambridge (1990)
- Haggard, P.: Conscious intention and motor cognition. *Trends in Cognitive Sciences* 9, 290–295 (2005)
- Haggard, P., Clark, S., Kalogeras, J.: Voluntary action and conscious awareness. *Nature Neuroscience* 5, 382–385 (2002)
- James, W.: *The principles of psychology*, vol. 2. Henry Holt & Co., New York (1890)
- Lau, H.C., Rogers, R.D., Ramnani, N., Passingham, R.E.: Willed action and attention to the selection of action. *NeuroImage* 21, 1407–1415 (2004)
- Lau, H.C., Rogers, R.D., Passingham, R.E.: On measuring the perceived onsets of spontaneous actions. *Journal of Neuroscience* 26, 7265–7271 (2006)
- Lau, H.C., Rogers, R.D., Passingham, R.E.: Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience* 19, 81–90 (2007)
- Libet, B.: Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences* 8, 529–566 (1985)
- Mele, A.: *Autonomous agents*. Oxford University Press, New York (1995)
- Mele, A.: Strength of motivation and being in control: Learning from Libet. *American Philosophical Quarterly* 34, 319–332 (1997)
- Nahmias, E.: Autonomous agency and social psychology. In: Marraffa, M., Caro, M., Ferretti, F. (eds.) *Cartographies of the mind: Philosophy and psychology in intersection*, pp. 169–185. Springer, Dordrecht (2007)
- Nahmias, E.: The psychology of free will. In: Prinz, J. (ed.) *Oxford handbook on philosophy of psychology*. Oxford University Press, Oxford (forthcoming)
- O'Connor, T.: *Persons and causes: The metaphysics of free will*. Oxford University Press, New York (2000)
- O'Connor, T.: Freedom with a human face. *Midwest Studies in Philosophy* 29, 207–227 (2005)
- Wegner, D.: *The illusion of conscious will*. MIT Press, Cambridge (2002)
- Wegner, D., Wheatley, T.: Apparent mental causation: Sources of the experience of will. *American Psychologist* 54, 480–491 (1999)