Chapter 7
The Neuroscience of Agency and Free Will

Markus E. Schlosser

Before I turn to my comments on Klemm's (Chap. 4, in this volume) target chapter, let me provide some notes about my background. I am a philosopher, and my main area of research is the philosophy of action. I started to read empirical research on human agency about four years ago in connection with taking up a research fellowship in a project on the philosophical implications of empirical studies of moral agency.[1] I was looking, first and foremost, for psychological and neuroscientific research on the various notions that lie at the core of philosophical theorizing about agency and free will—in particular, I was looking for research on intentional action, acting for reasons, decision making, long-term planning, free choice, and free action. I was struck by a number of things. First of all, although it is often remarked that neuroscience, and in particular cognitive neuroscience, is still in its infancy, I found an enormous amount of neuroscientific research on human behavior. Very little of it, however, was directly about the mentioned notions that are at the heart of philosophical accounts. One reason for this, no doubt, is that much of this research concerns the mechanisms and the details of movement control. Another reason, it seems, is that scientists tend to work with different concepts and conceptual frameworks. For instance, much of the literature that is relevant here can be found in the large body of research on executive control. The notion of executive control, however, is almost completely absent from the philosophical debate and from philosophical theorizing about agency. Further, it was surprisingly difficult to find overarching theories, overviews, and reviews that aim to integrate and unify the findings from different strands of research and different experimental paradigms. It was difficult, in other words, to see the big picture. Given this, I can only agree with Klemm, when he says that more neuroscience is needed (p. 1 in MS), and I would

[1] This project was entitled "Morality Beyond Illusions: Reassessing the Philosophical Implications of Empirical Studies of Moral Agency." It was funded by the Netherlands Organization for Scientific Research (NWO), led by Pauline Kleingeld, and hosted by the University of Leiden.

M. E. Schlosser (✉)
University College Dublin, Dublin, Ireland
e-mail: markus.schlosser@ucd.ie

C. W. Gruber et al. (eds.), *Constraints of Agency,* Annals of Theoretical Psychology 12,
DOI 10.1007/978-3-319-10130-9_7

add that overarching and integrative theories are needed in particular. In the more recent literature, one can find more attempts to integrate the findings from different strands of research, and Klemm's chapter is a welcome addition to this trend. In addition to the research mentioned by Klemm, one can find interesting theories and helpful overviews in Haggard (2008), Desmurget and Sirigu (2009), Cisek and Kalaska (2010), Gallivan et al. (2011), Momennejad and Haynes (2012), Desmurget (2013), and Brass et al. (2013).

Klemm provides his overview to the neuroscience of agency by way of ten axioms and ten propositions (at the request of the editor, as I understand). I found this to be an interesting and fruitful approach. I learned a great deal and I found myself in agreement with most of Klemm's claims.

There are, nevertheless, many issues in this rich chapter that deserve discussion. I decided to restrict my commentary to four main topics (in the following four sections). Most of my comments concern the particular claims and suggestions one can find in Klemm's contribution. But I will also use the opportunity to offer some more general reflections on the conceptual and methodological issues that arise for the scientific study of agency and free will. I will argue, in particular, that most neuroscientific studies of free will (and voluntary action) are based on an operational definition that is deeply flawed. With this, I hope to show that the philosophical concern with conceptual analysis and plausibility is not mere idle reflection if it concerns operational definitions that underlie experimental paradigms. Let me stress, though, that this criticism should also be understood as a plea for more interdisciplinary interaction. Neuroscience has delivered fascinating findings about agency, which should be of interest to philosophers. Experimental findings, however, are only as good as the experimental designs that deliver them, and experimental designs are only as good as the operational definitions and conceptual frameworks that underlie them. Given this, neuroscientists may well have something to gain by considering philosophical accounts of agency and free will, which are, after all, based on centuries of philosophical reflection and debate on the concepts in question.

7.1 Defining Agency

In his chapter, Klemm offers various reflections on the concept of agency and several characterizations of the nature of agency. The chapter does not provide a definite definition of agency, and it is not clear whether Klemm takes any one of the given characterizations as central. This is not a problem, I think, as one can distinguish plausibly between different *kinds* of agency, and as one need not give conditions that unify all kinds of agency in order to say interesting things about certain kinds of agency. In this section, I will first offer some comments on Klemm's characterizations of agency. Then I will outline how action and agency are usually conceptualized within philosophy. This will provide the background for a comment on Klemm's account of the process and sequence of intentional and conscious agency.

In the first section of Klemm's chapter, one can find the following suggestions on how to think about agency: agency as acting in the world, agency as animal action that arises out of its nervous system, and agency as something that is applicable to the whole animal nervous system (pp. 1–2 in MS). All three suggestions sound plausible, but there are also a number of issues. First, any definition or characterization of agency in terms of *action* is unsatisfactory as long as we are not given a further definition or characterization of the nature of action. Klemm, it seems, takes it for granted that we know what action is—that we know what distinguishes action from mere movement or unintentional behavior (we will return to this below). Given this, the first characterization is uninformative, if not circular. The second characterization also presupposes the notion of action, but it provides the additional constraint that agency is action that arises out of the animal's nervous system. This seems plausible, but it is not sufficient and perhaps not necessary for agency. Many things arise out of an animal's nervous system: sweating, hiccups, seizures, reflex movements, and so on. Of course, none of these things are proper actions. But this takes us right back to my first point: What is action? On the other hand, the condition (action that arises out of the animal's nervous system) does not seem necessary for mental agency and shared agency. Mental agency comprises things such as making a decision or trying to remember something. It does not seem correct to say that mental acts arise *out of* the nervous system—at least not in the sense in which movements arise out of the nervous system. Shared agency arises when agents act together as a group. This kind of agency does not arise out of any *one* nervous system. Similar worries apply to the third characterization. The whole animal system may undergo various changes that are not agency (sweating, hiccups, and so on). Further, mental agency does not seem to involve the whole animal system, and shared agency does not arise from one particular animal system.

This just shows, I take it, how difficult it is to come up with one definition or characterization of agency that fits all cases. Given this, it is only plausible to distinguish between *kinds* of agency. This is also what Klemm does when he distinguishes the agency of "higher animals" from simpler forms of agency (pp. 2–3 in MS). The distinguishing feature, according to Klemm, is that only higher animals act in accord with *intentions*. This is in line with the philosophical conception of action, to which I turn now.

As mentioned, it is plausible to hold that there are different kinds of agency. One definition of a kind of agency can be derived from the account of action that is widely shared and taken for granted within philosophy. On this standard view, all actions are intentional *under some description* (Anscombe 1957; Davidson 1963). The easiest way to explain this, without going into the technical details, is by way of an example. As I type these words on my computer, I am wearing down the keys of my keyboard. Typing these words is an intentional action of mine. Wearing down the keys is also something that I do, but it is not an intentional action. What makes this (wearing down the keys) an action, according to the standard view, is the fact that it is intentional under another description (namely, that of typing these words). So, on this view, intentionality distinguishes actions from other movements, behaviors, or events (such as slipping, falling, sweating, coughing, and so on). Further, most

versions of the standard view explain intentionality in terms of the agent's mental states and events and in terms of their causal roles. Very roughly, some movement (or event) is intentional, on this view, if it is caused and guided by an *intention,* and if this intention is based on the agent's *reasons* in the minimal or subjective sense that having the intention is caused and rationalized by the agent's desires and beliefs (for more on this, see, for instance, Enç 2003; Mele 2003).[2]

This philosophical account of action yields a straightforward definition of a higher kind of agency: intentional agency. On this approach, the exercise of this higher kind of agency consists simply in the performance of intentional actions (as defined above). This raises the question of what simpler forms of agency consist in, which I shall not discuss here. Let me suggest only that the notion of *goal-directedness* provides perhaps the best starting point here: It seems that an animal's movements can be goal directed even if they are not in any clear sense based on intentions, desires, and beliefs (for more on this, see Barandiaran et al. 2009, for instance).

With this as a background, let me now turn to Klemm's characterization of the process of intentional and conscious agency. On a number of occasions, Klemm points out that intentional and conscious agency is a temporally extended process, and he suggests that this process begins with an intention and ends with the execution of the action. Roughly, he suggests the following sequence (see p. 3 and 29 in MS):

> Intention, evaluation, decision, planning, execution.

Let me first point out here that this is out of line with the philosophical conception of intentional agency. On this conception, intentions are based on the agent's reasons, which are usually construed as the agent's desires and beliefs (or as the things that are represented by the agent's desires and beliefs). This would seem to correspond to the element of evaluation in Klemm's account. And this would mean that the two views disagree on the *order* of the elements in the sequence. Further, Klemm pulls apart intention and decision, whereas most philosophers hold that making a decision just *is* forming an intention. On their view, deciding to go the cinema tonight is nothing over and above forming the intention to go to the cinema tonight, for instance. Of course, once you have decided to go to the cinema, you have to make further decisions on what to see, how to get to the cinema, and so on. You have to make, what we may call, sub-decisions, which consist in the formation of sub-intentions (on how to the implement the goal). These sub-intentions fill out the further details concerning the means or the manner of attaining the goal. Given that such sub-intentions are usually also based on reasons, we can reconstruct the philosophical conception of the process as follows:

> Reasons (desires, beliefs, evaluation), decision (formation of an intention), sub-decisions concerning the means (formation of sub-intentions based on further reasons), execution.

[2] Note that this view applies also to mental and shared actions, provided that we can form some mental states intentionally and provided that the members of a group can be said to have shared intentions, goals, and beliefs (all of which might be based on agreement or some kind of voting system).

It seems that the difference between the two views is, in essence, the following. On Klemm's view, decision making concerns the question of how to implement an intention that stands at the beginning of the processes. According to the philosophical conception, reasons are at the beginning of the process.

I would like to make two points here. First, it is worth noting that most scientists who work on neuroscientific models of economic decision making use a conceptual framework that is closer to the proposed philosophical account of the sequence than to Klemm's. In particular, most models in neuroeconomics take also reasons (values or preferences) as the starting point of the sequence (for overviews and reviews, see Glimcher et al. 2009). Second, even though Klemm is very critical of the well-known (and notorious) neuroscientific experiments on free will (Libet 1985; Soon et al. 2008), his own reconstruction of the process of agency seems to exhibit one of their shortcomings, because these experiments also overlook or neglect the point that intentions and intentional actions are usually *based on reasons*. (I will say more about this in the following section.)

7.2 The Neuroscientific Study of Free Will

Most psychologists and neuroscientists seem to think that the belief in conscious agency and free will is illusory. Klemm is highly critical of their claims (see propositions 7 and 8, pp. 16–23 in MS), and I agree with most of what he has to say about this (for more on my take on these issues, see Schlosser 2012a, b, 2013, 2014). But, I also think that Klemm does not go quite far enough in his critique, and I would like to raise some more general conceptual and methodological issues here.

I agree with Klemm (p. 19 in MS) that the choices that are studied in the Libet experiment (1985) and in the follow-up experiment by Soon et al. (2008) are not representative of the choices that we make in our everyday lives—they are, at least, not representative of the more significant choices for which we hold each other responsible. However, when Klemm goes on to offer some further comments on this, he assumes a model of decision making in which the expected utilities of competing courses of action are evaluated before the final decision is made (see, in particular, the caption to Fig. 4.3, p. 21 in MS). It seems to me, however, that one problem with the mentioned experiments is that such a model of decision making does not apply at all, because participants are not presented with any real alternatives that can be evaluated or ranked. In the Libet experiment, participants are asked to perform a predefined movement when they feel like doing so. In the Soon et al. experiment, they are asked to press a button with either their left or right index finger when they feel like doing so. In both cases, participants have absolutely no reason to move now (rather than at some other time) and no reason to use one index finger (rather than the other one). Of course, in a sense they do have alternatives. But they do not have real alternatives in the sense that the options are indistinguishable in terms of their value and in terms of their consequences. There is simply nothing that can be compared and evaluated,

and so participants *cannot even begin* to engage in a process of *proper* decision making. So, not only are those choices not representative. They are not based on processes of proper decision making at all.

This assessment is fully in line with the model and the experiment provided by Schurger et al. (2012). According to their model, the decision in the Libet experiment is not based on any evidence (value or reason) at all. It is, rather, determined by random fluctuations in neuronal activity. They tested this model by means of a simulation of the Libet experiment and they conducted an experiment which confirmed the model. As they point out, their model is also consistent with the existence of the kind of pre-decision biases that were found in the Soon et al. experiment, which may "reflect stochastic fluctuations rather than an intentional (pre-conscious) decision-process" (Schurger et al. 2012, p. 6). Moreover, there is simply no reason to think that this stochastic model of decision making applies to other tasks and ordinary decisions, because in other tasks and in ordinary situations there is usually some evidence (value or reason) that the agent takes into account.

This brings me to a second point about the neuroscientific experiments on free will (and voluntary action). There is a very widespread trend or tradition in neuroscience to define free will (and voluntary action) by contrasting it with actions that are triggered or driven by external causes. I do not know whether Libet was the first who operationalized free will in this way. But this dichotomy between free (and voluntary) versus externally triggered (or driven) action plays a central role in Libet's argument against free will, and it provides the operational definition that underlies the design of the Libet experiment and of numerous other experiments on voluntary action thereafter (see Libet 1985; Jahanshahi and Frith 1998; Dreiber et al. 1999; Jenkins et al. 2000; Haggard 2008; Passingham et al. 2010; Hughes et al. 2011, for instance). An explicit commitment to this approach can be found in the review article by Patrick Haggard, who writes that a "scientifically […] satisfactory approach defines voluntary action by contrasting it with stimulus-driven actions" (2008, p. 934).

Elsewhere, I have argued at length that this conceptualization of free will is deeply flawed (Schlosser 2014). I will not try to summarize the full argument here, as my main point can be made effectively by means of an example. Suppose that you are sitting at your desk, working on something. At some point, the phone rings. Depending on the particular circumstances, and depending on your habits, you might respond in different ways. You might, for instance, immediately pick up the phone, perhaps because you have the habit of doing so. This does not mean that you would always respond in this way. For instance, if you have an urgent deadline to meet, you might either ignore the phone or you might pause for a moment and briefly consider whether or not you have the time to talk to someone right now. Comparisons between such possibilities support some important observations. First, in some cases, you may respond habitually or automatically, whereas in others, you may respond after a brief moment of deliberation. But, in each case in which you pick up the phone, you respond to

an external factor (cause, cue, or trigger). This highlights a first shortcoming of the mentioned dichotomy between free (or voluntary) versus externally triggered (or driven) actions. There are significant differences between different ways of responding to external factors: They are all *responses* to something, but they are not all *triggered* in an automatic fashion. Moreover, even if a response is triggered, it is simply not obvious that it is therefore involuntary or unfree. Suppose, for instance, that you are expecting an important phone call. Because of this, you might immediately pick up the phone as soon as it rings, and you might have the intention to pick up the phone as soon as it rings. In a sense, at least, your response would be triggered by an external cue. Does this mean that picking up the phone would therefore be involuntary or unfree? I do not think so. It is, at least, far from obvious that this response would be involuntary or unfree, especially if we take into account the fact that it might be *based on a prior intention* to respond quickly.

Given this, any approach that defines free will (or voluntary action) by contrasting it with actions that are performed in response to external factors would appear to be flawed. To take another example, suppose you take an umbrella in the morning in response to seeing dark clouds over the sky. This appears to be a free and voluntary action, and it may well be a free and voluntary action even if you did not consciously deliberate about whether or not to take an umbrella. It is, I contend, simply a mistake to assume that free choices (and voluntary actions) must not have external causes. In fact, some reflection on everyday decisions suggests that our choices and actions should usually have external causes, because they should be responsive or sensitive to external factors (such as dark clouds over the sky). Choices and actions that are altogether insensitive to environmental circumstances do even seem dysfunctional and random. Given this, it is rather unfortunate that the existing neuroscience of free will (and voluntary action) is largely about such choices—choices that are not based on any reasons and that are not made on the basis of anything that has significance or value.

It should be clear that this is not merely a conceptual or semantic issue. Far from it, the operational definition of free will and the design of the neuroscientific experiments are based on this problematic conception of free will. To his credit, Klemm does not reproduce this mistake. He says, for instance, that

> […] willed action requires some input to the brain circuits that generate willed action. Such input may come from an external contingency or may be generated internally from some emotional drive or motivation. (p. 20 in MS)

This suggests that, on Klemm's view, free and voluntary actions may well have external causes, and so it seems that Klemm departs from the common neuroscientific practice of defining free and voluntary actions by contrasting them with externally caused actions. This point deserves emphasis, because Klemm himself does not make this explicit, and because this point really does, in my opinion, uncover a very serious conceptual and methodological shortcoming of the neuroscientific experiments on free will and voluntary action.

7.3 Metaphysical Presuppositions: Dualism and Incompatibilism

In the two sections on the scientific challenges to conscious will and free will (pp. 16–23 in MS), Klemm addresses some issues that are connected to the mind–body problem and to the question of whether free will is compatible with determinism, which are traditional philosophical issues. I largely agree with the points he makes here, but I would nevertheless like to add a few remarks and observations.

When psychologists and neuroscientists draw their radical conclusions about the illusion of conscious will and free will, they sometimes presuppose dualism about the mind–body problem (Libet 2001; Haggard & Libet 2001; Wegner 2002) and they often presuppose incompatibilism about free will and determinism (Libet 2001; Haggard & Libet 2001; Haynes 2011). Again, Klemm does not seem to share these presuppositions, and again he deserves credit for this.

What if, Klemm asks (Chap. 4, in this volume, p. 16 in MS), consciousness is constituted by neural events? If consciousness is constituted by neural events, then it does not have to intervene, somehow, in neural processes. In my opinion, that is exactly the right starting point for thinking about the role of consciousness in the initiation and guidance of action; if consciousness is to play a role, we better construe it as something that is constituted or realized by neural events (states or processes).

Further, Klemm makes some remarks, which suggest that he is sympathetic to compatibilism about free will and determinism. It is not clear to me whether he meant to suggest this, and I do not know whether or not he is a compatibilist. But when he asks, rhetorically, whether we are slaves of reason when we make wise choices (p. 23), he certainly sounds like a compatibilist to me. Compatibilists are always keen to point out that not all kinds of causal determination rule out free will. They argue, in particular, that our choices may well be free if they are determined by our *reasons,* because determination by reasons is *persuasion,* not *coercion*. Klemm further notes that we would lack free will if our choices were immutable (p. 22). Again, this is something that compatibilists like to stress, because an agent (organism or system) that is causally determined need not be immutable at all. Determinism is perfectly compatible with development, learning, and frequent change.

No matter whether or not Klemm is a compatibilist, it is worth noting that some of his arguments and remarks are fully in line with compatibilism. More importantly, it is worth stressing that other neuroscientists should consider compatibilism more carefully as well. They should, at least, be aware of the fact that their conclusions about free will are often based on an unquestioned and unjustified assumption of incompatibilism just as much as they are based on empirical evidence.

7.4 The Nature and Role of Consciousness

As mentioned, Klemm suggests that consciousness is constituted by neural events. He holds, in particular, that it is constituted by certain circuit impulse patterns (CIPs). This leaves open the question of how we should think about consciousness to begin with. As Klemm notes, there seem to be two basic options (p. 24 in MS): We can think of consciousness as a certain type of *state*, or we can think of it as a certain type of *being*. Klemm does not consider the first option. He advances, instead, a version of the second: "consciousness is a nerve-impulse based brain state existing as a being that acts in the world" (proposition 10, pp. 24–26 in MS). I must say that I found it rather difficult to follow Klemm on this—both in the sense that I found it difficult to understand the view fully and in the sense that I found myself in disagreement with most of the things that I did understand. Before we turn to that, let me offer a brief remark on Klemm's representation of Descartes' view on consciousness.

According to Klemm, Descartes thought that the contents of consciousness are presented "on the stage of a 'Cartesian Theatre' for viewing by a virtual little man," and he goes on to say that this idea is nowadays "ridiculed by most scholars" (p. 24 in MS). Descartes, I should like to note, never held such a view and he never wrote anything that implies it. In fact, Klemm himself inadvertently ridicules Descartes' view by misrepresenting it in this way. Descartes never said or implied that consciousness is a "little man" and he never used the metaphor of a stage on which the contents of consciousness are presented for viewing. Yes, some philosophers (and scientists) talk that way about Descartes' views *in order to* ridicule them. But, this should not be taken to mean that he actually said such things.

At the core of Klemm's own view is the notion of an *avatar*, which is meant to replace and, in a sense, rehabilitate the notion of a "virtual little person" (or homunculus). An avatar is construed as a "being with agency": avatars "sense, evaluate, decide, and initiate and direct action" (p. 25 in MS). It is suggested that we should think about "consciousness as an avatar, generated by the brain as a set of CIPs to act on behalf of the best interests of the brain and body" (p. 25 in MS).

Let us first ask how this notion of an avatar can be combined with Klemm's naturalistic commitments—in particular, the commitment to the view that consciousness is constituted by neural events and processes (CIPs). On this, Klemm says that an avatar

> […] is a CIP representation of the body and what goes on inside and outside of the body, all referenced to the sense of self, which itself is a CIP representation of the "little person". I contend that the brain creates a conscious homunculus in the form of an avatar that it deploys to act on behalf of the embodied brain in ways not otherwise possible. (Klemm Chap. 4, in this volume, p. 25 in MS)

I find this rather difficult to understand. In particular, it seems to me that the notion of representation creates confusion here. Why is the sense of self a representation of a "little person"? If by sense of self we mean bodily awareness, then we can understand the sense of self in terms of perception, proprioception, and body schemata. If by sense of self we mean having an individual and narrative identity, then we can

understand it in terms of the contents of cognition and metacognition (beliefs about who I am, where I come from, and awareness of having those beliefs). If, finally, by sense of self we mean a sense of agency, then we can understand it in terms of having conscious intentions and in terms of anticipated and perceived feedback from the consequences of our movements. At no point, it seems to me, are we aware here of a "little person" when we have a sense of self or a sense of agency.

Further, it has not become clear to me what Klemm means by the claim that the avatar is acting *on behalf* of the embodied brain. According to the view, an avatar is consciousness construed as a conscious being with agency. But why should I think of my consciousness as acting on behalf of my embodied brain? This strikes me simply as a very odd thing to say. In my opinion, the problem with this suggestion stems, ultimately, from Klemm's decision to view consciousness as a being (entity or agent). If we take this as our starting point, then it is difficult to avoid the implication that I am actually *two* things: an embodied brain and a conscious agent. If I am, in this sense, two things, then it does make sense to say that one of the two is acting on behalf of the other. But for all I know, I am not two things, and my consciousness is not acting on behalf of my embodied brain when I am acting.

A much more plausible starting point is the assumption that conscious agency consists in the fact that conscious states or processes of mine, such as having or forming conscious intentions, play certain causal roles in the initiation and guidance of my actions. On this view, consciousness is not a being (entity or agent), but it consists in having conscious mental states that play certain causal roles. It is widely agreed that many mental states or processes are unconscious. When consciousness comes into being, it is not the case that a new being or agent comes into existence. Rather, when consciousness comes into being, certain mental states or processes become conscious. And I engage in conscious agency when such states or processes initiate and guide my actions. Of course, this view raises many questions and problems of its own, which are beyond the scope of this commentary. Let me stress, however, that it does avoid at least two of the issues that arise for the view proposed by Klemm. First, as the sketched view does not stipulate another being (entity or agent) apart from the embodied brain, it avoids the odd suggestion that consciousness is acting on behalf of the embodied brain. Second, it construes conscious agency in terms of the causal roles of conscious mental states. It offers, thereby, an explanation of what conscious *agency* consists in. In contrast, Klemm's view does not really explain conscious agency, because it refers to an entity, an avatar, which is *defined* as a conscious being *with agency*. This view, in other words, presupposes the notion it seeks to explain.

7.5 Concluding Remarks

In reply to critics, Benjamin Libet once made the following two points. First, he claimed that he had fully considered the implications of his experimental findings for the concept of free will (Libet 2002, p. 292). Second, he noted that the negative

criticisms of the findings and their implications have come mostly from "philosophers and others with no significant experience in experimental neuroscience" (p. 292). Libet, I think, committed two mistakes there. He did not fully explore the implications of the findings, as his conclusions about free will were based on an idiosyncratic and rather problematic definition of free will and on an unquestioned commitment to dualism and incompatibilism. Further, as I have pointed out, experimental designs are based, in part, on operational definitions, and operational definitions are based on conceptual frameworks. Yes, philosophers have usually no experience in experimental neuroscience, but they have experience in assessing the coherence and plausibility of conceptual frameworks and definitions. So, contrary to what Libet seemed to imply, the comments and criticisms from philosophers can be relevant to neuroscience, insofar as they can inform operational definitions. Klemm's chapter is to be commended for avoiding such mistakes. He does not define free will and voluntary action simply by contrasting it with externally driven action, and he explores a way of thinking about free will that departs from the unexamined commitment of other neuroscientists to incompatibilism.

To conclude, let me point out what I take to be the two most important points for future research on human agency and free will (from a philosophical point of view). First, I think scientists should reconsider the operational definition of free will (and voluntary action) in terms of the dichotomy between free (or voluntary) versus externally driven action. Second, before drawing radical conclusions about conscious agency and free will, scientists should consider the fact that such conclusions are usually based as much on metaphysical presuppositions (such as dualism and incompatibilism) as they are based on experimental results. As pointed out, Klemm avoids the corresponding mistakes, but he does not, in my opinion, emphasize the underlying issues strongly enough.

References

Anscombe, E. (1957). *Intention*. Oxford: Basil Blackwell.

Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior, 17,* 367–386.

Brass, M., Lynn, M. T., Demanet, J., & Rigoni, D. (2013). Imaging volition: What the brain can tell us about the will. *Experimental Brain Research, 229,* 301–312.

Cisek, P., & Kalaska, J. F. (2010). Interacting with a world full of action choices. *Annual Review of Neuroscience, 33,* 269–298.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy, 60,* 685–700.

Deiber, M. P., Honda, M., Ibanez, V., Sadato, N., & Hallett, M. (1999). Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: Effect of movement type and rate. *Journal of Neurophysiology, 81,* 3065–3077.

Desmurget, M. (2013). Searching for the neural correlates of conscious intention. *Journal of Cognitive Neuroscience, 25,* 830–833.

Desmurget, M., & Sirigu, A. (2009). A parietalpremotor network for movement intention and motor awareness. *Trends in Cognitive Sciences, 13,* 411–419.

Enç, B. (2003). *How we act: Causes, reasons, and intentions*. Oxford: Oxford University Press.

Gallivan, J. P., McLean, D. A., Valyear, K. F., Pettypiece, C. E., & Culham, J. C. (2011). Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *Journal of Neuroscience, 31,* 9599–9610.

Glimcher, P. W., Fehr, E., Camerer, C., & Poldrack, R. A. (Eds.). (2009). *Neuroeconomics: Decision making and the brain*. London: Elsevier Academic Press.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience, 9,* 934–946.

Haggard, P., & Libet, B. (2001). Conscious intention and brain activity. *Journal of Consciousness Studies, 8,* 47–63.

Haynes, J. D. (2011). Beyond Libet: Long-term prediction of free choices from neuroimaging signals. In W. Sinnott-Armstrong & L. Nadel (Eds.), *Conscious will and responsibility: A tribute to Benjamin Libet* (pp. 85–96). New York: Oxford University Press.

Hughes, G., Schütz-Bosbach, S., & Waszak, F. (2011). One action system or two? Evidence for common central preparatory mechanisms in voluntary and stimulus-driven actions. *Journal of Neuroscience, 31,* 16692–16699.

Jahanshahi, M., & Frith, C. D. (1998). Willed action and its impairments. *Cognitive Neuropsychology, 15,* 483–533.

Jenkins, I. H., Jahanshahi, M., Jueptner, M., Passingham, R. E., & Brooks, D. J. (2000). Self-initiated versus externally triggered movements. II. The effect of movement predictability on regional cerebral blood flow. *Brain: A Journal of Neurology, 123,* 1216–1228.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences, 8,* 529–566.

Libet, B. (2001). Consciousness, free action and the brain. *Journal of Consciousness Studies, 8,* 59–65.

Libet, B. (2002). The timing of mental events: Libet's experimental findings and their implications. *Consciousness and Cognition, 11,* 291–299.

Mele, A. R. (2003). *Motivation and agency*. Oxford: Oxford University Press.

Momennejad, I., & Haynes, J. D. (2012). Human anterior prefrontal cortex encodes the 'what' and 'when' of future intentions. *NeuroImage, 61,* 139–148.

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010). Medial frontal cortex: From self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences, 14,* 16–21.

Schlosser, M. E. (2012a). Causally efficacious intentions and the sense of agency: In defense of real mental causation. *Journal of Theoretical and Philosophical Psychology, 32,* 135–160.

Schlosser, M. E. (2012b). Free will and the unconscious precursors of choice. *Philosophical Psychology, 25,* 365–384.

Schlosser, M. E. (2013). Conscious will, reason-responsiveness, and moral responsibility. *Journal of Ethics, 17,* 205–232.

Schlosser, M. E. (2014). The neuroscientific study of free will: A diagnosis of the controversy. *Synthese, 191,* 245–262.

Schurger, A., Sitta, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences, 109*(42), E2904–E2913.

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience, 11*, 543–545.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge: MIT Press.