FREE WILL: ACTION THEORY MEETS NEUROSCIENCE

Alfred R. Mele, Florida State University

It is a truism that whether or not anyone has free will depends on what free will is. In section 1, I describe two competing theories about the nature of free will. In section 2, I sketch some conceptual background relevant to some striking claims Benjamin Libet has made about decisions and free will. The experiments on which Libet bases these claims are the topic of sections 3 and 4. I argue there that his data do not justify his claim that "the brain 'decides' to initiate or, at least, to prepare to initiate the act before there is any reportable subjective awareness that such a decision has taken place" (Libet 1985, p. 536) and do not justify associated worries about free will. I conclude, in section 5, with some remarks about the relevance of another experiment.

1. Free Will: Conceptual and Theoretical Background

Free will may be defined as the power to act freely. But what is it to act freely? Familiar philosophical answers fall into two groups: compatibilist and incompatibilist. Compatibilism and incompatibilism are theses about the relationship between free action and determinism. Determinism is the thesis that a complete statement of the laws of nature together with a complete description of the condition of the entire universe at any point in time logically entails a complete description of the condition of the entire universe at any other point in time. Compatibilism is the thesis that free action is compatible with the truth of determinism. Because they attend to what contemporary physics tells us, the overwhelming majority of contemporary compatibilists do not believe that determinism is true, but they do believe that even if it were

true, people would be able to act freely. Incompatibilism is the thesis that free action is incompatible with the truth of determinism. In the incompatibilist group, most answers to the question what it is to act freely come from libertarians. Libertarianism is the conjunction of incompatibilism and the thesis that some people sometimes act freely. Some incompatibilists argue that no one acts freely (Pereboom 2001). They argue that even the falsity of determinism creates no place for free action.

The compatibilist thesis usually sounds strange to nonspecialists. When people first encounter the pair of expressions "free will" and "determinism" they tend to get the impression that the two ideas are defined in opposition to each other, that they are mutually exclusive by definition. This is one reason that it is useful to think of free will as the power to act freely and to regard acting freely as the more basic notion – that is, as a notion in terms of which free will is to be defined. Consider the following conversation between two Detroit police officers who have a notoriously stingy friend named Stan. Ann: "Stan gave $20 to a homeless man today." Bill: "Why? Did he hold a gun to Stan's head?" Ann: "No, Stan gave him the money freely." Surely, Ann and Bill do not need to have an opinion about whether determinism (as defined above) is true to have this conversation. If what Ann says is true – that is, if Stan freely gave away $20 – and free will is the power to act freely, then Stan has free of will (or had it at that time). Even if "free will" is typically opposed to "determinism" in ordinary speech, "he did it freely" seems not to be. And even if "he did it freely" were typically opposed to determinism in ordinary speech, that would settle nothing. After all, in ordinary speech, deductive reasoning seems to be defined as reasoning from the general to the particular, and that certainly would only jokingly be said to constitute an objection to a logician's definition of deduction (according to

which “Ann is a police officer; Bill is a police officer; therefore Ann and Bill are police officers” is a valid deductive argument).

Compatibilist theories of free action emphasize a distinction between deterministic causation and compulsion (Frankfurt 1988, Smith 2003). If determinism is true, then my eating toast for breakfast today and my working on this article today were deterministically caused; and so were a certain compulsive hand-washer’s washing his hands dozens of times today, a certain delusional person’s spending the day trying to contact God with his microwave, a certain addict’s using his favorite drug while in the grip of an irresistible urge to do so, and a certain person’s handing over money to gunmen who convincingly threatened to kill him if he refused. But there is an apparent difference. I am sane and free from addiction, and I received no death threats today. The basic compatibilist idea is (roughly) that when mentally healthy people act intentionally in the absence of compulsion and coercion they act freely, and an action’s being deterministically caused does not suffice for its being compelled or coerced.

Many compatibilists have been concerned to accommodate the idea that, for example, if I freely spent the day working, I could have done something else instead. They grant that, if determinism is true, then there is a sense in which people could never have done otherwise than they did: they could not have done otherwise in the sense that their doing otherwise is inconsistent with the combination of the past and the laws of nature. But, these compatibilists say, the fact that a person never could have done otherwise in that sense is irrelevant to free action. What is relevant is that people who act freely are exercising a rational capacity of such a kind that if their situation had been different in any one of a variety of important ways, they would have responded to the difference with a different suitable action (Smith 2003). For example, although I spent the day working, I would have spent the day relaxing if someone had

bet me $500 that I would not relax all day. This truth is consistent with determinism. (Notice that if someone had made this bet with me, the past would have been different from what it actually was.) And it reinforces the distinction between deterministic causation and compulsion. Offer a compulsive hand-washer $500 not to wash his hands all day and see what happens.

Like compatibilists, libertarians tend to maintain that when mentally healthy people act intentionally in the absence of compulsion and coercion they act freely, but they insist that the deterministic causation of an action is incompatible with the action's being freely performed. Some libertarian theories of free action assert that agents never act freely unless some of their actions are indeterministically caused by immediate antecedents (Kane 1996). Whereas the laws of nature that apply to deterministic causation are exceptionless, those that apply most directly to indeterministic causation are instead probabilistic.[1] Typically, events like deciding to help a stranded motorist – as distinct from the physical actions involved in actually helping – are counted as mental actions. Suppose that Ann's decision to help a stranded motorist is indeterministically caused by, among other things, her thinking that she should help. Because the causation is indeterministic, she might not have decided to help given exactly the same internal and external conditions. In this way, some libertarians seek to secure the possibility of doing otherwise that they require for free action.

There are a variety of kinds of libertarian and compatibilist theories about free will. Each kind has been challenged, of course. Reviewing the major details is a project for a separate article. My brief remarks in this section provide enough theoretical background on free will to set the stage for what follows.

2. Some Conceptual Distinctions

A 1983 article by physiologist Benjamin Libet and colleagues (Libet, Gleason et al 1983) has been described as "one of the most philosophically challenging papers in modern scientific psychology" (Haggard et al 1999, p. 291). A striking thesis of that 1983 article is that "the brain . . . 'decides' to initiate or, at the least, prepare to initiate [certain actions] at a time before there is any reportable subjective awareness that such a decision has taken place" (p. 640; aslo see Libet 1985, p. 536).[2] In a recent article, Libet pointedly asserts: "If the 'act now' process is initiated unconsciously, then conscious free will is not doing it" (2001, p. 62; also see 2004, p. 136).

Because Libet uses such terms as "decision," "intention," "wanting," "wish," and "urge" interchangeably, some conceptual preliminaries are in order in interpreting his work.[3] Most people recognize that deciding to do something differs from having an urge or wanting to do something. For example, you might have an urge to scream at an irritating colleague but decide not to. And you might want to have a second helping of dessert but decide to stop at one. If, as I believe, to decide to A is to perform a momentary mental action of forming an intention to A (Mele 2003, ch. 9), then in deciding to stop at one dessert you form an intention to stop at one. Your having that intention also differs from your merely wanting to stop at one dessert. After all, you might want to have another dessert (it is very tempting) while also wanting to refrain from having it (you are concerned about your weight); but intending to have a second helping of dessert while intending to refrain from doing that, if it is possible at all, would be a sign of a serious disorder. Similarly, you might want to meet one friend at a 7:00 movie tonight, want to meet another at a 7:00 play, and be unsettled about what to do. At this point, you want to do each of these things and lack an intention about which of them to do.

In saying that deciding is momentary, I mean to distinguish it from, for example, a combination of deliberating and deciding. Someone who is speaking loosely may say, "I was up all night deciding to sell my house" when what he means is that he was up all night deliberating or fretting about whether to sell his house and eventually decided to sell it. Deciding to A, on my view, is not a process but a momentary mental action of forming an intention to A, 'form' being understood as an action verb.

Not all intentions are formed in acts of deciding. For example, "When I intentionally unlocked my office door this morning, I intended to unlock it. But since I am in the habit of unlocking my door in the morning and conditions . . . were normal, nothing called for a decision to unlock it" (Mele 1992, p. 231). If I had heard a fight in my office, I might have paused to consider whether to unlock the door or walk away, and I might have decided to unlock it. But given the routine nature of my conduct, there is no need to posit an act of intention formation in this case. My intention to unlock the door may have been acquired without having been actively formed. If, as I believe, all decisions are prompted partly by uncertainty about what to do (Mele 2003, ch. 9), in situations in which there is no such uncertainty, no decisions will be made.

Some of our decisions and intentions are for the nonimmediate future and others are not. I might decide on Monday to attend a lecture on Friday, and I might decide now to phone my sister now. The intention formed in the former decision is aimed at action four days in the future. The intention I form when I decide to phone my sister now is about what to do now. I call intentions and decisions of these kinds, respectively, distal and proximal intentions and decisions (Mele 1992, pp. 143-44, 158). Proximal decisions and intentions also include decisions and intentions to continue doing something that one is doing and decisions and intentions to start A-ing (e.g., start climbing a hill) straightaway.

A distinction between relatively specific and relatively unspecific intentions also is in order. Ann now intends to buy a particular Ford Escort that she saw yesterday. That is a more specific intention than the intention she had three months ago, in January, to buy a new car this year. In another illustration, Bob has agreed to be a subject in an experiment in which subjects are instructed to salute whenever they feel like it on at least forty occasions during a two hour period. When Bob begins his participation in the experiment he has a relatively unspecific intention to salute many times during the next two hours. At various times during the experiment he has specific proximal intentions to salute.

3. Libet's Work

This section develops an interpretation of Libet's work that is sensitive to the conceptual points just made. In some of his studies, subjects are instructed to flex their right wrists or the fingers of their right hands whenever they wish. Electrical readings from the scalp – averaged over at least 40 flexings for each subject – show a "negative shift" in "readiness potentials" (RPs) beginning at about 550 milliseconds (ms) before the time at which an electromyogram shows relevant muscular motion to begin (1985, pp. 529-30).[4] Subjects are also instructed to "recall . . . the spatial clock position of a revolving spot at the time of [their] initial awareness" (p. 529) of something, x, that Libet variously describes as an "intention," "urge," "wanting," "decision," "will," or "wish" to move (see n. 3). On average, "RP onset" preceded what subjects reported to be the time of their initial awareness of x (time W) by 350 ms. Time W, then, preceded the beginning of muscle motion by about 200 ms.

Figure 1

-550 ms	-200 ms	0 ms
RP onset	time W	muscle begins to move

(Libet finds independent evidence of a slight error in subjects' recall of the times at which they first become aware of sensations [pp. 531, 534]. Correcting for that error, time W is -150 ms.)

At what point, if any, does a specific intention to flex arise in Libet's subjects? Again, Libet writes: "the brain 'decides' to initiate or, at least, to prepare to initiate the act before there is any reportable subjective awareness that such a decision has taken place" (1985, p. 536). If we ignore the second disjunct, this quotation (given its context) apparently offers the answer that a specific intention to flex appears on the scene with "RP onset," about 550 ms before relevant muscular motion and about 350 to 400 ms before the agent becomes aware of the intention (see Libet 1985, p. 539); for to decide to initiate an act is to form an intention to initiate it.[5] But are decision and intention the most suitable mental items to associate with RP onset? Again, Libet describes the relevant occurrence of which the agent later becomes aware not only as a "decision" and the onset of an "intention" to move, but also as the onset of an "urge," "wanting," and a "wish" to move. This leaves it open that at -550 ms, rather than acquiring an intention or making a decision of which he is not conscious, the person instead acquires an urge or desire of which he is not conscious – and perhaps an urge or desire that is stronger than any competing urge or desire at the time, a preponderant urge or desire. It is also left open that what emerges around -550 ms is a pretty reliable causal contributor to an urge.

I believe that if Libet himself were to distinguish between intending and wanting (including having an urge) along the lines I sketched, he might find it more credible to associate

the readiness potentials with the latter than with the former. To explain why, I turn to another experiment reported in Libet 1985 (and elsewhere).

Libet proposes that "conscious volitional control may operate not to initiate the volitional process but to select and control it, either by permitting or triggering the final motor outcome of the unconsciously initiated process or by vetoing the progression to actual motor activation" (1985, p. 529; also see 1999, p. 54, 2004, pp. 139, 142-43, 149). "In a veto, the later phase of cerebral motor processing would be blocked, so that actual activation of the motoneurons to the muscles would not occur" (1985, p. 537). Libet offers two kinds of evidence to support the suggestion about vetoing. One kind is generated by an experiment in which subjects are instructed both to prepare to flex their fingers at a prearranged time (as indicated by a revolving spot on a clock face) and "to veto the developing intention/preparation to act . . . about 100 to 200 ms before the prearranged clock time" (p. 538). Subjects receive both instructions at the same time. Libet writes:

> a ramplike pre-event potential was still recorded . . . resembl[ing] the RP of self-initiated acts when preplanning is present. . . . The form of the 'veto' RP differed (in most but not all cases) from those 'preset' RPs that were followed by actual movements [in another experiment]; the main negative potential tended to alter in direction (flattening or reversing) at about 150-250 ms before the preset time. . . . This difference suggests that the conscious veto interfered with the final development of RP processes leading to action. . . . The preparatory cerebral processes associated with an RP can and do develop even when intended motor action is vetoed at approximately the time that conscious intention would normally appear before a voluntary act. (1985, p. 538)[6]

Keep in mind that the subjects were instructed in advance not to flex their fingers, but to prepare to flex them at the prearranged time and to "veto" this. The subjects intentionally complied with the request. They intended from the beginning not to flex their fingers at the appointed time. So what is indicated by the RP? Presumably, not the acquisition or presence of an intention to flex; for then, at some point in time, the subjects would have both an intention to flex at the prearranged time and an intention not to flex at that time. And how can a normal agent simultaneously be settled on A-ing at t and settled on not A-ing at t?[7] That is, it is very plausible that Libet is mistaken in describing what is vetoed as "intended motor action" (p. 538; my emphasis).

If the RP in the veto scenario is not associated with an intention to flex at the appointed time, with what might it be associated? In the passage I quoted from Libet 1985, p. 538, Libet compares "the 'veto' RP" with (a) "'preset' RPs that were followed by actual movements" and (b) "the RP of self-initiated acts when preplanning is present." The RP referred to in a is produced in experiments in which subjects are instructed to watch the clock and flex when the revolving spot reaches "a pre-set 'clock time'" (Libet et al 1982, p. 325). "The subject was encouraged to try to make his movement coincide as closely as possible with the arrival of the spot at the pre-set time." The RP referred to in b is produced in two kinds of studies: studies in which subjects are neither instructed to flex at a prearranged time nor explicitly told to flex spontaneously (Libet et al 1982, p. 326), and studies in which subjects instructed to flex spontaneously reported that they experienced "some 'pre-planning'," even if only in "a minority of the 40 self-initiated acts that occurred in the series for that averaged RP" (p. 328). "Even when some pre-plannings were recalled and reported, subjects insisted that the more specific

urge or intention to actually move did not arise in that pre-planning stage" (p. 329). Reports of "pre-planning" seem to include reports of thoughts about when to flex and reports of anticipations of flexing (pp. 328-29). Libet and his coauthors remark that "Subject S.B. described his advance feelings [of pre-planning] as 'pre-tensions' rather than pre-plannings to act" (p. 329). This subject may have meant that he experienced tension that he expected to result in flexing.

The RPs referred to in a and b have a very similar form (Libet et al 1982, pp. 330, 333-34; Libet 1985, p. 532). RPs with that form are called "type I RPs" (p. 326). They have significantly earlier onsets than the RPs produced in studies in which subjects instructed to flex spontaneously report that they experienced no "pre-planning" – "type II RPs." "The form of the 'veto' RP" is the form of type I RPs until "about 150-250 ms before the preset time" (1985, p. 538). What does the veto group (group V) have in common until that time with the three kinds of subjects who produce type I RPs: those with a pre-set time for flexing (group PS), those who are neither told to flex at a pre-set time nor instructed to flex spontaneously (group N), and those who are instructed to flex spontaneously but who report some "pre-planning" (group PP)?

Presumably, subjects in group PS are watching the clock with the intention of flexing at the pre-set time. But it certainly does not follow from that and the similar RPs in groups N and PP – and V for a time – that members of each of these groups are watching the clock with a similar intention to flex. For one thing, as I have explained, it is very likely that group V – subjects instructed in advance to prepare to flex and then veto the preparation – are watching the clock without an intention to flex at the targeted time. Given that the members of group V lack this intention, we should look for something that groups V and PS actually have in common that might be signified by the similarity in the RPs until "about 150-250 ms before the preset time."

One possibility is that members of both groups have urges to flex (or to prepare to flex) – or undergo brain events that are pretty reliable relatively proximal causal contributors to such urges – that are associated with an RP and regularly play a role in generating subsequent flexings in the absence of "vetoing."[8] In the case of group V, perhaps a subject's wanting to comply with the instructions – including the instruction to prepare to flex at the appointed time – together with his recognition that the time is approaching produces a growing urge to (prepare to) flex, or a pretty reliable causal contributor to such an urge, or a simulation of such an urge, or the motor preparedness typically associated with such an urge. And the "flattening or reversing" of the RP "at about 150-250 ms before the preset time" might indicate a consequence of the subject's "vetoing" his preparation.

What about groups N and PP? It is possible that they, along with the subjects in groups PS and V, begin acquiring urges to flex at a greater temporal distance from 0 ms than do subjects instructed to flex spontaneously who report no pre-planning. That difference may be indicated by type I RPs' having significantly earlier onsets than type II RPs. Another possibility is consistent with this. Earlier, I distinguished proximal from distal intentions, and Libet himself recognizes the distinction (see Libet et al 1982, pp. 329, 334; Libet 1989, pp. 183-84). Presumably, subjects in group PS respond to the instruction to flex at a pre-set time with an intention to flex at that time. This is a distal intention. As the pre-set time for flexing draws very near, that intention may become, help produce, or be replaced by a proximal intention to flex, an intention to flex now, as one naturally says (see Libet 1989, p. 183; 1999, p. 54; 2004, p. 148). That may happen around the time subjects in group V veto their urge to flex or closer to 0 ms. And it may happen at or around the time subjects in groups N and PP acquire a proximal intention to flex. They may acquire such an intention without having had a distal intention to

flex soon: recall that members of group V probably had no distal intention to flex soon and that their RPs are very similar to those of groups N, PP, and PS until "about 150-250 ms before the preset time." All this is consistent with the similarities in RPs in the various groups of subjects, on the assumption that no segment of the RPs before about -150 to -250 ms for subjects in group PS specifically represents subjects' distal intentions to flex at the pre-set time – as opposed, for example, to something that such intentions have in common with distal urges to flex (or to prepare to flex) at the pre-set time – even though those intentions are present.

The main difference between type I and type II RPs, in Patrick Haggard's words, is that the former have "earlier onsets than" the latter (Haggard and Libet 2001, p. 49). The earlier onsets may be correlated with earlier acquisitions of urges to flex soon – urges that may be brought on, variously, by the instruction to flex at a pre-set time (group PS), the instruction to prepare to flex at a pre-set time and to veto that later (group V), unsolicited conscious thoughts about when to flex (groups N and PP), or unsolicited conscious anticipations of flexing (groups N and PP). (Of course, it is possible that some such thoughts and anticipations are instead products, in part, of urges to flex soon.) These urge inciters (or perhaps urge products, in the case of some experiences in groups N and PP) are absent in subjects instructed to flex spontaneously who report no "pre-planning" – at least, if their reports are accurate. If type I RPs indicate urges, or urges together with proximal intentions that emerge later than the urges do, the same may be true of type II RPs. The difference in the two kinds of RP may mainly be a matter of when the urge emerges – that is, how long before 0 ms. Once again, Libet describes in a variety of ways the mental item that is indicated by RPs. Even if "intention" and "decision" (to flex) are not apt choices, "urge" and "wanting" are still in the running.

If "RP onset" in cases of "spontaneous" flexing indicates the emergence of an urge to flex soon, proximal intentions to flex may emerge at some point between RP onset and time W, at time W, or after time W: at time W the agent may be aware only of an urge that has not yet issued in a proximal intention. Again, Libet asserts that "In a veto, the later phase of cerebral motor processing would be blocked, so that actual activation of the motoneurons to the muscles would not occur" (1985, p. 537). Perhaps, in non-veto cases, activation of these motoneurons is the direct result of the acquisition of a proximal intention (Gomes 1999, pp. 68, 72; Mele 1997, pp. 322-24). Libet suggests that this activation event occurs between 10 and 90 ms before the muscle begins moving and apparently favors an answer in the 10 to 50 ms range (p. 537). Elsewhere, he asserts that the activation event can occur no later than 50 ms before the onset of muscle motion (2004, pp. 137-38).

Although I will not make much of the following point, it should be observed that urges that may be correlated with RP onset at -550 ms might not be proximal urges, strictly speaking. Possibly, they are urges to flex very soon, as opposed to urges to flex straightaway. And perhaps they evolve into, or produce, proximal urges. Another possibility is that urges to flex very soon give rise to proximal intentions to flex without first evolving into or producing proximal urges to flex. Some disambiguation is in order. A smoker who is rushing toward a smoking section in an airport with the intention of lighting up as soon as he enters it wants to smoke soon. That want or desire has a specific temporal target – the time at which he enters the smoking section. A smoker walking outside the airport may want to smoke soon without having a specific time in mind. Libet's subjects, like the latter smoker, might at times have urges or desires to flex that lack a specific temporal target. Desires to A very soon, or to A, beginning very soon, in this sense of "very soon," are roughly proximal action-desires.

Libet's experimental design promotes consciousness of urges and intentions to flex, since his subjects are instructed in advance to be prepared to report on them – or something like them – later, using the clock to pinpoint the time they are first noticed. For my purposes, what is of special interest are the relative times of the emergence of a (roughly) proximal urge or desire to flex, the emergence of a proximal intention to flex, and consciousness of the intention. If RP onset indicates the emergence of proximal, or roughly proximal, urges to flex, and if acquisitions of corresponding proximal intentions directly activate the motoneurons to the relevant muscles, we have the following picture of subjects instructed to flex "spontaneously" who report no "pre-planning" – subjects who produce type II RPs:

Figure 2

a. -550 ms: proximal or roughly proximal urge to flex emerges

b. -90 to -50 ms: acquisition of corresponding proximal intention[9]

c. 0 ms: muscle begins to move.

Possibly, the intention is consciously acquired. My point here is simply that this figure is consistent with Libet's data on type II RPs and on time W.

In an alternative picture, the acquisition of a proximal intention to flex sends a signal that may be regarded as a command to flex one's wrist (or finger), and that signal helps produce finer-grained command signals that directly activate the motoneurons to the relevant muscles. This picture moves the time of the acquisition of a proximal intention further from 0 ms than -90 to -50 ms, but it does not move it anywhere near -550 ms. On this, see section 5 below.

I mentioned that Libet offered a second kind of evidence for "veto control." Subjects instructed to flex "spontaneously" (in non-veto experiments) "reported that during some of the trials a recallable conscious urge to act appeared but was 'aborted' or somehow suppressed before any actual movement occurred; in such cases the subject simply waited for another urge to appear, which, when consummated, constituted the actual event whose RP was recorded" (1985, p. 538). RPs were not recorded for suppressed urges. But if these urges fit the pattern of the unsuppressed ones in cases of "spontaneous" flexing, they appeared on the scene about 550 ms before the relevant muscles would have moved if the subjects had not "suppressed" the urges, and subjects did not become conscious of them for about another 350 to 400 ms. Notice that it is urges that these subjects are said to report and abort or suppress. This coheres with my "urge" hypothesis about groups V, PS, N, and PP. In group V (the veto group), as I have explained, there is excellent reason to believe that no proximal intention to flex is present, and the RPs for this group resembled the type I RPs for these other three groups until "about 150-250 ms before the preset time." If it is assumed that these RPs represent the same thing for these four groups until the RPs for group V diverge from the others, these RPs do not represent a proximal intention to flex before the point of divergence, but they might represent a growing urge to (prepare to) flex, or a pretty reliable relatively proximal causal contributor to such an urge, or the motor preparedness typically associated with such an urge. And if at least until about the time of divergence there is no proximal intention to flex in any of these groups, we would need a special reason to believe that the type II RPs of the spontaneous flexers indicate that proximal intentions to flex emerge in them around -550 ms. In section 5, I show that there is independent evidence that their proximal intentions emerge much later.

Does the brain decide to initiate actions "at a time before there is any reportable subjective awareness that such a decision has taken place" (Libet, Gleason et al 1983, p. 640)? Libet and his colleagues certainly have not shown that it does, for their data do not show that any such decision has been made before time W or before the time at which their subjects first are aware of a decision or intention to flex. Nothing justifies the claim that what a subject becomes aware of at time W is a decision to flex that has already been made or an intention to flex that has already been acquired, as opposed, for example, to an urge to flex that has already arisen. Indeed, the data about vetoing, as I have explained, can reasonably be used to argue that the "urge" hypothesis about what the RPs indicate is less implausible than the "decision" or "intention" hypothesis. Now, there certainly seems to be a connection between what happens at -550 ms and subsequent muscle motion in cases of "spontaneous" flexing. But it obviously is not a temporally direct connection. Between the former and latter times, subjects apparently form or acquire proximal intentions to flex, in those cases in which they do intentionally flex. And, for all Libet's data show, those intentions may be consciously formed or acquired.

4. Free Will

When Libet's work is applied to the theoretically subtle and complicated issue of free will, things can quickly get out of hand. The abstract of Haggard and Libet 2001 opens as follows: "The problem of free will lies at the heart of modern scientific studies of consciousness. An influential series of experiments by Libet has suggested that conscious intentions arise as a result of brain activity. This contrasts with traditional concepts of free will, in which the mind controls the body" (p. 47). Now, only a certain kind of mind-body dualist would hold that conscious intentions do not "arise as a result of brain activity." And such dualist views are rarely

advocated in contemporary philosophical publications on free will. Moreover, contemporary philosophers who argue for the existence of free will typically shun substance dualism. If Libet's work is of general interest to philosophers working on free will, the source of the interest must lie elsewhere than the theoretical location specified in this passage.

In a recent article, Libet writes: "it is only the final 'act now' process that produces the voluntary act. That 'act now' process begins in the brain about 550 msec before the act, and it begins unconsciously" (2001, p. 61).[10] "There is," he says, "an unconscious gap of about 400 msec between the onset of the cerebral process and when the person becomes consciously aware of his/her decision or wish or intention to act." (Incidentally, a page later, he identifies what the agent becomes aware of as "the intention/wish/urge to act" [p. 62].) Libet adds: "If the 'act now' process is initiated unconsciously, then conscious free will is not doing it."

I have already explained that Libet has not shown that a decision to flex is made or an intention to flex acquired at -550 ms. But even if the intention emerges much later, that is compatible with an "act now" process having begun at -550 ms. One might say that "the 'act now' process" in Libet's spontaneous subjects begins with the formation or acquisition of a proximal intention to flex, much closer to the onset of muscle motion than -550 ms, or that it begins earlier, with the beginning of a process that issues in the intention.[11] We can be flexible about that (just as we can be flexible about whether the process of my baking my frozen pizza began when I turned my oven on to pre-heat it, when I opened the oven door five minutes later to put the pizza in, when I placed the pizza on the center rack, or at some other time). Suppose we say that "the 'act now' process" begins with the unconscious emergence of an urge to flex – or with a pretty reliable relatively proximal causal contributor to urges to flex – at about -550 ms and that the urge plays a significant role in producing a proximal intention to flex many

milliseconds later. We can then agree with Libet that, given that the "process is initiated unconsciously, . . . conscious free will is not doing it" – that is, is not initiating "the 'act now' process." But who would have thought that conscious free will has the job of producing urges? In the philosophical literature, free will's primary locus of operation is typically identified as deciding (or choosing); and for all Libet has shown, if his subjects decide (or choose) to flex "now," they do so consciously.

Libet asks (2001, p. 62), "How would the 'conscious self' initiate a voluntary act if, factually, the process to 'act now' is initiated unconsciously?" Here is one answer. An "'act now' process" that is initiated unconsciously may be aborted by the agent; that apparently is what happens in instances of spontaneous vetoing, if "'act now' processes" start when Libet says they do.[12] Now, processes have parts, and the various parts of a process may have more and less proximal initiators. A process that is initiated by the welling up of an unconscious urge may have a subsequent part that is directly initiated by the conscious formation or acquisition of an intention. "The 'conscious self'" – which need not be understood as something mysterious – might more proximally initiate a voluntary act that is less proximally initiated by an unconscious urge. (Readers who, like me, prefer to use 'self' only as an affix may prefer to say that the acquisition or formation of a relevant proximal intention, which intention is consciously acquired or formed, might more proximally initiate an intentional action that is less proximally initiated by an unconscious urge.)

Recall that Libet himself says that "conscious volitional control may operate . . . to select and control ['the volitional process'], either by permitting or triggering the final motor outcome of the unconsciously initiated process or by vetoing the progression to actual motor activation" (1985, p. 529; also see 1999, p. 54). "Triggering" is a kind of initiating. In "triggering the final

motor outcome," the acquisition of a proximal intention would be initiating an action in a more direct way than does an urge that initiated the process that issued in the intention. According to one view of things, when proximal action-desires help to initiate overt actions they do so by helping to produce pertinent proximal intentions the formation or acquisition of which directly initiates actions.[13] What Libet says about triggering here coheres with this.

Nothing warrants Libet's claim that type II RPs are correlated with decisions or intentions rather than with urges strong enough to issue pretty regularly in related intentions and actions. Moreover, that, in certain settings, (roughly) proximal urges to do things arise unconsciously – urges on which the agent may or may not act about half a second after they arise – is no cause for worry about free will. Even if one grants Libet much more than his critics tend to grant him, as I have done, it can be shown that his data fall well short of providing good grounds for accepting his main theses.[14]

5. Further Testing

I have argued that the "urge" hypothesis about what the type II RPs indicate in Libet's studies is less implausible than the "decision" or "intention" hypothesis. Is there an independent way to test these hypotheses – that is, to gather evidence about whether it is (roughly) proximal urges that emerge around -550 ms in Libet's studies or instead decisions or intentions?[15] One line of thought runs as follows: (1) all overt intentional actions are caused by decisions (or intentions); (2) the type II RPs, which emerge around -550 ms, are correlated with causes of the flexing actions (because they regularly precede the onset of muscle motion); so (3) these RPs indicate that decisions are made (or intentions acquired) at -550 ms. I have shown that this line of thought is unpersuasive. A lot can happen in a causal process that runs for 550 ms, including a

subject's moving from having an unconscious roughly proximal urge to flex to consciously deciding to flex "now" or to consciously acquiring a proximal intention to flex. One can reply that, even so, 3 might be true. And, of course, I can run through my argumentation about the veto and related matters again to remind the imaginary respondent why 3 is improbable. But what about a test?

If makings of proximal decisions to flex or acquisitions of proximal intentions to flex (or the physical events that realize these things) cause muscle motion, how long does it take them to do that? Does it take about 550 ms? Might reaction time experiments show that 550 ms is too long a time for this? Some caution is in order here. Lüder Deeke has distinguished among three kinds of decision: decisions about what to do, decisions about how to do something, and decisions about when to do something (1996, pp. 59-60). In reaction time experiments, subjects have decided in advance to perform the assigned task – to "A," for short – whenever they perceive the relevant signal. When they perceive the signal, there is no need for a proximal decision to A, as Deeke observes (p. 59). (If all decisions are responses to uncertainty about what to do and subjects are not uncertain about what to do when they perceive the signal, there is no place here for proximal decisions to A.) However, it is plausible that when they perceive the signal, they acquire an intention to A now, a proximal intention. That is, it is plausible that the combination of their conditional intention to A whenever they perceive the signal (or the neural realizer of that intention) and their perception of the signal (or the neural realizer of that perception) produces a proximal intention to A. The acquisition of this intention (or the neural realization of that event) would then initiate the A-ing.[16] And in at least one reaction time experiment (described shortly) that is very similar to Libet's main experiments, the time between the "go" signal and the onset of muscle motion is much shorter than 550 ms. This is evidence

that proximal intentions to flex – as opposed to (roughly) proximal urges to flex – emerge much closer to the time of the onset of muscle motion than 550 ms. There is no reason, in principle, that it should take people any longer to start flexing their wrists when executing a proximal intention to flex in Libet's studies than it takes them to do this when executing such an intention in a reaction time study.

The line of reasoning that I have just sketched depends on the assumption that, in reaction time studies, a proximal intention to A is at work. An alternative possibility is that the combination of subjects' conditional intentions to A whenever they perceive the signal and their perception of the signal initiates the A-ing without there being any proximal intention to A. Of course, there is a parallel possibility in the case of Libet's subjects. Perhaps the combination of their conditional intentions to A whenever they feel like it – conscious intentions, presumably – together with a relevant feeling (urge) initiates a flexing without there being any proximal intention to flex. If that possibility is an actuality, then Libet's thesis is false, of course: there is no intention to flex "now" in his subjects and, therefore, no such intention is produced by the brain before the mind is aware of it.

The reaction time study I mentioned is reported in Haggard and Magno 1999:

> Subjects sat at a computer watching a clock hand . . . whose rotation period was 2.56 s. . . . After an unpredictable delay, varying from 2.56 to 8 s, a high-frequency tone . . . was played over a loudspeaker. This served as a warning stimulus for the subsequent reaction. 900 ms after the warning stimulus onset, a second tone . . . was played. [It] served as the go signal. Subjects were instructed to respond as rapidly as possible to the go signal with a right-key press on a computer mouse button. Subjects were instructed

> not to anticipate the go stimulus and were reprimanded if they responded on catch trials. (p. 103)

"Reaction times were calculated by examining the EMG signal for the onset of the first sustained burst of muscle activity occurring after the go signal" (p. 104). "Reaction time" here, then, starts before any intention to press "now" is acquired: obviously, it takes some time to detect the signal, and if detection of the signal helps to produce a proximal intention, that takes some time too. The mean of the subjects' median reaction times in the control trials was 231 ms (p. 104). If a proximal intention to press was acquired, that happened nearer to the time of muscle motion than 231 ms and, therefore, much nearer than the 550 ms that Libet claims is the time proximal intentions to flex are unconsciously acquired in his studies. Notice also how close we are getting to Libet's time W, his subjects' reported time of their initial awareness of something he variously describes as an "intention," "urge," "wanting," "decision," "will," or "wish" to move (-200 to -150 ms). If proximal intentions to flex are acquired in Libet's studies, Haggard and Magno's results make it look like a good bet that they are acquired around time W. How seriously we should take his subjects' reports of the time of their initial awareness of the urge, intention, or whatever, is a controversial question, and I will say nothing about it here.

Patrick Haggard, in his contribution to a recent discussion with Libet, asserts that "conceptual analysis could help" (Haggard and Libet 2001, p. 62). This article may be read as a test of his assertion. In my opinion, the result is positive. Attention not only to the data but also to the concepts in terms of which the data are analyzed makes it clear that Libet's striking claims about decisions, intentions, and free will are not justified by his results. Libet asserts that his "discovery that the brain unconsciously initiates the volitional process well before the person becomes aware of an intention or wish to act voluntarily . . . clearly has a profound impact on

how we view the nature of free will" (2004, p. 201). Not so. That, in certain settings, (roughly) proximal urges to do things arise unconsciously or issue from causes of which the agent is not conscious – urges on which the agent may or may not subsequently act – is a cause neither for worry nor for enthusiasm about free will.[17]

REFERENCES

Adams, F. and A. Mele. 1992. “The Intention/Volition Debate.” Canadian Journal of Philosophy 22: 323-38.

Brand, M. 1984. Intending and Acting. Cambridge: MIT Press.

Coles, M. and M. Rugg. 1995. "Event-Related Brain Potentials: An Introduction." In M. Rugg and M. Coles, eds. Electrophysiology of Mind. Oxford: Oxford University Press.

Deeke, L. 1996. “Planning, Preparation, Execution, and Imagery of Volitional Action.” Cognitive Brain Research 3: 59-64.

Frankfurt, H. 1988. The Importance of What We Care About. Cambridge: Cambridge University Press.

Gomes, G. 1999. “Volition and the Readiness Potential.” Journal of Consciousness Studies 6: 59-76.

Haggard, P. and Eimer, M. 1999. “On the Relation Between Brain Potentials and the Awareness of Voluntary Movements.” Experimental Brain Research 126: 128-33.

Haggard, P. and B. Libet. 2001. “Conscious Intention and Brain Activity.” Journal of Consciousness Studies 8: 47-63.

Haggard, P. and E. Magno. 1999. Localising Awareness of Action with Transcranial Magnetic Stimulation.” Experimental Brain Research 127: 102-07.

Haggard, P., C. Newman, and E. Magno. 1999. “On the Perceived Time of Voluntary Actions.” British Journal of Psychology 90: 291-303.

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

Libet, B. 2004. Mind Time. Cambridge, Mass. Harvard University Press.

___. 2001. “Consciousness, Free Action and the Brain.” Journal of Consciousness Studies 8: 59-65.

___. 1999. “Do We Have Free Will?” Journal of Consciousness Studies 6: 47-57.

___. 1992. “The Neural Time-Factor in Perception, Volition and Free Will.” Revue de Métaphysique et de Morale 2: 255-72.

___. 1989. “The Timing of a Subjective Experience.” Behavioral and Brain Sciences 12: 183-84.

___. 1985. “Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action.” Behavioral and Brain Sciences 8: 529-66.

Libet, B., C. Gleason, E. Wright, and D. Pearl. 1983. “Time of Unconscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential).” Brain 106: 623-42.

Libet, B., E. Wright, and A. Curtis. 1983. “Preparation- or Intention-to-Act, in Relation to Pre-Event Potentials Recorded at the Vertex.” Electroencephalography and Clinical Neurophysiology 56: 367-72.

Libet, B., E. Wright, and C. Gleason. 1982. “Readiness Potentials Preceding Unrestricted ‘Spontaneous’ vs. Pre-Planned Voluntary Acts.” Electroencephalography and Clinical Neurophysiology 54: 322-35.

Mele, A. n.d. “Decisions, Intentions, Urges, and Free Will: Why Libet Has Not Shown What He Says He Has.” In J. Campbell, M. O’Rourke, and D. Shier, eds. Explanation and Causation: Topics in Contemporary Philosophy (MIT Press, forthcoming).

___. 2003. Motivation and Agency. New York: Oxford University Press.

___. 1997. “Strength of Motivation and Being in Control: Learning from Libet.” American Philosophical Quarterly 34: 319-33.

___. 1992. Springs of Action: Understanding Intentional Behavior. New York: Oxford University Press.

Pereboom, D. 2001. Living Without Free Will. Cambridge: Cambridge University Press.

Smith, M. 2003. "Rational Capacities." In S. Stroud and C. Tappolet, eds. Weakness of will and practical irrationality. Oxford: Clarendon Press, pp. 17-38.

Trevena, J. and J. Miller. 2002. "Cortical Movement Preparation Before and After a Conscious Decision to Move." Consciousness and Cognition 11: 162-190.

NOTES

1. So if the occurrence of x (at time t1) indeterministically causes the occurrence of y (at t2), then a complete description of the universe at t1 together with a complete statement of the laws of nature does not entail that y occurs at t2. There was at most a high probability that the occurrence of x at t1 would cause the occurrence of y at t2.

2. In a later article, Libet writes: "the brain has begun the specific preparatory processes for the voluntary act well before the subject is even aware of any wish or intention to act" (1992, p. 263).

3. Some passages in which two or more of these terms are used interchangeably are quoted in Sections 3 and 4 below. Libet, Gleason et al report that "the subject was asked to note and later report the time of appearance of his conscious awareness of 'wanting' to perform a given self-initiated movement. The experience was also described as an 'urge' or 'intention' or 'decision' to move, though subjects usually settled for the words 'wanting' or 'urge'" (1983, p. 627).

4. For background on the generation, analysis, and use of electroencephalograms (EEGs) and "event-related brain potentials," including readiness potentials, see Coles and Rugg 1995.

5. I say "apparently," because an author may wish to distinguish an intention to flex one's wrist from an intention to initiate a flexing of one's wrist. I discuss initiation in Section 4. For completeness, I observe that if we instead ignore the quotation's first disjunct, it makes a claim about when an intention to prepare to flex – or to prepare to initiate a flexing of one's wrist – arises.

6. For a more thorough discussion of the experiment, see Libet et al 1983 or Libet, Gleason et al 1983.

7. I do not wish to exclude the possibility of such settledness in commissurotomy cases.

8. Another is that they have an intention to prepare to flex, if preparing is understood in such a way that so intending does not entail intending to flex.

9. Recall that Libet suggests that the activation event occurs between 10 and 90 ms before the onset of muscle motion (1985, p. 537) and later revises the lower limit to 50 ms (2004, pp. 137-38).

10. When does the action begin in all this – that is, the person's flexing his wrist or fingers? This is a conceptual question, of course: how one answers it depends on one's answer to the question "What is an action?" Libet identifies "the actual time of the voluntary motor act" with the time "indicated by EMG recorded from the appropriate muscle" (1985, p. 532). I favor an alternative position, but there is no need to disagree with Libet about this for the purposes of this article. Following Brand 1984, Frederick Adams and I have defended the thesis that overt intentional actions (i.e., intentional actions that essentially involve peripheral bodily motion) begin in the brain, just after the acquisition of a proximal intention; the action is proximally initiated by the acquisition of the intention (Adams and Mele 1992). (One virtue of this view is that it helps handle certain problems about deviant causal chains: see Mele 2003, ch. 2.) The relevant intention may be understood, in Libet's words, as an intention "to act now" (1989, p. 183; 1999, p. 54; 2004, p. 148), a proximal intention. (Of course, for Libet, as for me, "now" need not mean "this millisecond.") If I form the intention now to start running now, the action

that is my running may begin just after the intention is formed, even though the relevant muscular motions do not begin until milliseconds later.

11. A central point of disagreement between Haggard and Libet is usefully understood as a disagreement about when the "'act now' process" begins (see Haggard and Libet 2001; also see Haggard and Eimer 1999). Haggard apparently views the onset of lateralized response potentials (LRP), which happens "later than RP onset," as the beginning of the process (2001, p. 53; also see Trevena and Miller 2002).

12. Notice that in addition to "vetoing" urges for actions that are not yet in progress, agents can abort attempts, including attempts at relatively temporally "short" actions. When batting, baseball players often successfully halt the motion of their arms while a swing is in progress. Presumably, they acquire or form an intention to stop swinging while they are in the process of executing an intention to swing.

13. See Mele 1992, pp. 71-77, 143-44, 168-70, 176-77, 190-91. Those who view the connection as direct take the view that actions begin in the brain. See n. 11.

14. For example, unlike many critics, I did not challenge Libet's method for timing the relevant conscious experiences.

15. Again, a more cautious formulation of the urge hypothesis is disjunctive and includes the possibilities that what emerges around -550 ms is a (roughly) proximal urge to flex, a pretty reliable relatively proximal causal contributor to such an urge, a (roughly) proximal urge to "prepare" to flex, a simulation of an urge of either kind, and the motor preparedness typically associated with such urges.

16. Hereafter, the parenthetical clauses should be supplied by the reader. Intentions, in my view, are realized in physical states and events, and their causes are or are realized in physical states and events. I leave it open here that although intentions enter into causal explanations of actions, the causal work is done, not by them (qua intentions), but by their physical realizers. I forego discussion of the metaphysics of mental causation, but see Mele 1992, ch. 2.

17. Parts of this article derive from Mele n.d. [This is a draft to be revised in light of discussion at the conference.]