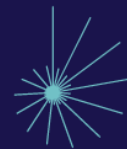


*Recent Work on
Agency, Freedom,
and
Responsibility:
A Review*

July 2019

Robyn Repko Waller
Department of Philosophy, Iona College



JOHN TEMPLETON
FOUNDATION

Table of Contents

I. Introduction	3
II. Background: Free will as a control over what?	4
III. Metaphysical issues: What does a characterization of free will look like?	6
IV. Neuroscientific findings: How do humans actually exercise agency?.....	19
V. Psychology of free will: how do humans actually exercise agency?	34
VI. Theological issues: can the nature of the Divine limit or rule out free will?	45
VII. Conclusion.....	48
VIII. References.....	49

I. INTRODUCTION

Free will is a familiar, ingrained feature of human life. When we decide what to do – for instance, decide to pick up that cup of coffee, we may unreflectively believe that it was up to us what we decided. When we move about in the world – for example, running along the trail, we may unreflectively believe that it is up to us when and how we act. Here ‘up to us’ indicates a kind of control over our decisions and actions. Intentional actions are a kind of goal directed activity over which the agent has some control.¹ We humans as social beings take it that, typically, other humans too have this kind of control over their decisions and actions. Moreover, we take it that humans extend this control not over just innocuous matters, like sipping coffee, but over morally weighty ones too, such as whether and how to care for and treat our dependents, affiliates, and distant others.

If we believe that at least some of our decisions and actions, especially our intentional bodily movements and their consequences, are up to us, we may believe that we are freely deciding or acting, at least at times. Accounts of free will differ widely, but most free will theorists take free will to be the ability to act freely. An agent who has free will has the appropriate kind of control over her decisions and actions to be an apt target of the responsibility practices (see, e.g., Mele 2013a for background on this linking of free will and moral responsibility).² Here by responsibility practices, we’ll understand these practices broadly to include not only moral practices of blaming or praising an agent (e.g., for moral bad and good actions), but also legal practices of punishment as well as divine judgments and punishments of an agent’s actions (e.g., for sins and saintly acts). Others take the control typically associated with responsibility to be valuable in its own right as grounding our human nature as self-governing agents.

Yet, on reflection, our conception of ourselves as free and responsible agents raises many questions: What kinds of abilities do free agents need? What kinds of conditions make possible or prohibit our free agency? Do we humans, in fact, possess those abilities in the right conditions – i.e., do we exercise free will? If not, should we still be held responsible? What would institutions, such as the everyday moral system and legal system, look like without the assumption of free will?

This paper surveys recent work on these questions from a wide array of disciplines, including philosophy, psychology, and neuroscience. In Section II, I will provide a background on the discussion of free will and why possessing free will is held to be crucial for our conception of ourselves as self-governing and responsible agents. The remainder of the paper (Sections III-VI) will be devoted to exploring innovative new research which grapples with our place in the world as free agents. In those sections I will outline the proposed threat of determinism, of three kinds – metaphysical, neurological/physical, and divine, to our ability to exercise control over our lives. I will then, correspondingly, divide these research contributions

¹ Non-human animals engage in goal-directed behaviors in some sense, but here the focus is a subset of goal-directed behaviors, human intentional actions.

² See, e.g., Fischer and Ravizza (1998). If one is an apt target of the moral responsibility practices, then one qualifies as the type of agent deserving of moral blame (for a morally bad action) and moral praise (for a morally good action) (see, e.g., Fischer and Tognazzini 2011; Watson 1996).

(Sections III-VI) in terms of their free will focus: metaphysical, neurological/psychological, and divine issues.

II. BACKGROUND: FREE WILL AS CONTROL OVER WHAT?

Exercising one's free will involves an exercise of control, or influence, over oneself. But what exactly must an agent control in order to exercise free will? One answer is that free will requires control over one's actions (and perhaps the consequences of those actions). This sense of free will is sometimes referred to as free action.

Here by 'actions' we mean overt actions, such as bodily movements and the effects of those movements. This makes sense: When we hold people responsible – e.g., blame them, punish them, feel gratitude toward them, we tend to hold them responsible for their conduct, or actions, and the repercussions of those actions. And so on (at least) a backward-looking notion of responsibility,³ we care about whether the agent had control over her actions. Why?

For instance, suppose on a crowded sidewalk Jan shoves into a group of pedestrians, causing an elderly pedestrian in the group to fall onto a crosswalk and sustain a serious injury. If we come to find out that Jan had a seizure and her “shoving” was the result of that unexpected seizure, it would be strange and inappropriate to feel (or express) blame or anger or indignation at Jan. (Although we may still rightly feel that the incident was unfortunate for all involved.) Her bodily movement and the resulting shove were not 'her own' or 'up to her' – she the agent did not authorize or causally initiate the action, she did not desire or intend to shove into or recklessly endanger others. She did not control her action. Instead it was a passive, involuntary movement.

Now suppose instead that circumstances were different: Jan is late to a Broadway play, one for which she has paid significant money. She doesn't want to miss the curtain's going up and the doors closing for seating. She runs down the crowded sidewalk and, knowingly, straight into a group of pedestrians. The elderly pedestrian falls and sustains a serious injury. Here of course the group of pedestrians and especially the injured party should be permitted, and likely would, express blame, anger, resentment (from the injured party), and indignation (from the spectators) at Jan for her reckless action and the resulting injury. To achieve her goal of not missing the opening curtain, Jan intentionally ran down the crowded sidewalk, knowing that she may well knock into others and that someone may get injured (even if Jan didn't intend for others to get injured). Hence, her shoving of the pedestrians and the elderly pedestrian's injury are the result of how she intentionally moved her body, an exercise of control over her actions. Moreover, the

³ See Pereboom 2015 for a clear discussion of distinct kinds of backward-looking and forward-looking senses of responsibility. On a forward-looking account of responsibility, we look to the effect of holding the agent responsible on future states of affairs – e.g., it will aid in preventing her (or others) from acting in a similar manner in the future. On a backward-looking account of responsibility, we look to see that the response, or treatment of her, fits or is deserving given her past action and state of mind – e.g., she knowingly expressed ill-will (or lack of good will) toward others in her action, so the response is a fitting expression of condemnation of that ill-will (or lack of good will).

action expresses something about Jan as an agent — she was reckless, knowingly taking the risk of injuring others. In this way the action was up to Jan.

So it seems that the agent's state of mind — goals, desires, perceptions — at the time of action, in addition to how she moves her body, plays a central role in whether and to what extent she controls her actions and so acts freely. And that state of mind may well be heavily influenced by the kind of person, morally and pragmatically, that Jan is. For instance, if Jan were the kind of person who often put others before herself in life or perhaps just is a laidback individual, she wouldn't have thought to rush to the theatre. We tend to believe, and endorse in our attributions and judgments of others, that people have characters, or a stable set of traits, that inform their decisions and dispose them to act in certain ways.⁴ But now one may reasonably ask the following: even if an agent acts freely, how much control does the agent have over her state of mind?⁵ In particular, are agents' decisions really up to them? Here we can separate two further senses in which agents might be said to have free will:

(1) The agent exercises control over her decision, or broader, motivation for her action. Call this freedom of the will, or volitional freedom.

(2) The agent exercises control over the psychological features of her that cause, or more broadly, inform her decision (and so action). Call this ultimacy, short for ultimate sourcehood.

Hence, agents who act on similar decisions in similar circumstances may differ in the control that they exercise over their will and character. (Here by 'decision,' we mean practical decisions, or decisions about what to do, as opposed to decisions about what is the case — see, e.g., Mele 2010a; Mele 2010c). Everyday considerations and criminal law are full of talk of mitigating circumstances due to lack of control — someone's mindset and actions might be judged to be bad, but that agent have been subject to horrific childhood neglect and violence. To the extent that the development of character and self-control capacities are out of one's hands, we tend to soften or change our judgment and treatment of those agents. For instance, we might excuse a freshly bereaved person for acting hostilely to those in his close social circle — although friends may ordinarily take offense and so express anger and moral blame for slights. Here stress and extreme emotion may be driving the person to act in ways out of line with who he is. Or, in some cases, it may be on the whole inappropriate to blame the offending agent in question at all for any ill conduct. For example, in cases of immaturity due to youth or lack of moral (or reasoning) capacity due to disorder, certain agents can be outside of the realm of the responsibility practices (at least morally) because they lack the relevant control capacities altogether: The moral reaction of others to a toddler or

⁴ For early psychological work that challenges this assumption, see Mischel 1973. For recent challenges to this commonsense assumption, see, e.g., Doris 2002 and Mele and Shepherd 2013, among others.

⁵ This paper focuses on questions of control in relation to free will and moral responsibility. I have included in these questions some issues of how and whether agents know of the consequences of their actions. There is, however, a tradition in the literature of discussing freedom (control) requirements for responsibility as separate from epistemic (or knowledge) requirements for free will. Whether or not philosophers of free will do or should distinguish the two is debatable in itself (see Mele 2010d).

schizophrenic patient who causes serious injury is of a different kind than the reaction to a healthy neurotypical adult who acts similarly.⁶ (See Glannon 2012a for a discussion of how academic theories of control bear upon the agency of those with mental disorder).

A central focus, then, in the discussion of free will are the following issues regarding control:

What exactly must an agent control to qualify as a responsible agent for her actions? (e.g. free action, volitional freedom, ultimacy)

What are the abilities necessary or sufficient (or both) for the (relevant) free will and so responsible agency?

What are the necessary and prohibiting conditions for free will? That is, which conditions are required for or rule out agents having free will?

Which conditions may enhance or limit our exercise of free will?

For each section III through VI, I will provide a brief taxonomy of accounts of free will in terms of the issue of determinism. Determinism will be explained in terms of metaphysical determinism, physical/neurological determinism, and divine foreknowledge and intervention. Each kind has relevance for various ways an agent might be held responsible (e.g., morally, legally, punishment and reward for sins). This taxonomy will help guide readers in situating the accounts of free will offered and assessed in the review. There recent contributions to the literature will correspondingly be organized in terms of metaphysical, neurological/physical, and divine issues of free will.

III. METAPHYSICAL ISSUES: WHAT DOES A CHARACTERIZATION OF FREE WILL LOOK LIKE?

We've discussed how much of the free will literature revolves around the question of what kind of control an agent must have to be responsible for her conduct (and to be a self-governing agent). The debate in both the philosophical literature centers on the following question of

Can agents have free will if determinism is true (i.e., if the world is deterministic)?

Call this question the *compatibility question*, as distinct from the question of whether determinism is true — that is, whether the world is in fact deterministic or not. This question is also distinct from the following related one:

⁶ See Strawson 1963, a foundational paper in the literature, for an account of moral responsibility practices in terms of reactive attitudes and conditions for excuse and exemption.

Do some actual agents have free will?

Call this the *existence question*.

I have not yet said what determinism is, which is an important thesis in the larger discussion. However, given that this paper concerns at least three related senses of determinism and their compatibility with free will, I will temporarily put the details aside. Instead, let's label the positions in the free will debate according to their stance on the compatibility and existence questions.

Compatibilists hold that free will is compatible with determinism.

That is, compatibilists answer 'yes' to the compatibility question. (Compatibilism might sound odd to readers, given that the popular debate is often framed as freedom from determinism – or freedom versus determinism, but the following discussion of determinism will motivate why compatibilism is a genuine possibility, at least to some philosophers and perhaps nonphilosophers too.) It is important to note that most compatibilists do not think that the truth of determinism is required for agents to act freely, but merely that determinism in itself is no threat to agents exercising free will. Philosophers sometimes state this idea in terms of 'possible worlds': For the compatibilist, there is a possible world in which determinism is true of that world and some agent acts freely in that world. Further, compatibilists tend to answer 'yes' to the existence questions: actual agents have the control capacities to be considered free agents. In contrast,

Incompatibilists hold that free will is incompatible with determinism.

That is, incompatibilists answer 'no' to the compatibility question: if determinism is true of that world, then no one acts freely in that world. Incompatibilists are divided further into camps based on their stance on the existence question.

Incompatibilists who argue that some actual agents do in fact possess free will are called *libertarians*.

Incompatibilists who argue that no actual agents have free will are termed *free will skeptics*.

The latter group are skeptical that anyone has the control over themselves required to be a responsible agent or to be self-governing. One may be a free will skeptic because either one takes determinism to be true – *hard determinism* – or one also takes indeterminism to preclude free will – *hard incompatibilism*.

Given this talk of determinism, it is time to characterize three kinds of determinism free will theorists discuss as potentially incompatible to free will. Doing so will make it clear why compatibilists and incompatibilists of different stripes lay out the control capacities and enabling conditions they do for their accounts of free will. In this section on metaphysical issues, we will consider the following kind of determinism as relevant for free will:

Metaphysical determinism: A complete true description of the laws of nature and state of the world at a time [logically] entails all truths about that world.⁷

One way to think about this idea is in terms of the events of the universe being rewound and replayed.⁸ This “rollback” explanation of metaphysical determinism is expressed well in Nahmias et al. 2006:

Imagine there is a universe (Universe C) that is re-created over and over again, starting from the exact same initial conditions and with all the same laws of nature. In this universe the same initial conditions and the same laws of nature cause the exact same events for the entire history of the universe, so that every single time the universe is re-created, everything must happen the exact same way. For instance, in this universe whenever a person decides to do something, every time the universe is re-created, that person decides to do the same thing at that time and then does it...For instance, in this universe a person named Jill decides to steal a necklace at a particular time and then steals it, and every time the universe is re-created, Jill decides to steal the necklace at that time and then steals it. (p. 38)

Notice that this kind of determinism is all inclusive — it’s not just the facts about the agent such as her desires and values and situation that remain the same in each rollback. Rather, if *everything* about that world is held fixed, then the same event must occur. Two things are especially noteworthy about this: First, metaphysical determinism is consistent with an agent’s doing something else if something major or even minor had been different prior to her decision and action (e.g., the reasons she considered for stealing the necklace). Second, metaphysical determinism is consistent with physicalism, the view that the facts about the world are exhaustively fixed by the physical facts of the world.

Some compatibilists, historically and in contemporary debates, have embraced that all that is required for free will is what we have called *free action*. The agent need only exercise control in the sense that her action is the product of her internal state of mind.⁹ If Jill steals a necklace, then she did so freely insofar as she stole the necklace due to her desire or intention to do so. If she had been in a different state of mind — say, considered the disappointment that her family might feel about her being a thief and entertained anticipated guilty and shame — then she might have not stolen it. But if, on the contrary, we find out that she was coerced or compelled to steal on that occasion — say, at gunpoint by another person or due to a severe kleptomaniac compulsion, then we might say that she does not have the right control over action to act freely at that time. Compatibilists of this stripe then bid you to notice that we can distinguish coerced and compelled actions from agent-driven actions even in deterministic worlds. So determinism is not a threat to free will.

⁷ van Inwagen (1983)

⁸ The former is an entailment claim, while the latter is in terms of causation talk.

⁹ For an overview of Hobbes and Hume’s views on free will, see Russell (1988)

Incompatibilists typically agree that lack of compulsion or coercion is necessary for agents to have free will, however incompatibilists don't take this kind of freedom of action to be enough control for free will. Instead they press: How can one be in control of what one does here and now if any action one takes is the consequence of the laws of nature and the (complete) conditions of the world before one's birth? This idea is captured in van Inwagen's Consequence Argument, an informal version of which is expressed as follows:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore the consequences of these things (including our own actions) are not up to us (1983, p. 16).

The argument assumes, then, that for us to have free will our actions have to be 'up to us' in the sense that, roughly, there is something we could do now to change or affect which actions we take.¹⁰ If the Consequence Argument is taken as a sound argument, this suggests two things: First, that the control relevant for free will is the control to decide which decision or action (or inaction) to take and that that requires some 'deep openness' of the future. As Mele puts deep openness in his *A Dialogue on Free Will and Science*:

What you need is that more than one option is open to you, given everything as it actually was at the time – your mood, all of your thoughts and feelings, your brain, your environment, and indeed, the entire universe and its entire history (2014b, p. 8).

This deep openness is ruled out in deterministic worlds. Second, note that on this picture of free will, the agent needs to do more than exercise *free action* to have free will: the agent needs *volitional freedom* and (or) *ultimacy*.

Hence, incompatibilists begin to fill out their account of volitional freedom as, first, requiring indeterminism when an agent decides or acts freely. Second, the abilities of the agent which underscore control are then specified. The further details of any given free will account rely on its answer to what I'll call the *criteria question*:

What are the agential abilities necessary and (or) sufficient for free will?

Event-causalists argue that free decisions and actions must be indeterministically caused by mental states, such as beliefs, desires, intentions, consideration of reasons, etc.¹¹ On these accounts, agents exercise free

¹⁰ This kind of agential control is sometimes understood in terms of agent's having other alternatives to their actual action, or the ability to do otherwise. Frankfurt 1969 offers a well-known proposed counterexample to the claim that an agent is morally responsible for her action only if she could have done otherwise. There is a thriving literature revolving around these counterexamples (see, e.g., Hunt and Shabo 2013 for one such paper out of many).

¹¹ Some incompatibilists are noncausalists about free action (see, e.g., McCann 1998).

will when their state of mind brings about their decision and action (in the absence of manipulation, compulsion, or coercion), AND a different decision or action could have occurred holding fixed everything prior to the decision (or action).¹²

Other incompatibilists, agent-causalists, set the bar higher. They are concerned that if all of our decisions and actions are just the result of events and processes in the world, albeit a special kind – mental events and processes – then agent as a self or as a person isn't in control or self-governing. Hence the agent-causalists hold that the agent is something more than or distinct from these beliefs, desires, and values. Agent-causalists argue that free decisions and actions must be caused, or brought about directly, by the agent herself and not (just) her mental states. Here the agent could be, but need not be, understood as a soul. The idea is that the agent who exercises free will is a substance that has an active power to determine her decisions and actions. Otherwise, the thought goes, what would action be but just another event in a world of events (e.g., fires burning, fish swimming, etc.)? In this way, incompatibilists, especially agent-causal incompatibilists, contend that free will involves ultimate sourcehood, the agent contributes to her decisions and actions independently of past events.¹³

We should pause to note the following at this stage: First, it bears keeping in mind that the incompatibilist can be either a libertarian or a free will skeptic. That is, answer 'yes' or 'no' to the question of whether agents actually met these conditions for free will (the *existence question*). Second, importantly, compatibilists may also agree that free will involves volitional freedom or ultimacy beyond free action. However, such compatibilists will argue that the conditions for having volitional freedom and ultimate sourcehood are compatible with determinism: Mele characterizes compatibilist accounts broadly as being of the following form: "a competent decision-maker who hasn't been manipulated and isn't being compelled or coerced" (2014b, p. 6). Roughly, in addition to lack of manipulation and compulsion, agents have volitional freedom or ultimate sourcehood in the compatibilist sense that they understand reasons for actions and act because of those reasons or act on desires that are really their own (e.g., they endorse or approve of those desires).¹⁴ And, in fact, many event-causalist incompatibilists too endorse these kinds of capacities in their answer to the criteria question. The major difference, then, is that the incompatibilists require a lack of determinism in addition to these agential abilities underscoring free will. (For an extremely accessible and clear guide to the basic philosophical positions in the free will and their differences, see Mele 2012j.) Third, it's important to note that metaphysical accounts of free will are typically tied to an agent's status as a responsible agent in the moral sense: whether she is deserving of praise (for a morally good action) or blame (for a morally bad action). (For a discussion of the relationship between free will and moral responsibility, see Mele 2015b and Mele 2017d, Ch. 5). To a lesser extent, similar conditions of control are subject of legal mens rea – the quality of mind of a defendant for a crime – and so pertinent for issues of criminal responsibility (e.g., desert of legal punishment).¹⁵

¹² For event-causal accounts, see, for example, Kane (1996), Balaguer (2004), Franklin (2011), Franklin (2013).

¹³ For agent-causal incompatibilist accounts of free will, see, for example, O'Connor (2000) and Clarke (2003).

¹⁴ For a sample of such compatibilist accounts, see Fischer and Ravizza (1998), Wallace (1994), and Frankfurt (1971).

¹⁵ See Mele (2012a) for a discussion of control and criminal responsibility. See also section IV.4.

The kinds of free will accounts introduced here all specify the metaphysical conditions that would need to be met for agents to have free will and how free will is related to metaphysical determinism. We'll now explore how recent research has provided further insight into these metaphysical issues of free will.

III.1 What is intentional action?

We've discussed the kinds and amount of control that various theorists hold one must have to be responsible for one's action. However, another key question is for which items are agents in control in the relevant respect and responsible? It seems plausible that one can exercise control over decisions and bodily movements, if one controls anything. So agents can be responsible for these items if so. Beyond this, philosophers disagree about whether and to what extent agents are responsible for their attitudes, implicit and explicit, and the consequences of their actions (see, e.g., Levy 2014 for work on responsibility for attitudes). Nevertheless, to understand what it means to be responsible for decisions and intentional actions requires that we have a handle on what decisions and intentional actions are. Here I outline a sampling of recent work on these topics from action theory.

Let's examine Jan's decision to run through the crowd of pedestrians: imagine that Jan, stuck on the subway, deliberated about what to do: should she run to the theatre to make the curtain's going up or walk at a safe pace? Jan considers her reasons for each option. She might believe that if she doesn't rush, then she will miss the start of the show and will not be seated. She wants to make the start of the show. She believes she paid good money for her tickets and (falling prey to the sunk cost fallacy) desires not to waste her money. On the other hand, she believes that it is dangerous to run through crowds, and she desires to act safely. Nevertheless, she decides to rush to the theater. Here Jan makes a practical decision, she forms an intention about what to do. This intention indicates Jan's settledness on a plan of action. More broadly, agents who have intentions to act are settled on acting in such-and-such way (Mele 2014b; Mele 2017d, Ch. 2). Here Jan's deciding was an active matter, a mental action. What role did her reasons play in her deciding as she did? Here action theorists disagree: McCann (2012; 2013b) argues that while reasons explain why Jan decided as she did, they did not cause her decision. Reasons on a noncausal view can serve justifying and explanatory roles, but are not causes of actions. In contrast, others, such as Mele (2013g; 2013h; 2017d, Ch. 3), articulate and defend a version of causalism, that not only can reasons be cited in explanations of decisions and actions, but further that reasons can be among the causes of decisions and actions.

Imagine that when Jan begins to run, she startles a pigeon and the pigeon flies away. Moreover, she also lifts up her left foot. Is Jan's startling a pigeon and lifting her left foot distinct from her action of beginning to run? That is, are they the same or different actions? She began to run intentionally — doing so was initiated by her intention to run. Did she also startle the pigeon and lift her left foot intentionally? The former question refers to the matter of action individuation, or how to distinguish or individuate actions. The latter question is related to the issue of what makes an action an intentional action. The literature on these questions is vast, but some recent philosophical writing on the topic include McCann 2013a, Mele 2014a, Mele 2017d. Recently, philosophers and psychologists have surveyed non philosophers' judgments as to whether certain actions are intentional in an effort to understand what kinds of factors drive judgments

of intentionality (for a seminal study, see Knobe 2003, and for related studies see Mele 2012b; Mele 2012h).

III.2 Philosophical accounts of self-control and self-deception

Action theorists may refer, at times, to the empirical literature on how people actually act or what non-philosophers believe about actions, but generally action theory is a normative field. Action theorists seek to understand decisions, intentional action, free actions, etc., in terms of rational agency. One way to see this is to think of agency on a continuum. On one end of the spectrum, an agent may be (ideally, if not ever in practice) perfectly self-controlled, an exemplar of rational agency. Such an agent demonstrates *enkrateia* (to use Aristotle's term [Nicomachean Ethics]), or self-control (as I will refer to it). That is, the self-controlled agent's actions always align with what she judges it best to do, on the whole, given how she is psychologically (her beliefs, values, desires). If you are an agent with ideal self-control, then if you judge it best to stop eating pastries at the coffee shop (despite your strong desire to eat them), then you will pass up the pastries every visit to the coffee shop. The other end of the spectrum is *akrasia*, or weakness of will (as I will refer to it): The weak-willed agent acts contrary to his on-the-whole judgment of what it is best to do. The extreme weak-willed agent always eats the pastry when in the coffee shop despite his judgment that he shouldn't. Normal human agents fall somewhere in the middle of this spectrum, in terms of their trait self-control. Here self-control as a character trait is distinct from the virtue of temperance (Mele 2016b).

Contemporary action theorists have typically, although not universally, focused on whether any particular action (as opposed to the agent) is weak-willed (see, e.g., Mele 2011c for a concise introduction of this historically well-established topic and Mele 2012l for a book length treatment). This is sometimes also referred to as weak-willed action.¹⁶ One of the major questions regarding weakness of will/*akrasia* is whether agents can act freely when they intentionally act contrary to their judgment about what's best to do — referred to as strict *akratic* action (Mele 2011c). Another point of discussion concerns whether how closely exercises of ideal self-control are tied to the influence of reason to the exclusion of the 'passions' or desires (Mele 2011d). Further, we can distinguish how agents can exercise self-control in their actions at a time (synchronic) versus over time (diachronic). Each kind of exercise of self-control raises conceptual issues for which philosophers have offered proposals (e.g., Mele 2013e; Mele 2012l). (For an interview-style overview of this research, see the interview with Mele in *The Philosopher's Magazine* (2013m).) Levy (2016a) interestingly applies a broader notion of autonomy-related self-control to the issue of whether addicts can give informed consent for recovery treatments and programs, such as supervised heroin injections. Other intriguing topics regarding rationality, which will not be explored here given space constraints, include philosophical accounts of self-deception (see, e.g., Mele 2012f). The issue of self-control, broadly construed, from a psychological perspective will be taken up again in section V.1 below.

¹⁶ See, however, Holton 1999 for a distinct usage of this term.

III.3 Arguments for and against compatibilism

Compatibilists run the gamut from those who argue that mere free action is sufficient for free will to those who argue for determinism-friendly volitional freedom or determinism-friendly ultimacy. What is common to all such accounts, though, is that the control capacities relevant for moral responsibility are compatible with determinism. That is, even if our universe is deterministic, this is no barrier in and of itself to agents acting freely.

One of the dominant objections to compatibilism in the contemporary literature is a family of arguments called ‘manipulation arguments.’ Manipulation is a species of influence. By manipulation here, I mean cases in which a person wants another person, the agent, to do something or be some way, but does not want to overtly force her to do it (or be that way). Manipulation can range from subtle suggestion and nudging to full-blown science-fiction-like agent-design. The relevance of manipulation to a person’s autonomy is a critical debate undertaken by, among other disciplines, philosophy and political science (e.g., manipulation arguments against certain accounts of free will and moral responsibility; the relationship between coercion and manipulation in political philosophy). In particular, in the free will domain incompatibilists exploit what is thought to be a common intuition about the responsibility-undermining feature of manipulation. Assume that our agent Jan satisfies some set of sufficient compatibilist conditions for free will and moral responsibility. These conditions include capacities such as appreciation of and ability to act for reasons or a well-ordered set of desires and background conditions such as lack of coercion. We might well judge that she acted freely and deserves blame for injuring the elderly pedestrian and acting so recklessly. But imagine we find out that Jan’s history is a bit unusual. Perhaps a goddess — who has perfect knowledge of the laws of nature and state of the world in her deterministic universe — designed Jan’s zygote just so Jan would eventually, forty years after her birth, rush into the crowd and injure the pedestrian. Or alternatively, perhaps the night before the Broadway play, Jan had gone to bed a moral saint — with only virtuous desires, intentions, and values — but had her psychology modified by a team of nefarious scientists who intended to make her into the sort of person who would genuinely endorse performing reckless and selfish actions. In essence the incompatibilist strategy is to illicit an intuition that the agent isn’t acting freely or responsible in these cases, but that these agents may satisfy compatibilist sufficient conditions for free will. There is a wide variety of such subtle manipulation arguments; Mele (2013c; 2013d; Mele 2016a) provides fuller versions of the kinds of manipulation mentioned here as well as other types of manipulation that may challenge compatibilist accounts of free will and responsible. Mele also outlines the strategies that compatibilists have adopted to respond to manipulation arguments and notably offers a compatibilist account that is designed to handle such challenges. (For a book-length treatment of manipulation, see Mele 2019.)

Key to the debate regarding manipulation arguments are both specialist and nonspecialists’ intuitions about manipulation cases. Hence, psychologists and philosophers have turned their attention to people’s intuitions about cases of manipulation and the psychological mechanisms that drive people to judge that manipulation compromises agency and responsibility when it does so (for a general discussion of how empirical surveys of nonphilosophers’ judgments on free will and responsibility contribute to the philosophical debate, see Bjornsson and Pereboom 2016). For instance, Sripada (2012) uses statistical

analysis of non philosophers' judgements about manipulation cases to argue that intuitions of no (or mitigated) responsibility aren't driven by the lack of agential control due to determinism. Rather, participants seem more concerned that such manipulated agents aren't acting from their true, or "deep self," and this factor drives judgments of responsibility. Bjornsson (2016) conducted similar studies to those of Sripada but concludes on the basis of his statistical analysis that what's driving folks to judge that manipulated agents are not (or less) responsible isn't the deep self factor; It's that the agent's actions can be explained by factors outside of her control. Further, Phillips and Shaw (2015), using surveys of non philosophers' judgments about cases, have supported the claim that when agents are subject to threats and manipulation, nonphilosophers judge such agents as less blameworthy than they would otherwise. Their findings suggest that this reduction in blameworthiness is due to the perception that the manipulator is controlling the agent and his actions. These psychological findings can either lend support or undermine key factors that philosophers argue either are (or ought to) drive judgments of control and responsibility in the face of manipulation.

III.4 What are the arguments for and against incompatibilism?

We've noted that typically incompatibilists reject the possibility of free will and moral responsibility in a deterministic world because (they argue) determined agents aren't in control of their decisions and actions. Hence, in their accounts of the conditions for free will incompatibilists require that agents' decision-making and action production isn't deterministic. That is, such processes and events are indeterministic. Incompatibilists argue that this element of indeterminism is a necessary condition of free will so that the agent is the source of her decision (ultimacy) and/or has genuine alternative options open to her (a version of volitional freedom). Consider Jan's decision to and act of recklessly running through a crowded group of pedestrians to get to the play on time: If determinism is true, then it follows that the event of deciding to run through the crowd had to occur at time t , say 6:45 pm on January 14, 2019, (and this is the crucial bit) holding fixed the laws of nature and the state of the world prior to her decision. So if Jan's behavior is undetermined, then instead of Jan's deciding to run through the crowd — what she actually did — something different could have occurred *at that exact time holding fixed the laws of nature and the state of the world prior to her decision*. For instance, perhaps Jan might have decided not to run through the crowd, kept deliberating, or even got knocked over by someone dressed in an Elmo costume. In any case, on an incompatibilist picture of decision-making things could have been different holding everything about the world fixed. (Recall that this is referred to above as *deep openness*.)

Here, however, some compatibilists and hard incompatibilists find fault with the incompatibilist strategy. The crux of the objection is that even if determinism doesn't secure an agent adequate control to act freely, indeterminism — the rejection of determinism — won't either. Why might indeterminism undermine the control required to act freely and be morally responsible? Here compatibilists have advanced versions of what's called the 'luck argument.' Crude versions charge that incompatibilist free decisions are too 'random' or luck-infested to give agents the control needed to be accountable for their actions. More carefully, on the contrastive explanation version, one might argue that incompatibilism leaves us with no satisfying explanation of why Jan ran through the crowd, as opposed to decided to not or kept deliberating. After all, all of the reasons that Jan considered in the actual case (e.g., that she believes that she would be

late to the show if she did not hurry, that she believes that she paid good money for the tickets, that she desires to be on time and not ‘waste’ money, that she desires not hurt someone by knocking them over, etc.) she considers in the cases of other possible outcomes mentioned. In fact everything about Jan’s whole psychology and indeed the world are the same in the counterfactual case where she does not decide to run through the crowd, per deep openness. Therefore, there is nothing we can point to at all that explains why Jan decided to run through the crowd (in the actual case) as opposed to not (in the other possible cases). Everything happens the same prior to each outcome.

Furthermore, suppose we make Jan’s decision or action a significantly morally weighty one: Jan decides to and does shove someone into oncoming traffic. According to another version of the broad luck problem, if we consider that everything about the universe could be exactly the same, including everything Jan considers before she decides to and does shove someone into traffic, it isn’t fair to blame her for doing so (see, e.g., Waller 1988). The same past exactly could have led to a morally better outcome – for example, she could have decided to and could have helped an elderly pedestrian cross the street instead. Hence, action control on the incompatibilist picture seems not satisfying in terms of an explanation of how an agent controls her actions or whether it’s justified to blame agents at all. For an illuminating account of the contemporary history of the luck problem and how the objection has been discussed with regard to Kane’s event-causal libertarian view, see Mele 2014d and Kane 2014. For a discussion of the problem as applied to agent-causal libertarian, see Mele 2006; 2015a.

The luck problem has been around in some form for the lengthy duration of philosophy of free will, but one major voice in the contemporary discussion of the issue is Alfred Mele. Mele argues that even if the differences between agents’ actual decisions and other possible outcomes is “just a matter of luck” on an incompatibilist picture, still such agents can act freely – perform actions that are ‘up to them’ in some significant sense (Mele 2010f; Mele 2013k). Mele outlines the general worry for incompatibilism without appeal to any notion of chance, luck, or explanations of action. He terms this version of the objection the “Continuation Problem” (Mele 2010f; 2013k; Mele 2015d; 2017d, Ch. 6). Going back to Jan and her decision to run through the crowd of pedestrians at 6:45 pm, we will say that the genuine possibilities – those alternative ways she could have exercised agency at that time, such as deciding not to run through the crowd – are *continuations* of some (long) common past. Mele then compares these continuations with the possible outcomes, or continuations, of a genuinely indeterministic random number generator that at a time yields the number 17. It could have, given the exact same state, yielded the number 68 or 45, etc., at that exact time. These are all possible continuations of the (long) common past state of the generator. (In other versions Mele compares the continuations of an agent’s indeterministic decision to the outcomes of a spin of a “tiny indeterministic neural roulette wheel” (see, e.g., Mele 2013k). The challenge for the incompatibilist is as follows: insofar as the possible continuations of the past are similar to the possible continuations of the indeterministic generator in nature of origin, then how can incompatibilist free agents exercise the appropriate control over their actions to be acting freely and held accountable? (For other more general audience retellings of the continuation problem, including one in dialogue form, see Mele

2014b, Ch. 3, and Mele 2011b).¹⁷ Both compatibilists and libertarians have offered responses on behalf of incompatibilism to the luck problem or a refined version such as the continuation problem (see, e.g., Mele and Kearns 2014 for a back and forth on the issue). As noted, Mele himself argues that there is an acceptable libertarian account of free will that can address this challenge (Mele 2006; Mele 2017d; Mele 2017c).

Recently, Levy (2011) and Perez de Calleja (2014) have proposed a similar luck problem for compatibilist accounts of free will. Think back to Jan’s decision to run through the crowd in a deterministic world. Briefly, Perez de Calleja (2014) argues that there are relevant possible scenarios (here philosophers sometimes speak in terms of possible worlds) in which the agent (e.g., Jan) considered all of the same reasons, was in the same mood, and so on, but she decided differently due to some difference in the broader state of the world or difference in the laws of nature. Perez de Calleja argues that it’s lucky (bad luck, in this case) that the agent acts as she does (e.g., Jan runs through the crowd and causes the elderly pedestrian an injury). After all, the difference in these worlds is not due to differences in the agent’s reasons and broader motivation and background conditions. Hence, compatibilist accounts of free will face a challenge from luck as well. Mele (2015c) addresses this challenge from luck to compatibilism, arguing that such a challenge, although different in important respects from the luck problem for libertarianism, should at the very least motivate compatibilist to offer strong revisions to their sufficient conditions for free will.

Recall that some incompatibilists argue that thinking of free actions as just events caused in the same way as other natural events (e.g., formation of a hurricane, a meteor shower) is to miss what is distinctively *agential* about free will. Hence they argue for agent powers as the control required for free will – an agent (not some collection of events, like consideration of reasons) brings about a free decision or action. Some of these agent-causal incompatibilists are motivated to do so because of what’s called “the disappearing agent problem” (for a recent exposition of, see Pereboom 2014). The disappearing agent problem is meant to be a kind of luck-based objection to event-causal libertarians. Proponents of this objection argue that if we accept that indeterminism is required for free will, then it won’t help to propose that free decisions and actions are those that are caused by the agent’s mental states (e.g., Jan’s desire to be on time, Jan’s belief that she will be late unless she runs). Decisions and actions caused indeterministically by such events leave no room for the agent to “settle” which way she decides. Why? Let’s say Jan’s decision to run through the crowd was indeterministically caused. Prior to her decision to run, her considerations in favor of doing so (her belief that she would be late, her desire not to be late, etc.) and her considerations against her doing so (her belief that someone might get hurt, her desire not to cause others injury) each made her 50% likely to decide one way or the other. She does decide to run through the crowd. But in what sense did she, the agent, settle on that decision? Hence, perhaps what we need, theorists such as Pereboom argue, is that the agent herself settles the matter – brings about the decision to run through the crowd.

¹⁷ For a German-language paper on the continuation problem, see Mele 2012n.

Mele (2017d; 2017b; 2017a) presses Pereboom to be clear on what he means by “settles.” Mele charitably posits that perhaps Pereboom means complete control, where the former Jan, the one deciding on an event-causal picture, lacks complete control, whereas the latter Jan, the one deciding on the agent-causal picture does not. Mele argues that, first, we do not know what complete control amounts to and, second, that on some readings of what it might mean, complete control does not appear to be required for free will and moral responsibility. Mele gives the example of a basketball player who has a 99% chance of sinking a free throw and a 1% chance of breaking down and being unable to sink the free throw. Here the basketball player may be said to lack complete control over whether he sinks the shot, but it doesn’t seem inappropriate to say that he freely sank the shot. Further, Mele argues, even if we degrade his chances of making the shot to 90% (and his chance of breaking down to 10%), it still doesn’t seem appropriate to say that he freely made the shot. Mele concludes that even if indeterministic decisions and actions are luck-infested, there are still adequate event-causal libertarians accounts of free will, one of which he proposes.¹⁸

Pereboom is a free will skeptic who is motivated primarily by ultimacy concerns: He argues that in order to have free will, we’d need to be agent-causes of our decision, but we aren’t. Philosophers have offered a number of both philosophical and data-driven responses to these positions (e.g., Sartorio (2013b)’s argument that prominent incompatibilist arguments for ultimacy are question-begging). Here we’ll focus on the empirically-informed responses: Nahmias and colleagues (see, e.g., Nahmias et al. 2007) argue that incompatibilists are committing a mistake when they judge that determinism is incompatible with free will. Here these folks are confusing determinism with bypassing – the idea that our mental life (desires, beliefs, intentions, etc.) aren’t among the causes of our behavior. Citing surveys in which non philosophers make judgments about agents in deterministic universes, Nahmias and colleagues support that when determinism is described in a way that makes it clear that mental life is in the causal loop, participants are much more likely to judge that determined agents act freely and are responsible. Hence, the real threat to free will isn’t determinism, but rather determinism plus bypassing. Roughly, we require that our mental states make a difference to how we decide and act.

Bjornsson (2015; Bjornsson and Pereboom 2014) agrees that incompatibilist judgments can be explained in a way that is consistent with the compatibilism being true. However, he disagrees that it is perceived bypassing that is motivating incompatibilism. First, he questions whether the wording of the bypassing statements in Nahmias and company’s studies is really being read the way it is meant to be read – that is, whether participants’ endorsement of the items implies that they endorse bypassing. Further studies fail to support the Bypassing Hypothesis (Bjornsson 2015). Second, Bjornsson (2015; Bjornsson and Persson 2012) offer a competing alternative explanation of incompatibilist judgments. The business of responsibility judgments is to offer explanations of behavior: “People take an agent to be morally responsible for an object only if a relevant motivational structure of the agent is taken to be part of a significant or salient explanation of that object” (Bjornsson and Pereboom 2014, p. 34). The role that mental states must play in the causal chain leading to an agent’s free action will be taken up further in next subsection.

¹⁸ Note, though, that Mele also proposes an adequate compatibilist account of free will, and so is not a compatibilist.

III.5 Further Metaphysical Issues: Causation, Time, and Omissions

Advancements in philosophy of free will extend beyond fresh perspectives on the well-established debate regarding the compatibility of determinism and free will. Here we'll look at a snapshot of recent philosophical work on the place of free agents in the causal and temporal order of the universe. That is, how free agency fits in with what we know about the nature of causation and time. Further, we'll consider metaphysical accounts of not acting, or omissions, and whether and how one is accountable for one's omissions.

Recall the consequence argument for the incompatibilism of free will and determinism. One assumption of this argument is the fixity of the past — that is, the relevant events of the past must have happened at a time when human agents had no control, such as a distant time before humans. This assumption plus other claims generates the conclusion that we have no control over the consequences of those events, what we do now. But, argues Sartorio (2015a), there are possible scenarios in which this is not the case — e.g., possibility of time travel in a deterministic world. In these scenarios human agents can affect the distant past. Hence, on Sartorio's view, the real proposed threat of determinism is not determinism simpliciter. Rather, the proposed threat is determinism plus the idea that all events can be causally explained in terms of natural events, and human agents aren't suitably part of the causal chains leading to action. In other words, human agents seem causally impotent (2015a, p. 264). Hence, Sartorio counsels that we need to turn our attention to the finer details of causation and the causal role that agents need to play to exercise free will. In her book *Causation and Free Will* (2016; 2015b) she offers a compatibilist proposal to this effect as well as a critical assessment of the incompatibilist literature. Specifically, she proposes the “actual-causal-sequence theory” of free will. This theory takes seriously that free will is grounded in the actual causal facts — how actions are actually caused. This carefully laid-out account includes a substantive view of causation, with the free actions involving human agents in some sense making a difference in the causal chain (Sartorio 2013a). Sartorio (2015c; 2016) argues that her compatibilist account is consistent with the condition that free agents are sensitive to and act for reasons.

As we've noted, the fixity of the past — the idea that the events of the distant past are fixed and so out of our control — is a prevalent assumption. In contrast, we tend to assume that the future is open — it is not yet fixed and can be affected by events now. For instance, both compatibilist and incompatibilists can agree that the events happening now, my typing these words, cause the words to then appear on my computer screen at a slightly later time. Rather, compatibilists and incompatibilist disagree on what needs to be the case prior to my acting (in the past) —whether a deterministic state of the world prior to and at the time of my action threatens free will. So, despite their differences, both set of theorists appear to assume an asymmetry of time and a corresponding unidirectionality of causal influence. The arrow of time points from the past to the future, and past events and the laws of nature influence and constrain what happens in the future; the past is fixed. Ismael, whose work bridges metaphysics and physics (see, e.g., Ismael 2016a; 2016b; 2017a; 2017b; 2017c), draws on physical theory to reject a number of these operative notions and attempt to restore free agency even if determinism is true. She argues that we should understand free agency as the exercise of control by agents — understood in terms of first-person perspectives of self-

governing physical systems – in a broader physical system (2016c, Ch. 1-3). Her case rests on overturning the assumption of a temporal asymmetry at the microphysical level, as well a physics-informed account of causation and laws of nature (2013; 2016c, Ch. 4-7).

Most of the discussion on free will and moral responsibility focuses on our practical decisions and intentional actions. But sometimes human agents fail to act – either intentionally or due to negligence or forgetfulness. When agents omit to act, an omission, are they responsible for omitting to act? What is an omission? For example, suppose that Alex goes out of town and asks Nancy, his neighbor, to water his plants while he’s gone. Nancy promises to do so, but she forgets. She omits to water the plants. The plants die. What exactly is Nancy’s not watering the plants? Is Nancy morally responsible for not watering the plants? Should we blame her? But plenty others failed to water the plants, including Queen Elizabeth II. Are they too to blame for the plants dying? Perhaps it’s tempting to say that of course Queen Elizabeth II isn’t to blame – she didn’t promise to water the plants. But trickier and real life cases abound – e.g., the construction worker who negligently throws the heavy slate material off of the roof, omitting to check for passersby beneath, thereby killing someone below. Insofar as his dropping the heavy slate results from something he omitted to do, his responsibility for the consequences of his omission are seemingly tied to his responsibility for his omission – a non-action. Clarke (2015; 2017) takes on these thorny moral and legal issues of negligent action and failures to act, given culpable or non-culpable ignorance. Clarke (2014b) gives an account of what omissions are, rejecting the view that omissions are merely some kind of (negative or positive) action. He explores how intentional omissions, on this view, are linked to intentions, and proposes conditions under which one can be morally responsible for omitting to act (2014a; 2014b). He argues moral responsibility for omissions is closely tied to conditions for criminal responsibility for failures of acting (2014b).

IV. NEUROSCIENTIFIC FINDINGS: HOW DO HUMANS ACTUALLY EXERCISE AGENCY?

We’ve mentioned that metaphysical determinism is consistent with physicalism, the idea that all facts are fixed by the physical facts about the world (e.g., those specified by a complete physics or more broadly the complete physical sciences). Some theorists worry about the standing of free agents in a purely physical, mechanistic world. It may well be palatable to understand the occurrence of events like fires or snowstorms or the actions of animals in naturalistic terms, but we might intuitively feel that free decisions and actions, and so the exercise of free will, are special in some way. Specifically, exercise of free will is closely linked with exercise of the mind and brain and our mental states. For instance, it is often assumed that an agent must make some or all of her decisions consciously if she is to decide and act freely. We certainly experience ourselves as consciously deciding what to do at least some of the time. This requirement of conscious mental states is closely tied to the focus on the deliberative process of considering and weighing reasons for performing one action versus alternative actions in philosophical free will accounts. Hence, one question is whether we can exercise free will in the face of potential neurological determinism:

Neurological determinism: The initiator of action, decisions to act, and overt bodily movements are the product solely of neurological (and other physiological) processes and events.

That is, our decisions and bodily movements are the effects of neurological and physiological event-causes in our bodies.

Neurological determinism by itself might be underwhelming as a threat to free will for some: Of course, one could reply, it's just obvious that our brain and body are causally involved in our decision-making and action production. But if one conjoins neurological determinism, so characterized, with one or more of the following, the situation may appear bleaker for free will:

Dualist threat: Free will requires a non-physical event of conscious decision among the initiators of action, but the initiator of action, decisions to act, are neural events.

Here dualism is a distinction in kind between mental and physical properties (or substances). That is, the mental (e.g., consciousness of pain or sensory experiences) are different in kind from the physical (e.g., firing of such and such neurons in this brain area). Alternatively, we could pair neurological determinism with the following:

Epiphenomenalist threat: Free will requires that conscious decisions are among the causes of action, but conscious decisions don't cause subsequent action.

Epiphenomenalism is the view, more broadly, that mental events are caused by physical events but that mental events themselves don't have any physical effects — don't cause any physical events. In this context, regardless of whether conscious decisions are caused by neural processes and events, they aren't part of the "chain" of events that leads to the efficacious decision and action. Our experience of ourselves as first consciously deciding to act and then acting *as a result of that conscious decision* is mistaken. (For a similar, particularly illuminating breakdown of the proposed scientific threats to free and responsible agency, see Nahmias 2014.)

Neuroscientists have taken up this question of how our overt bodily movements are produced and what role, if any, conscious decisions and intentions, play in producing our movements. *Section IV* will canvass recent empirical work by neuroscientists on the reality of our human agency and whether we exercise the kinds of control required to decide and act freely. Philosophical reflections on the significance of those empirical studies for our view of ourselves as free agents will also be addressed; Issues of neurological determinism have been raised as potential threats for both moral and criminal responsibility. *Section V* will examine empirical work by psychologists more broadly on the reality of human agency — what capacities for control do we in fact have, apart from any issues of metaphysical or neurological determinism. The question of our actual psychological capacities as human agents is critical to understanding whether we measure up to the image of our human nature as self-governing beings who obtain our goals via planning and self-control. And, as Mele, Vohs, and Baumeister (2010) helpfully note, issues of *how* consciousness, choice, and intentional action work is itself a substantive question, quite apart from the question of whether we actually exercise free and responsible agency.

The literature on neuroscience of voluntary agency, including action initiation, intention, and sense of agency is vast. Here we'll focus on research in neuroscience taken to bear upon questions of free will (For a concise, informative summary and significance of the broader neuroscientific work on voluntary action, see Roskies 2013. For a helpful lexicon of what many of the terms in the science of free will debate mean, at least as used in the work cited here, see Haggard et al. 2015). Glannon (2015) helpfully orients us when it comes to the neuroscience of free will: Whereas the metaphysical debate regarding free will often focuses on external impediments or enhancements to human agents' exercise of free will (e.g., metaphysical determinism or indeterminism, manipulation), neuroscientific research speaks to potential *internal* impediments to or internal mechanisms of agents' free decision-making.

The neuroscience of free will owes much of its motivation to the influential Libet paradigm. In 1983, Libet and his colleagues pioneered neuroscientific studies of action to test the widely accepted claim, introduced above, that conscious intentions¹⁹ initiate or cause action. Libet's work built upon significant advancements in neuroscience of agency: Earlier studies had established, using electroencephalography (EEG), that a slow negative buildup of brain activity in the supplementary motor area (SMA) and pre-SMA, termed the 'Bereitschaftspotential' or Readiness Potential (RP), preceded simple movements in the laboratory (Kornhuber and Deecke 1965). The RP was taken to be preparatory activity for action. Further, Libet, Wright, and Gleason (1982) extended these results to show that the RP preceded simple endogenously generated actions, such as wrist and finger flexes. Here 'endogenously generated' refers to actions for which the cue for acting does not come from the experimenter or other external stimuli but from within the participant. For instance, in 'spontaneous' trials, the participant decides on the spot when to flex within a preset window of time. One standard assumption about endogenously-generated actions, prior to Libet and in line with commonsense, is that the participant consciously wills or decides when to act, and that that volition or decision initiates the preparatory processes leading to action. Hence, Libet and his colleagues (1983) sought to determine when the will or decision comes on the scene in relation to the RP and movement, where by will or decision the researchers took this broadly to refer to the disjunction of intention, urge, or desire to flex.

Libet and his colleagues found that when participants were asked to perform a series of endogenously-generated spontaneous wrist flexes as they felt the desire or intention to do so and report the timing of their awareness of their intention to flex, EEG readings showed brain preparation for the flexing, the RP, prior to the participants' reported awareness of their intention to flex (1983). Participants completed dozens of trials, or task time intervals in which they made a movement. The beginning of each trial was signaled by a beep. Post flexing, participants reported the time of their awareness of their intention to flex in each trial by indicating where a moving spot was on a modified clock face when they were first aware of their intention to flex. This modified clock -- now referred to as a Libet clock -- makes a complete revolution every 2.56 seconds. Averaging over 40 trials, the average time of RP ramp up for spontaneous

¹⁹ Libet's studies concern proximal intentions to act, intentions to act now. Proximal intentions can be distinguished from distal intentions to act, intentions to act sometime in the future (Mele 2009, 10). A similar distinction is also made by Bratman (1984) (present-directed intentions and future directed intentions) and Pacherie (2006) (P-intentions and F-intentions).

trials was about a half second (517 ms) before recorded movement. In contrast, the average reported time of first awareness of intention to move, labeled 'W,' occurred about 200 ms (206 ms) prior to movement ('M'). Libet and colleagues found, then, that the RP preceded the reported time of conscious intention to move by a third of a second.

Libet and co-authors took these findings to indicate that unconscious brain processes decide (form an intention) to flex, with conscious intentions to flex temporally later and inefficacious (1983; Libet 1985). This conclusion rests on the assumptions that the RP represents neural preparation for action and that W is an accurate measure of the initial presence of a conscious intention. (For a discussion of whether W, an average of reported times, accurately reflects the actual onset of a conscious intention, see Mele 2010a; Mele 2010c; Mele 2010g; Mele 2012i.) On the basis of his results some theorists – including Libet himself (e.g., 1999) – have concluded that our conscious intentions never cause our overt actions (see also Wegner 2002). Libet reasoned that if an RP is present for all overt movements and W appears after the onset of the RP in his studies, then RPs occur prior to conscious intentions and initiate action for all intentional actions, even complex ones (Libet 2005, 560).

Several philosophers and scientists, Libet included, have applied Libet et al. (1983)'s conclusions about the role of conscious intentions in action production to argue for free will skepticism. The thought is that if our conscious intentions never cause our actions (epiphenomenalism), then we lack the control required to act freely and to be morally responsible for our actions. The experience of being an agent might suggest that our conscious intentions cause our actions and hence seem to support our intuitions of free will, but our experience is mistaken. Some have labelled this view "willusionism" (e.g., Nahmias 2012). For instance, Libet has argued that there is no free will, only 'free won't.' That is, no one is morally responsible for their intentions to move, as such items are acquired prior to awareness, but agents can override their unconsciously acquired intentions and hence prevent themselves from so moving (Libet 2005). If agents possess this capacity then individuals can be held responsible for their bodily actions (1999, 54). For a critical empirical study of Libet's 'free won't' or veto power, see Hallett (2011).

To situate Libet-inspired free will skepticism, it's helpful to understand Libet's free will skepticism. Libet was a neuroscience, but he hints at a particular metaphysical account of free will in parts of his work. Libet's writing supports, at least in places, that he was an incompatibilist and a dualist about conscious intentions and related neural activity. For instance, he argues that if conscious intentions were efficacious in action production, then these intentions should "command the brain to perform the intended act" (Libet 2005, 553). Elsewhere he notes the following:

Are freely voluntary acts subject to macro-deterministic laws or can they appear without such constraints, nondetermined by natural laws and 'truly free'? (2005, 551)

I myself proposed an experimental design that could test whether conscious will could influence nerve cell activities in the brain, doing so via a putative 'conscious mental field' that could act without any neural connections as the mediators (Libet 1994) (Libet 1999, 56).

Hence, Libet's results are consistent with both the dualist threat and the epiphenomenalist threat introduced above. That is, if the **RP** is the biomarker of the neural decision to act and that neural decision causes action, the initiator of action is a physical event and conscious decisions may be out of the loop. The Libet paradigm is a thriving one, and his original findings have been replicated and extended using multiple neuroimaging technologies. Below are neuroscientific work on human decision-making and action production that have been broadly inspired by Libet's original study. Due to the vastness of this literature, only select recent Libet-inspired studies will be addressed here. For a clear and up to date sketch of methodological criticisms of the original Libet et al. 1983 study, see Maoz et al. (2014).

IV.1 Neural Signature of Intention to Action?

Several studies have questioned the link between **RP** and conscious intentions in Libet studies. For instance, one might theorize that the **RP** as preparation for action is a sufficient cause of **W** (assumed time of conscious intention to act): If an agent's brain exhibits an **RP** prior to voluntary movement and that **RP** is intention-specific, then a conscious intention will result. One way to test whether **RPs** are related to conscious intentions is to see if there is a relationship between the time of onset, or initial ramp up, of the **RP** and **W**, the average reported time of first awareness of intention to act. An earlier experiment by Haggard and Eimer (1999) supported that the timing of the **RP** is not correlated with **W** time, but that timing of the **LRP** is: **LRP** is the lateralized motor activity. When an agent makes a unilateral movement (e.g., left movement or right movement), the **LRP** represents the difference in electrical potential to the left and right of the vertex prior to that movement. Recent work has attempted to replicate Haggard and Eimer's findings (Schlegel et al. 2013; Alexander et al. 2015). The results were mixed: Although they replicated the finding that **RP** is not correlated with **W**, they also found no covariance of the **LRP** and **W**, in contrast to Haggard and Eimer (1999).

Even if the timing of the **RP** ramp up and reported time of conscious intention are not correlated, one might still expect that **RP** is sufficient for the presence of a conscious intention to act. That is, if an agent exhibits an **RP** prior to voluntary movement and the **RP** is related to an intention to act, then the agent will have *some* conscious intention related to her so acting (at some time prior to movement). Even this assumption, taken for granted in earlier Libet-inspired work, has been thrown into question with new empirical evidence: Schlegel and colleagues (2015; forthcoming) examined what the biomarkers of agency look like when hypnotized agents act. The researchers compared participants who performed a series of simple movement tasks — a cued stress ball squeeze with the left or right hand — either under hypnotic suggestion or under normal instruction. The hypnotized participants had no awareness of any intention — no conscious intention — to squeeze the ball (and were told a cover story for the “twitching” of their hands). Schlegel and collaborators found no significant differences in both **RP** and **LRP** profiles for the two conditions. These findings question whether there is any connection between the **RP** and **W**.

Given this evidence, perhaps neuroscientists ought to be skeptical whether the **RP** cognitively represents an intention to act. But even if the **RP** isn't the biomarker of intention to act, it may well still be the biomarker of preparation to act or a bias towards a particular action (Mele 2009; Mele 2012i). Here we can distinguish the two as follows: an intention to act implies a *settled* plan to act *in a particular way* (e.g.,

with the left hand, at this time, etc.). In contrast, preparation to act might involve attention to task-related aspects; bias can be understood as an inclination, but not settledness, to perform a particular action (e.g., left button press) over others (e.g., right button press). How does this proposed weaker relation between RP and action fare? Not well. Several studies provide evidence that RP is not action-specific cognitive activity. Alexander et al. 2015; 2016 support that the RP isn't representative of movement-specific activity, even preparation to move, but that the LRP may be: When participants engage in movement tasks (e.g., cued button pushes) versus non-movement tasks (e.g., mental addition and subtraction), the analogue of the RP is present for both kinds of tasks, but the analogue of the LRP is only present for motor tasks. Perhaps, then, the RP represents cognitive preparation of a more general nature (for any kind of task). This leaves open the LRP is movement-specific.

Some neuroscientists have studied the neurological signature of intention to act using nonhuman primate populations. Using these subject pools, researchers are permitted to conduct research via experimental single cell recordings, and thus can investigate the neural markers of decision and agency at the level of neural activity of individual neurons. Such single cell recordings and intracranial work are permitted only in clinical populations in humans, and here the electrode placement is pre-specified by the patient's clinical needs (e.g., brain areas affected by epilepsy). Roskies (Shadlen and Roskies 2012; 2014a; 2015) details work by Shadlen and colleagues, who trained monkeys to perform a visual discrimination task. Monkeys were to watch a screen in which a group of dots were moving either to the left or right. The dot array stimuli differed on the dimension of coherence of the dot motion, with 0% coherence meaning 50% of the dots moved to the left and 50% of the dots moved to the right and 100% coherence meaning that 100% of the dots moved in one direction (left or right). The monkeys reported whether over the trial, the dots moved net motion left or right. They reported their decision via an eye movement, or eye saccade, to the left or right. Given the measured firing rates of neurons during the task, the researchers found that populations of neurons in the middle temporal (MT) and medial superior temporal (MST) visual areas encoded, or represented, the stimulus motion in some region of visual space. In contrast, populations of neurons in the lateral intraparietal area (LIP) integrated this sensory information from MT and MST, and the monkey produced an eye movement to report a decision when the firing rates of the LIP neurons reached a certain threshold. Hence, Shadlen and Roskies argue that the activity of LIP neurons encode, or represent, accumulating evidence for each potential decision (left or right) and further the decision itself to produce an eye saccade. Roskies (2014a) stops short of endorsing that this neural representation of motor decision has evidential bearing on whether the brain is a deterministic system or not (see section IV.4). Although it's tempting to read this work as supporting a picture of decision-making as one in which the activity of populations of neurons represents the consideration of reasons and moment of decision (settledness on a course of action), Roskies (2014b) cautions that we would need substantial inferences and further evidence to do so. However, this kind of work is indicative in general of a growing trend toward neuroprediction of decisions, to which we will now turn.

IV.2 Neuroprediction of Decision

One of the worrisome aspects of the Libet study, on the face of it, is that neuroscientists are able to find brain activity relevant for subsequent action prior to the agent's reported time of being aware of decide to

act. This finding opens the door for neuroprediction of decision and action — scientists being able to predict, with varying degrees of accuracy, how one will act prior to one’s reported awareness of plans to act in that way. Using more recent neurotechnology, neuroscientists have been working to “encode intention” using brain-computer interfaces. Using EEG, Salvaris and Haggard (2014) trained a computational model that could predict with over 83% accuracy (for some participants) whether a participant was about to press a left button or a right button. Notably, this neuroprediction was greater for instructed actions (e.g., push left on the GO signal) than for free actions (e.g., choose to push left or right on the Go signal). Experimenters could predict with a mean of 82% accuracy (SD of 6%) the direction of button press in the cued condition versus a mean of 75% accuracy (SD of 5%) the direction of button press in the free condition. Maoz et al. (2012) had similar accuracy rates of prediction of left or right choice using intracranial recording. Human epilepsy patients, with electrodes implanted based on clinical need, played a game for monetary reward and loss in which the winner was rewarded for holding up a different hand (left or right) from her opponent but penalized for holding up the same hand as her opponent. Acting as the opponent and using a model trained on the activity recorded from the electrodes, the neuroscientists were able to predict, with 83% accuracy over each trial (here each one-time face-off of players), which hand the patient would raise.

IV. 3 Neuroscience of ‘Free Won’t

As noted, Libet argued that although agents don’t possess free will in the sense of conscious intentions initiating actions, agents still have a ‘free won’t.’ That is, agents can consciously override, within a small window of time (roughly after the ramp of the RP but no later than 50 ms prior to action) the initiated preparation to move. Agents can consciously decide not to move; This is referred to as ‘intentional inhibition.’ Intentional inhibition is tricky to study in the lab for numerous reasons (see Mele 2009, p. 54), but one way to do so is with Go/No Go tasks. The participant prepares to move, but if he is given the No Go signal, he doesn’t move. Of course, if he is given the Go signal, he moves. Libet (1985) thought free won’t to be the last refuge of efficacious conscious willing in action production. However, some neuroscientists have found evidence that unconscious processes, such as subliminal primes, can significantly bias intentional inhibition (Parkinson and Haggard 2014).

Some have interpreted these subliminal primes as setting an intention to so act, with the participant changing that intention or decision (consciously, or more likely, not) later in the trial (Furstenberg et al. 2013; 2014; Ramaswamy et al. 2013). For instance, Furstenberg and colleagues (2015a; 2015b) extended the effect of subliminal prime on decision to “picking” scenarios, scenarios in which agents have no reason to favor any of the action options. They found that even in an arbitrary choice task (here the choice between pushing a left or right button), participants were biased toward action content of the subliminal prime. That is, the chosen movement was ‘congruent’ with the prime (e.g., prior arrow primes for a left button press and the participant presses the left button) (57.6% of the choice trials). Moreover, using the EEG data and in particular the LRP, they argued that in incongruent choice trials — trials where the participant chose to act contra to the prime (e.g., prior arrow primes for a left button press and the participant presses the right button), there was a neural signature indicative of the brain’s changing intentions (e.g., from left press to right press). Hence the implication drawn is that the participant can

override an intention to act with a conflicting intention to act, albeit not consciously, at least for cued actions and even arbitrary choices.

One must caution not to read this result as too sensational, as the authors specify their use of intention as inclusive of “immediate nonconscious and nonexecuted neural preparations to perform a motor plan and not necessarily an explicit control mechanism” (Furstenberg et al. 2015b). This is consistent with the earlier stated claim that the LRP is movement-specific neural preparation notably short of settledness to so move. Elsewhere, however, Furstenberg makes a stronger claim and argues for a conceptual framework in which we should understand these prior vetoed intentions as “non-executed, nonconscious intentions” (2014). This would deny intentions, conceptually and empirically, a role in the actual action production process, but recover some conceptual and empirical role for practical intentions, neurally realized, as defeasible plans for action.

Notably, philosophers such as Mele (2009; 2013l) have been critical of the empirical paradigms of vetoing. First, Mele argues with reference to Libet (1985)’s studies of vetoing, if one is told to prepare to move at a time knowing that one isn’t going to move at that time, then one does not intend to move. Second, Mele questions whether the intentional inhibition paradigm – preparing to move and then either vetoing on a No Go signal or moving on a Go signal – used in more recent experiments is really indicative of an intention to move versus either a mere simulation of motor intention or general task preparation. Mele (2013l) elucidates some strategies that participants would have to follow during the intentional inhibition task in order to attempt to veto an intention to move.

IV.4 Significance for free and morally appraisable agency

Not disregarding its status as a fruitful paradigm for neuroscientific knowledge of human agency, the Libet study and related empirical work have also attracted much criticism. Some scientists endorse the conclusion that the Libet study poses some combination of a deterministic, dualist, or epiphenalist threat to the existence of free will, while others – particularly philosophers of action – are skeptics of Libet-inspired free will skepticism. Koch (2012), a neuroscientist, like Libet endorses some version of a dualist free will view, which he develops and defends in the context of recent empirical findings. Notably, philosophers have not been uncritical of the high bar set for free will by scientists (see, e.g., Glannon 2012c). In particular, endorsing a dualist, libertarian view of free will strongly impacts their interpretation of the Libet-inspired work, and so the claims that are disseminated to the public (Mele 2009; Mele 2012k; Nahmias 2012; Nahmias 2014). Other scientists, such as Tse (2013a; 2013b), are more critical of Libet’s findings. Tse (2013a) defends in detail a reductionist account of mental causation, consciousness, and indeterminism which draws heavily upon both cellular and systems neuroscience especially work in vision science. Vision science is a crucial area for understanding neural correlates of consciousness (e.g., Kohler et al. 2015; Reavis et al. 2015; Reavis et al. 2016; Kohler et al. 2017). Here we will canvass philosophers’ critical reactions to neuroscience of free will and neuroscientists’ claims about free agency.

IV.4.1 Criticism of the dualist threat

Irrespective of individual scientists' own views on free will, in their writing scientists often take for granted that the majority of people, particularly nonexperts, subscribed to some form of a soul-based libertarian account of free will.²⁰ This is on par with the kind of free will position that Libet hints at in his work mentioned above. Their next move, typically, is to evoke the dualist threat to free will from scientific work such as Libet's. Mudrik and Maoz (2014) refer to this as "neuroscience's closet dualism" (or the "dual subject fallacy") and counsel that speaking of the agent and brain in such terms at best is obfuscating the empirical research and at worst dangerous, both for replicability and public perception.

In response to those who do argue that the pressing threat of Libet's findings is the dualist threat, Mele (2012g; 2014f) disputes that non philosophers are typically dualist libertarians about free will in the way that Libet and other neuroscientists assume they are. Here he cites empirical work by philosophers and psychologists, including a study of his own, that supports that a large proportion of non philosophers are compatibilist, who, further, find free will and responsibility compatible with decisions being brain processes. Rather, when asked to detail instances of free or unfree actions that they've performed, participants tend to discuss morally positive actions that involve goal achievement and prior deliberation (Stillman, Baumeister, and Mele 2011). Further, recent work suggests that a lack of constraint, rather than the presence of a soul, is more central to nonphilosophers' conception of free will (Vonasch et al. 2018). These elements of choice and meaningful agency, on the face of it, seem compatible with a deterministic world.

Nadelhoffer (2014), a free will skeptic on grounds independent of neuroscientific determinism, replies to Mele (2014f) and argues that it is not clear, empirically speaking, where nonphilosophers stand on the relationship between determinism, physicalism, and free will. He outlines results from the "Free Will Inventory Project," a co-developed new measure of nonphilosophers' endorsement of statements regarding free will, for support (see also Schooler et al. 2015 and Nadelhoffer et al. 2014) for more on the development of the Free Will Inventory). Specifically, he argues that, in a survey of 330 participants from the general population, there is strong agreement with statements expressing dualism of the nonphysical soul or mind from the physical human body and brain. Further, when asked to endorse views directly (not in the form of judgments about hypothetical particular agents' actions), the majority of participants agreed with items expressing that free will requires deep openness in terms of an ability to choose differently and be the source of one's action. However, the results also indicated that the role of having a nonphysical soul or mind for free and responsible agency on a folk (nonphilosopher) theory is less clear: Similar proportions of participants agreed, disagreed, or neither agreed nor disagreed that souls were required for free will and for moral responsibility. Hence, Nadelhoffer concludes that the empirical studies of nonphilosophers' view of free will don't unequivocally support either the neuroscientists' or Mele's contentions about folk views.

²⁰ See Mudrik and Maoz (2014) and Mele (2014f) for a variety of scientific excerpts guilty of this practice.

In a response to Nadelhoffer's reply, Mele (2014e) answers that it is in keeping with both his results and Nadelhoffer's that most nonphilosophers can believe in nonphysical souls or minds without believing that humans' exercising free will requires humans having such nonphysical features. Indeed, Mele notes, perhaps many non philosophers hold that human agents having other capacities is enough for those agents to have free will, even if humans also have souls or immaterial minds. It should be noted that the literature on whether nonphilosophers believe in souls doesn't seem to be univocal – Nahmias (2014) reports that a sizable group of philosophers as well as nonphilosophers (in his studies) don't believe that humans need to possess a soul or immaterial mind in order to act freely. (See also Nahmias and Thompson 2014 for a discussion of free will and naturalism among the folk.)

In fact, work by Malle and colleagues suggest that non philosophers have a low bar for responsible agency compared to the heavy metaphysical criteria in philosophical accounts of free will and moral responsibility (Roskies and Malle 2013; Monroe and Malle 2015; Monroe 2014). In one study, when participants were asked the open-ended question "explain what it means to have free will," only one of 180 participants gave a metaphysically demanding answer (e.g., no external influence on decisions). Rather, the majority of participants cited psychological capacities such as the ability to make decisions and acting as you want without constraints (Monroe and Malle 2010). While this outlined study was conducted with university students, a related study among general community members found similar results, with additional emphasis on lack of constraints and deliberation. When asked what it would take to build a free agent, community participants most frequent response involved building in a capacity for choice. In a further innovative study, Monroe, Dillon, and Malle (2014) asked participants to make judgments (e.g., Is the agent to blame?) of various kinds of agents. Participants read about one of the following agents: a neurotypical human agent, a weak-willed agent (recall section III – one who acts contrary to her judgment of what is best to do), a cyborg, artificial intelligence in a human body, or an advanced robot. Echoing results from Mele (2014e) and Nadelhoffer (2014), participants' judgments did not reflect a requirement of a soul for free will and moral responsibility. The major factor driving judgments of control and responsibility appears to be whether the agent had the capacity for choice and intentional action.

Recently, Ismael has argued that much of this talk, by both non philosophers and philosophers, of locating conscious will neurally is misguided. Ismael (2015; 2016c, Ch. 1-3) brings to bear work on self-organizing versus self-governing systems to argue that the "self" or agent from decisions are made should not be strictly speaking be thought of in terms of the physical concrete brain. Rather, the brain-body system has a self-governing structure with a corresponding reflective standpoint through which coordination of the brain-body system is to be understood. Hence, the agent or self in decision-making is to be understood as the reflective perspective or point of view. She argues that those who look for the self at a neural level won't find it, but it would similarly be a mistake to discount the self as nonexistent.

IV.4.2 Criticism of the epiphenomenalist threat

Several philosophers, using the field's toolkit of conceptual distinction in combination with critical assessment of the empirical findings, have offered strong objections to the neurological determinist threats to free will. For instance, Mele (2009; 2010a; 2010b; Mele 2010e; Mele 2012d; Mele 2012e; Mele 2012m;

Mele 2016c) has developed a detailed argument that, first, the Libet studies have not demonstrated that conscious intentions are causally inefficacious and, second, that research supports that some conscious intentions – specifically intentions to perform actions later at particular time or place (e.g., Gollwitzer 1993)— are casually efficacious. That is, Mele provides grounds for skepticism regarding the epiphenomenalism threat. One strand of his case against epiphenomenalist interpretations of the Libet results is the overlooked possibility that we can grant the existence of unconscious intentions and accept a timelag between the onset of intention and consciousness of an intention without embracing free will skepticism (Mele 2013f). For instance, in a recent elaboration of his response to the epiphenomenalist threat to free will from Libet results, Mele (2014f) argues that neither the compatibilists nor the event-causal libertarian should be worried: First, nothing in the Libet and related findings about RP (and relevant fMRI-detected brain activity) provide evidence that the neural correlate of conscious intention is out of the causal chain leading to action. If so, conscious decisions may indeed be part of the action-production loop via its grounding neural activity. But keeping the agent’s conscious mental life “in the loop” – having the causal chain run through conscious agency’s neural correlate – may be all that compatibilist accounts of free will require. (For a related recent paper on scientific epiphenomenalism and free will, see Mele 2018b.)

Of course, this path to explaining away the epiphenomenalist threat to free will may well depend on the details of how consciousness is realized in the brain. A number of plausible theories of consciousness are on offer. For instance, some recent empirical work supports the higher order theories of consciousness – that being conscious of a mental state is to have a mental state directed at, or about, that mental state of which one is aware (Lau and Rosenthal 2011). Levy (2014) agrees that whether agents are conscious is relevant for their moral responsibility status. However, he argues that the requisite consciousness for responsibility isn’t that agents must be aware of their mental states as a whole but only that they are aware of the content of those explicit attitudes that causally contribute to their behavior. Many of the objections detailed here against the original Libet studies and its earlier offshoots apply to more recent Libet-inspired studies of volition, including those using functional magnetic resonance imaging (fMRI), depth electrodes, and subdural electrodes (Mele 2018a). (For versions of Mele’s arguments mentioned here in a general audience format, see Mele 2014b; Mele 2014c.)

Of note, proposed empirical support for the epiphenomenalism threat isn’t restricted to the results of the Libet paradigm. Wegner (2004) outlines a theory of apparent mental causation, supported by a wealth of psychological studies, some akin to the Ouija Board effect. In effect, in the lab the experience of consciously willing can come apart from actual control over effects in the world. For instance, Wegner and Wheatley (1999) had pairs of participants and confederates (posing as participants) jointly operate a computer mouse and move a cursor over objects. When asked to what degree they “intended” to stop the cursor on various objects, participants reported a strong sense of willing even in the absence of any actual control over the cursor motion (as the confederate was solely in control). Mele (2012c) tackles the question of whether these kinds of results mean that conscious intentions (and their neural realizers) are never the cause of action. He answers in the negative: Such theories of apparent mental causation rely, much like Libet does, on a strong notion of what the mental will must be in order for it to be causally efficacious, and moreover assumes the all action must be produced in the same way. Neither assumption is adequately

supported. Further, Mele and Shepherd (2013) argue that even given the pessimistic wider set of studies from psychology on the influence of situational factors on our behavior, of which the agent is seemingly unaware, the results point to an optimistic outcome as well: With knowledge of how human agents actually make decisions, we as agents can adjust our behavioral strategies and so mitigate the effects of unconscious or situational factors that may contribute to morally bad decisions and actions.

Interestingly, Nahmias (2014) makes the case that the only possible scientific epiphenomenalist threat to free will would be if conscious agency (and its neural correlates) were bypassed – that is, not part of the causal chain leading to action. However, none of the extant studies he surveys support this occurring (2014). In fact, recent findings support that nonphilosophers are not bothered by the kind of neuroprediction that neuroscientists are aiming for in Libet-inspired studies: When asked about hypothetical cases in which agent’s decisions and actions are predicted perfectly by neuroscientists, the vast majority of participants still agreed that the agent acted freely and was responsible for her conduct (Nahmias and Thompson 2014). Of course, Nahmias notes, if one is an agent-casualist who takes free will to require ultimacy in the extreme sense – the agent determines her action apart from any influence including from her own brain, then science of agency does threaten free will (2012). Hence one’s answer to the criteria question for free will capacities is critical to understanding why some take neuroscience findings to be so problematic for the existence of human free will and others don’t (Nahmias 2011).

IV.4.3 Neuroscience and evidence for determinism

Returning to our focus on Libet-inspired work, several philosophers have also mounted skeptical arguments that neuroprediction of voluntary actions in the lab (see section III.2) provides an answer to whether the human brain and mind is a deterministic system. Mele argues that no empirical findings in the Libet paradigm have found brain activity prior to reports of conscious decision that predicts the kind (e.g., left or right button push) or time of action with *certainty*. (In fact, as described in III.2 above, neuropredictive accuracy has not been demonstrated above 83%.) Hence, *no determinants* of actions in the strong sense have been found. This leaves open the door for libertarian free will (Mele 2014f). Nahmias (2014) expands this criticism beyond Libet paradigms, and argues that the cognitive sciences, such as neuroscience, are not equipped to show that metaphysical determinism is the case for human decisions and actions. Moreover, he argues, some philosophers as well as a majority of non philosophers (in his studies) have compatibilist intuitions and judgments. Glannon (2015) provides further reason to be skeptical of strong claims of neurological determination of decisions and actions. Such claims are grounded in neuroimaging studies, but notably the kinds of information gathered from neuroimaging is merely correlative: Data such as the BOLD signals of fMRI (markers of changes of blood oxygenation as indicative of dynamic brain activity) are merely correlated with reported or task-relevant mental states (e.g., judgments about stimuli). Correlation is not sufficient for causal determination.

Notably, Roskies (2014a; 2014b) offers an even stronger claim against incompatibilist scientists who argue that the physical sciences provide evidence of determination of human decisions and actions: She argues that there is a critical distinction between metaphysical determinism (as defined in section III) and *apparent determinism*. We tend, she argues, to think of predictability as the marker of determinism, and

so unpredictability as the marker of indeterminism, but it may be the case that what we, given our current knowledge of the physical world, taken as unpredictable may yet be fundamentally, or metaphysically, predictable with complete knowledge. That is, apparently indeterministic systems can be metaphysically deterministic. (For instance, in Shadlen and Roskies 2012, the two co-authors disagree as to whether noise at the level of neural firing rates is supportive of metaphysical or merely apparent indeterminism. Roskies argues the latter.) Relatedly, it may be the case that what we experience as predictable physically at a macroscopic level is in fact fundamentally unpredictable at a microscopic level of analysis. Hence apparently deterministic systems can be metaphysically indeterministic. She then notes, first, that physics hasn't yet found a complete physical theory, so the verdict is out on whether the world is metaphysically deterministic or indeterministic. Furthermore, current neuroscience can only tell us about apparent determinism. The criteria for neuroscientific evidence of complete predictability, of the variety relevant for human free will, is quite demanding.

Moreover, Roskies and Nahmias (2017) argue against the position that even if Libet-inspired studies don't provide evidence of metaphysical determinism, they nevertheless could, in theory, provide support for a more restricted, or "local" determinism. Perhaps, then, we should cease to conduct the search for truths about free will in reference to the question of determinism, especially with regard to neuroscience of agency. Instead, we ought to look toward the mesh (or lack thereof) of psychological models of decision-making and the physiological underpinnings of our volitional capacities (Roskies 2012a). In "Libertarianism and Human Agency," Mele (2013b) does the important work of attempting an account of this interdisciplinary connection, which seeks to understand what potential evidence of indeterministic decision-making and action-production at the neural level would mean for the satisfaction of philosophical libertarian accounts of free will (those discussed in section III.4). More broadly, Roskies (2016a) is optimistic of the relevance of neuroscience to philosophy on certain fronts, including providing empirical evidence as potential counterexamples to philosophical claims.

It should be noted here, though, that not all empirical studies of nonphilosophers' judgments of free will support that non philosophers are compatibilists who aren't bothered by event-talk explanations of decision and action: Knobe (2013) argues, based on clever surveys about human and computer actions, that a sizable group of non philosophers are incompatibilists who favor a transcendental self view of the free agent. This kind of agent, he notes, is at odds with the scientific vision of decision- and action-production in terms of events and processes. Hence the empirical work is not univocal on the question of whether the folk embrace the challenge of neurological determinism to free will in one or more of its forms. Nevertheless, a more optimistic take-away may be to argue that recent advances in neuroscience may impact of view of ourselves as free agents by encouraging us to get rid of such inflated senses of free will required for responsibility in favor of more down-to-earth mechanism-friendly control capacities for responsible agency (Roskies 2013).

Even if, however, the case were clear that non philosophers don't take neurological determinism to threaten free will, as Mele (2014e) and Nahmias and Thompson (2014) have supported, other studies have suggested that non philosophers take neurological determinism to be threatening to some kinds of responsibility: When participants were primed to read either an explicit free will skeptic or free will neutral

excerpt from a prominent neuroscientists, those participants were significantly less inclined to recommend retributive legal punish (e.g., lengthy jail sentences) for a person who committed a violent crime (Shariff et al. 2014). In these cases there was meant to be no benefit of the punishment for the convicted individual or society, as he had received (participants were told) a “nearly 100% effective rehabilitation treatment” (p. 3). Hence any prison sentence would be, presumably, due to the judged desert of punishment, given the nature of the crime.

IV.4.4 Neuroscience and relevance for responsible agency

One might worry that a Libetian spontaneous wrist or finger flex is too simple or arbitrary of an action to generalize the results of the studies to the kinds of decisions that matter, those for which we are held responsible (Waller 2012; Waller forthcoming; Mele 2013i). Furthermore, past replications or extensions of the original Libet study suffer from similarly problematic operationalization of relevant practical decision and action – e.g., when to press a button (Banks and Isham 2009; Lau, Rogers, and Passingham 2007; Fried, Mukamel, Kreiman 2011; Jo et al. 2015), whether to press a left or right button (Haggard and Eimer 1999; Soon et al. 2008; Alexander et al. 2015), whether or not to press a button (Brass and Haggard 2007), cued simple movement (Alexander et al. 2015; Shlegel et al. 2015).

Recently, a few neuroscientists have begun to investigate the neurological underpinnings of value-laden decisions, those with options of differing value. Many of these studies have focused on decision-making in rhesus monkeys, given the permitted use of experimental single cell recording in this research subject pool. During a temporal discounting task – one in which the subject could choose a smaller reward now or a bigger reward later, researchers found distinct neuron types across brain regions for distinct aspects of the choice task (e.g., track choice over deliberation, track spatial bias in terms of location of option cue on screen) (Maoz et al. 2013). This kind of work gives us insight into the neural circuits active during decision-making in primates more broadly. Libet and others have proposed to tell us how motor decisions are represented just prior to action in humans, but his work is limited in scope to arbitrary choices. Further, this research goes beyond Libet to investigate how fine-grained aspects of that decision-making process are cognitively represented in the brain.

Much of the neuroscience of agency has to do with the neurological mechanisms by which human agents form intentions to act and control their overt movements. However, insofar as moral and legal responsibility practices make substantive assumptions about the required control for an agent to be held responsible, it is fair game to compare, using knowledge from neuroscience, whether we exercise that control. Here scientists themselves have weighed in on whether our responsibility practices are justified.

Let’s begin with the conditions of agency for criminal responsibility: To be found guilty of a crime, it must be shown beyond a reasonable doubt that the defendant satisfies both the actus reus and mens rea specified for the particular crime. (It should be added here as an additional condition that the defendant lacks an affirmative defense.) Here by actus reus is meant having completed the guilty action (e.g., killing an individual) and by mens rea is meant having certain mental states concerning that action (e.g., intending to kill or harm an individual). But now does the neuroscience of agency suggest that defendants of a certain

group – or maybe human agents on the whole – cannot satisfy these conditions? In an influential paper, Greene and Cohen (2004) argue that insofar as criminal law is, at least in part, driven by retributivist considerations, if neuroscience shows that free will is an illusion, then criminal law needs revision. This contention relies on the following: First, Greene and Cohen assume, without relevant data, that non philosophers are libertarian dualists about the required capacities and conditions for free will. Second, neuroscience of agency supports some version of free will skepticism regarding those conditions for free will. Third, if free will skepticism is right, then defendants won't have the control over their conduct to satisfy the mens rea and so are not deserving legal blame or punishment.

In response to Greene and Cohen's open empirical claim that the ordinary conception of free will is a dualist and libertarian one, Mele (2013j) presents results from his study that the majority of non philosophers still agree that an agent acts freely even if all of the agent's decisions and deliberation are physical events and processes. Mele also notes that further studies by Nahmias and colleagues (2005; 2006) support that a sizable portion of nonphilosophers are compatibilists. Hence, Greene and Cohen's assumption is not warranted. Regarding the broad claim that neuroscience of agency threatens criminal responsibility, Maoz and Yaffe (2014) argue that the current research on the neural mechanisms of agency do very little to excuse or exempt defendants as a whole: First, with regard to actus rea, they note that findings from broadly Libet paradigms do not imply that agents never act on their intentions and perform intentional actions. Hence, neuroscience is consistent with agents acting voluntarily. Second, although studies have assessed representation of risk and ambiguity in decision-making, there is nothing to suggest that neurotypical adults are unable to have the guilty mind required to be held responsible. Maoz and Yaffe note, however, that neuroscience of agency tells us quite a bit about the mechanisms underpinning self-control and so may be central to understanding and revising defenses, such as the insanity defense. Interestingly, some neuroscientists have proposed an explanation of why agents feel regret or disappointment with their decisions and action outcomes, even when those decisions and outcomes are limited: Using a choice paradigm, Filevich and Haggard (2013) presented participants with a set of options for action (e.g., which number to click) and allowed them to pre-select one option from the set before the clicker was active. When the options appeared in the active clicker period, sometimes (but not always) there were fewer options and the pre-selected option was no longer available. The difference in reaction time to click on an option suggests that participants still cognitively represent all of the earlier options, even if they are no longer available. Hence regret or disappointment may stem from, in part, the persisting representation of a no longer available option.

What can neuroscience tell us about agency apart from our interest in free will and responsibility? Roskies (2016b) has recently given an account of how the mechanistic, computational models offered by neuroscience of our decision-making processes can more generally inform, at the level of systems neuroscience, the notion of self-governance. We tend to value control over our mental life and actions in part for the intrinsically valuable practice of structuring and planning our activities. Even if neuroscience cannot definitively answer the compatibility question, dualist threats, and epiphenomenalist threats, as explored above, the empirical results concerning the neurological processes underpinning decision-making can inform philosophical and psychological theories of what makes a person (or system more generally) self-governing. Indeed, elsewhere, Roskies (2012b) has offered a fleshed-out version of a

naturalized account of self-governance in terms of sourcehood which is compatibilist (insofar as we needn't worry about metaphysical determinism) and rooted in psychological findings of self-control. In this way, Roskies argues that we human agents actually possess the control capacities (as Dennett might say) 'worth wanting.'

IV.5 Looking to the future: Neuroscience and free will

The work reviewed above has focused on studies of neuroscience of agency as conducted in experimental cognitive neuroscience labs. However, there is a broader field of neurotechnology that intersects with philosophical questions concerning our sense of identity and place in the world as agents. For instance, Glannon (2012b) discusses the use of these advancements in neuroscience, such as psychopharmacology (e.g., Adderall) and neuromodulation (e.g., deep brain stimulation), to enhance cognitive functioning in nonclinical populations (e.g., college students or shift-workers taking drugs to improve memory and concentration). He explores whether, given the neurological mechanisms affected and their connection to our will, whether such use in each kind of case hinders or enhances our exercise of free will, offering an analysis of each case in terms of the free will capacities impacted. (For a general audience version of Glannon's work on neurotechnology and free will, see Glannon 2014b).

Extending the discussion to clinical populations, Glannon (2013) addresses the use of neuromodulation, including brain-computer interfaces, with patients such as those with prosthetics. He outlines how such technology becomes part of the actor's sense of self as agent and how autonomy is enhanced and restored. Lipsman and Glannon (2013) argue that, indeed, the use of deep brain stimulation by populations such as patients with Parkinson's Disease presents new grounds and should be of interest to philosophical debates about the nature of personal identity. Such neurotechnologies don't just present philosophical concerns and questions but raise critical ethical legal challenges in their (potential) use for criminal populations, particularly those with psychopathic tendencies. Glannon (2014a) intriguingly asks about the ramifications, legally and ethically, of the use of psychopharmaceuticals in place of prison, as well as how prisoners' right to bodily autonomy should be understood in this context.

V. PSYCHOLOGY OF FREE WILL: HOW DO HUMANS ACTUALLY EXERCISE AGENCY?

Much like the neuroscience of free will, psychologists investigate issues of agency and responsibility with a focus on the descriptive, or 'what is the case' questions: how agents do in fact exercise their agency. The philosophy of free will, in contrast, is often focused on normative questions, or 'how things ought to be' questions. For instance, we've seen that philosophers of free will seek to give an account of what capacities or conditions would enable or prohibit the exercise of free will. This is a separate question from what is actually the case — whether we human agents do indeed possess those capacities in the right circumstances. Of course, philosophers might "reach across the aisle" and consult with the empirical findings from psychology, neuroscience, biology, or physics to give an informed response to the existence question. But it is, primarily, the scientific fields of psychology and neuroscience that are concerned with how agents

make decisions and act, including moral situations. As such, it is the results of the studies from these disciplines, while descriptive in nature, that are then used as an evidence base to inform interpretations of the results for our general decision-making and action production capacities. Further, these descriptive scientific findings are then brought to bear upon the existence question or other significant and intriguing descriptive matters, such as how our everyday behavior would change if we all became free will skeptics? Would we be more or less prosocial? More or less aggressive? To what extent would we forgive those who transgress moral norms? The recent work from psychology outlined here provides innovative self-report and behavioral methods to investigating these important descriptive questions regarding our agency and moral practices. (For a more detailed overview of recent psychological work on free will, inclusive of several of the issues discussed below, see Baumeister and Monroe 2014).

V.1 Trait Self-control and Ego Depletion

To understand how agents make practical decisions is, in part, to understand the capacities and mechanisms behind their decision-making and action production processes. We've seen work above on the neural mechanisms underlying practical decision and action from a systems neuroscience perspective. Here we'll consider both theoretical and empirical work from psychology on the cognitive capacities and mechanisms underlying agency at another level of description.

One assumption inherent in much of the philosophical free will literature canvassed above and indeed everyday explanation of action is that representational mental states, such as beliefs and desires (e.g., my belief that it is raining and my desire that I not get soaked) inform – indeed are causes of – of my action (e.g., bringing an umbrella outside). Hence my mental states represent, or are about, the world I encounter (objects, situations), and those representations causally impact my actions (see Crane 2014 for a good introduction to this and other issues of mental representation). For instance, one prevalent idea is that when we decide what to do, we not only consider, and hence, represent the world as it actually is, but we consider, or represent, the world as it could be. That is, we entertain counterfactuals about the world. Chater and Oaksford (2013) draw on the literature on computation and computer programs to provide a theoretical framework for understanding how systems – whether natural (e.g., human minds) or artificial (computers) encode, or represent, both the actual situation and counterfactuals in their causal models of the world. This theoretical work may find some application in the aforementioned work on persisting representations of alternative actions (Filevich and Haggard 2013).

Another everyday assumption about decision-making and intentional action is that we as agents need to exert willpower from time to time – we need to resist satisfying our strong desires or urges. For example, if I am committed to cutting back on eating pastries, then when I see that delicious-looking blackberry scone at the coffee shop, it will require willpower to resist ordering and eating a scone. Psychologists have studied a close cousin of the folk notion of willpower in research on self-control and ego depletion (for the seminal study, see Baumeister et al. 1998). The basic idea is that we, as unified agents, must exercise self-control during various tasks – e.g., restraining our impulses, regulating our emotions, making difficult decisions, performing cognitively demanding tasks. Research shows, though, that this capacity for self-control is not unlimited: After engaging in decisions and actions that require self-control, agents can suffer from 'ego depletion' – the agent's self-control is temporarily diminished and so the agent does not perform

as well on tasks that require self-control. That is, if one exercises self-control during some task, then it will be harder for one to do so on another task that required self-control. Recent work by Vohs et al. (2017) suggests that ego depletion can lead to a form of “mental passivity” in which agents are particularly disinclined to engage in subsequent tasks that required effort (e.g., eating food that requires some preparation, answering consumer questions) and more inclined to engage in passive tasks (e.g., watching TV). As we will outline in the research below, regardless of one’s trait self-control - or relatively stable over time ability to exercise self-control, successful exercise of self-control is dependent upon previous attempts to exercise self-control.

It’s interesting to note that although the effects of exercising self-control are known, the psychological mechanism underpinning the capacity for self-control is unclear and the jury is still out among psychologists. Some studies have supported that self-control is served by a limited set of resources, understood in terms of energy reserves. In this vein, some research suggests that the underlying mechanisms serving self-control have been tied to the physiological mechanisms governing blood glucose levels (Galliot et al. 2007). However, this account of self-control and its underlying physiological mechanisms is not univocally accepted and many reject a ‘resource’ mechanism account. For instance, Levy (2016b) offers a competing account of the relationship between glucose stores and exercise of self-control.

Setting aside the mechanism question, research supports that self-control is implicated in a variety of tasks, including navigating test anxiety in order to perform well on cognitive tests (Bertrams et al. 2013). Work by Hofmann et al. (2013) supports that those with high trait self-control – demonstrate a capacity for self-control over time – report greater affective well-being and higher life satisfaction. The authors attribute this relationship, in part, to the fact that those who regularly exercise self-control avoid, comparatively, motivational conflict and greater emotional distress than those agents who don’t. Moreover, Daly et al. (2014) found that, physiologically, those with high trait self-control exhibited lower resting heart rate and distinctive cortisol level slope over time as compared to other participants. This finding may be due to the ability of those with high trait self-control to maintain stable affect and so promotes health. Further, recent studies, using a wide array of methods (e.g., self-report, behavioral, laboratory setting and experience sampling) support that there is no limit, or ‘dose-dependent’-like relationship between trait self-control and subjective well-being. That is, it has not been observed to be the case that an abundance of self-control is associated with lower well-being (Wiese et al. 2018).

Interestingly, some have registered skepticism at the idea of Trait Self Control (Tangney, Baumeister, and Boone 2004) as predictive of temptation resistance and more broadly a tendency for self-control behaviors: Hoffman et al. (2012) found that trait self-control does not predict success in resisting impulses. Likewise, Levy (2017) argues that, empirically, the link between those scoring high on the Trait Self Control scale and exhibiting self-control behaviors is weak. Hence, to the extent that successful self-control is general domain (non self-control specific) skill-based, one might question the positing of self-control as a trait individuals can possess. Snoek, Levy, and Kennett (2016) apply this skepticism and general domain skill-based account of self-control to recovery from addiction, arguing that addicts are not necessarily deficient

in willpower. Rather, those who are more successful in recovery may be those who are better at maintaining willpower through the recruitment of skill-based strategies.

Still, setting aside debates regarding trait self-control, research supports that ego depletion hinders goal pursuit. Which factors led to successful goal pursuit? Ego depletion bears an interesting relation to the agent's cognitive and evaluative states. Job et al. (2010) support that an agent's belief about whether willpower is limited or not impacts whether she suffers from ego depletion. Those who did not believe that willpower was limited did not show ego depletion. Building on this finding, Vohs and colleagues (2012) provide evidence that if agents believe that willpower is unlimited or find a task especially meaningful, mild ego depletion is mitigated. That is, such participants are more likely to exercise self control after recent completion of a task requiring self-control compared to other ego-depleted agents. However, this effect of motivation and cognition does not hold for agents who have suffered more extreme ego depletion. Further evidence that meaningfulness of a task buffers ego depletion come from studies of decision-making requiring pain and effort. One might, intuitively, guess that the amount of effort and pain a task requires is negatively correlated with the frequency of engaging in that task – i.e., the more effort or pain one would need to go through to complete the task, the less likely one would be to complete the task. The conventional wisdom goes, people seek to avoid unpleasantness and pain. However, Olivola and Shafir (2011) found that if the experimental task was a prosocial one (here charity donations), especially one aimed at alleviating the suffering of other people, participants were more likely to give more money as the effort and pain involved increased. The researchers support the interpretation that the more a task aims to alleviate human suffering, the more meaningful people find the task and hence the more likely they are to commit more resources to the task. The researchers dub this “the martyrdom effect.”

Personality, such as whether an agent is action-oriented or state-oriented, is another factor in successful goal pursuit (Grope, Baumeister, and Beckmann 2014). More broadly, studies support that those with specified plans to obtain their goals are on average more successful in achieving those goals (see also discussion of implementation intentions in section IV.6). Masicampo and Baumeister (2012) add the critical conditional, though, that if one is under strict time pressure to accomplish a goal, specific plans may hinder, not aid, goal completion. Agents with insufficient time to carry out their plans succeed at higher rates when they are open to alternative means. More broadly, complex factors beyond energy reserves, such as whether one is changing focus from a more difficult or easier task to a new task – ‘task switching,’ affect one's ability to initiate and execute new decisions and actions (see, e.g., Barutchu et al. 2013a for a discussion of the mechanisms driving this effect. Barutchu et al. (2013b) discusses ego depletion in relation to task switching).

While much of the literature on ego depletion is conducted in the lab, Hofmann, Vohs, and Baumeister (2011) surveyed adult participants' desires and efforts to avoid temptation out in the world. Using mobile phones, researchers signaled participants throughout their day to report the occurrence of various desires, the strength of those desires, and the participants' success in avoiding satisfying those desires (when those desires conflict with goals). This broader method of sampling is called experience sampling. Self-reports suggested that desires for rest and sex were the strongest, whereas participants experienced more frequent self-control failures related to desires for media use and work. Using regression modeling, the study

identified frequency and timing of self-control exercise as a predictor of subsequent self-control failure, supporting the ego depletion theory. Baumeister and colleagues take self-control to be one of the key components of human agency and part and parcel of what it is to exercise free will (see section V.4). The capacity for self-control has been tied to the evolved capacity for social interaction and is utilized in the performance of a variety of kinds of human decisions and voluntary action (see Alquist and Baumeister 2012 for a review). How does self-control operate in social contexts?

We've noted that self-control capacities can strongly influence our ability to act in effortful, taxing, or tempting situations. These situations can include our interactions with social others. For instance, participants with low trait or state self-control more frequently commit social norm violations (Gailliot et al. 2012). Recall here that trait self-control is the stable capacity for self-control as evidenced over time (and not on a select occasion). Relying on the rich empirical findings concerning self-control and ego depletion, Baumeister (2013a) argues that antisocial and criminal behavior are mediated by an agent's failure to exercise self-control. Hence, criminal law might usefully use our understanding of ego depletion, a key psychological and physiological mechanism behind failure of self-control, to inform understanding of such behavior. And hence, he notes, on an optimistic note, that insofar as strengthening of trait self-control capacities is possible, strategies for strengthening self-control may be of benefit to certain populations, such as the incarcerated.

V. 2 Belief in free will and social behavior

The aforementioned surveys on non philosophers' judgments of free will and responsibility (see section IV.4) tell us a lot about the prevalence of compatibilist or incompatibilist, physicalist or dualist views of free will and moral responsibility among the general population. But a further question of import is whether non philosophers believe that free will exists – in particular, believe that some agents have the control over their decisions and actions required to be held accountable morally and legally. We might think of this as a *trait* – stable over time – belief in free will. But we could also inquire if non philosophers' *state belief* in free will – belief at a time – can be changed and whether these induced differences in belief in free will impacts social behavior and judgments regarding others. These questions have recently been taken up in force by social psychologists.

In a review of literature of the “correlates and consequences” of nonphilosophers' belief in free will, Baumeister and Brewer (2012) note that studies support that baseline belief in free will – prior to any experimental manipulation – is high among non philosophers. To measure belief in free will, researchers have used the Free Will and Determinism Plus (Paulhus and Carey 2011) or the Free Will and Determinism Scale (Rakos, Laurene, Skala, & Slane 2008). More recently, the Free Will Inventory has been developed to address some issues with older measures of free will beliefs (Nadelhoffer et al. 2014). Interestingly, although the belief in free will is generally high among the population, Ent and Baumeister (2014) found that one's embodied experience is related to the strength of one's belief in free will: Those with epilepsy, panic disorder, and strong felt physical desires reported lower belief in free will than others. In relation to the prior psychological literature that supports a metaphysics-light choice model of free decision-making among non philosophers (see section V.5), Feldman, Baumeister, and Wong (2014)

found evidence that those with a strong trait belief in free will associate exercise of free will with choice and enjoyment of making choices.

Research regarding the correlates of belief in free will is a reliable guide to how free will believers typically think and act. However, several factors correlate with belief in free will, and as such these factors may potentially explain the findings (e.g., conservatism correlates with belief in free will - Cary and Paulus, 2011; Crescioni et al. 2015). Therefore, psychologists must move beyond correlational studies and experimentally manipulate belief in free will to disentangle the contribution of belief in free will on thought and action. The main methods for induction of disbelief in free will include having participants read (and sometimes restate) claims affirming a lack of control and responsibility, whereas control participants read non-related statements, typically scientific ones. Baumeister and Brewer (2012) outline, among the many noted, the following main behavioral effects of induced disbelief in free will: increased tendency toward antisocial behaviors (in the lab) such as lying, cheating, and stealing (Vohs and Schooler 2008); expressions of aggression and unwillingness to help others (Baumeister, Masicampo, and DeWall 2009). In comparison, measured high belief in free will was associated with greater work ethic among laborers (Stillman et al. 2010). In correlation studies (Alquist and Baumeister 2010; Alquist, Ainsworth, and Baumeister 2012) participants who endorsed a stronger belief in free will conformed less with others' actions than those who had a weaker belief in free will. In follow-up experimental paradigms, participants who were induced to disbelieve in free will (e.g., read statements about lack of control) were more likely to conform with others' actions and judgments (e.g., offered the same or similar opinions of artwork to those opinions of confederates). MacKenzie, Vohs, and Baumeister (2014) found that expressions of gratitude are related to belief in free will: First, those with high trait belief in free will demonstrated a greater tendency to feel gratitude. Second, induction of disbelief in free will reduced expressions of gratitude for both past occurrences and occurrences in the lab. A strong belief in free will has been associated with several positive traits and outcomes, including reported life meaning and satisfaction, a tendency towards gratitude and forgiveness, and more reported commitment to relationships. Those who were induced to disbelief in free will reported lower levels of meaningfulness in life (Crescioni et al. 2015). (For a more recent review of the belief in free will literature that covers some of these studies, see Baumeister and Monroe 2014.)

Given these connections between belief in free will and social conduct as well as research outlined above on self-control, Baumeister and colleagues have often expressed free will as “an advanced form of action control that evolved to enable people to function and thrive in cultural groups” (Monroe, Vohs, Baumeister 2016, p. 41). However, Schooler registers some qualifications with this generalization of free will as promoting pro-sociality. In his online general discussion of research on belief in free will (2013), Schooler notes that some of the prominent findings, including his own seminal study on belief in free will and cheating (Vohs and Schooler 2008), have failed to replicate. That is, other researchers have been unable to obtain similar results when rerunning the experiments (in a relevantly similar manner). Instead, Schooler endorses a more subtle view: that whether belief in free will makes us better people depends on factors, such as personality.

Belief in free will has been linked to disposition to punish. Some studies suggest that a belief in free will is intimately tied to a desire to hold others responsible for bad actions and may function to justify one's punitive actions against those transgressors. For instance, Clark et al. (2014) found a relationship between punitive tendencies and belief in free will: Participants who either consider or are exposed to immoral behavior in the lab evidence increased belief in free will as compared to controls. This effect is supported by correlative work on country-level data, in which countries with higher rates of crime and homicide report stronger beliefs in free will in surveys. The authors argue that this effect is likely due to enhanced punitive motivations. Hence, the function of our belief in free will may be to aid in motivating punishment for morally bad actions. We've seen above, too, that induced disbelief in free will can lead to less punitive tendencies on retributivist grounds (Shariff et al. 2014; see section IV.4; see also Martin et al. 2017).

In line with these findings, some have argued, on partly conceptual grounds, that a belief in free will may foster negative and discriminatory attitudes towards the working poor and minorities (Miles 2011). The line of thought seems to be that if one believes that agents make their choices and act freely, then one believes that the agent is responsible for the consequences of those choices and actions. In response to this contention, Vonasch and Baumeister (2013) conducted surveys which suggest that this is not the case: Using the Free Will and Determinism Scale as a measure of free will beliefs, the researchers found that belief in free will is not correlated with such attitudes: First, degree of belief in free will is not correlated with sympathy with a person who works part-time by choice, but in fact stronger belief in free will is related to *greater* sympathy with a person who works part-time but is searching for a better job. Further, belief in free will was positively correlated with a desire for income equality. Scoring higher on another measured trait though, called Just World Belief, did track many of the negative and discriminatory attitudes that Miles incorrectly associated with belief in free will. Hence, Vonasch and Baumeister conclude that a strong belief in free will in itself does not lead to more discriminatory individuals in society even if belief in free will underpins punishing transgressors.

What overall message can we draw, then, from psychological work on belief in free will and its implications for social agency? Free will and responsibility skeptics of the philosophical sort have debated the repercussions of a societal belief in free will and/or moral responsibility (see, e.g., Nadelhoffer (ed.) (2013) for a survey of skeptical views and their implications). A critical aspect of this debate revolves around the issue of what human relationships and institutions would be like if society rejected free will or moral responsibility (see Strawson 1962 for a foundational paper on the topic). In review of the belief in free will, Shariff and Vohs (2014) argue that a society without the presumption of free will may well be less retributivist in punishment, and in this way more exonerating of wrong-doers. This claim is in line with the broader theme of Cohen and Greene (2004) as it pertained there to neuroscience of agency and criminal responsibility practices. However, they also argue that free will may be too beneficial of a notion to abandon completely, given the kinds of antisocial behaviors associated with reduced belief in free will.

Notably, belief in free will not only impacts social behavior, but also the capacities, neurological and psychological, that underlie an individual's exercise of agency, many of which have been discussed in section IV and V.1. For instance, induced disbelief in free will among participants leads to weaker RPs prior to voluntary movements in Libet-tasks (Rigoni et al. 2011; see section IV for a review of the abundant

Libet literature). One noted component of decision-making and action production reviewed above is counterfactual thinking (see V.1). Roughly, counterfactual thinking is the ability to simulate via mental representations alternative ways that events could be. Counterfactual thinking is often assumed to be critical to the process of practical deliberation, when an agent considers potential courses of action from a perspective that the future is open (Even if metaphysical determinism is true, we can have epistemic openness from the perspective of the agent who does not know the future.) One intriguing finding, then, is that both induced belief in free will and measured baseline high belief in free will is related to increased counterfactual thinking (Alquist et al. 2015).

V.3 What might a psychological account of free agency look like?

If we start with a ‘yes’ answer to the existence question (see section III) – that is, it is the case that free will exists, then we can turn to psychology to fill out exactly what kinds of psychological capacities for decision-making and acting that free agents rely upon. One psychologist who works extensively in this vein is Roy Baumeister, noted above in the work on self-control capacities and ego depletion. As Baumeister explains this in his (highly accessible) Slate article, “there is a *genuine psychological reality* behind the idea of free will” (2013b).

Baumeister offers a “scientific theory of free will” (2008; 2013; 2014a; 2014c; Baumeister and Monroe 2014): This type of free will theory takes it that no essential element of free will is “supernatural” or “noncausal” (Baumeister and Vonasch 2011; Baumeister 2014a). Insofar the psychology of free will does not tell us about the supernatural and tend to offer causal explanations, a scientific account of free will will exclude these elements. Another critical assumption is that free will comes in degrees (Baumeister, Clark, and Liguri 2014). Further, Baumeister conceives of a scientific account of free will as answering the free will criterion question with an account of free action – “the capacity for free action” (as opposed to ultimacy or volitional freedom) (Baumeister 2014a, p. 236). Hence, the appropriate question to ask is the following, according to Baumeister (2014a): “To what extent and in what sense(s) can humans act freely?” (p. 235). We’ll canvass some answers to this question, drawn predominantly from the work of Baumeister and colleagues, here. The proposed answers rely heavily on what psychology can tell us about our agential capacities.

One major component of Baumeister’s account of free will is the agent’s capacity for self-organization (Baumeister and Vonasch 2011). When we conceive of agents as systems, human agents are physical systems which act on representations (see section V.1) and often independently of environmental stimuli. This is the psychological analogue of the idea of endogenously-generated actions in cognitive neuroscience (see section IV). Moreover, on this view, an agent in the sense of a self-organized system is autonomous – it distinguishes itself from the environment and acts as a whole system, where “the whole is greater than the sum of its parts” (Baumeister and Vonasch 2011, p. 44). Hence, Baumeister focuses on personal-level representational mental states and their role in meaningful thought and intelligent behavior. He argues that the capacities underlying free action cannot be understood in terms of the working of physics or neurons but that “freedom exists at a high level of integrated self-organization” (2014a, p. 255). This

account is then supported with work on psychological models of human behavior as well as work on the role of consciousness and emotion in decision and action (see, e.g., Baumeister and Vonasch 2011).

Further, Baumeister (2013b; 2014a; 2014b; Baumeister, Clark, and Ligiuri 2014) draws on his body of work on self-regulation and ego depletion to argue that our free will is constituted, at least in part, as an evolved capacity for self-control and rational choice. This free will capacity likely evolved due to its role in facilitating social coordination in a cultural context. Insofar as ego depletion can occur in degrees, the capacity for free action can therefore be enhanced or diminished. The argument that free will is an evolved set of capacities for action control has been further supported with appeal to the psychological research on the prevalence and variety of belief in free will and the social and personal implications of those beliefs (Baumeister, Crescioni, and Alquist 2011; see Mele 2011a for a commentary).

We've noted in section III that, typically, free will is taken to be the control condition for moral responsibility. That is, an agent who exercises free will has the control over his or her actions required to be blameworthy (for a morally bad action) or praiseworthy (for a morally good action). Our moral responsibility practices are part of our broader system of morality – the kinds of moral codes we as a community live by and embody as well as our reactions to violations of that code. And so, just as we can, through the psychological sciences, offer a descriptive account of free will, we can also offer a descriptive account of morality. Psychology and other disciplines, such as cultural anthropology, can tell us the kinds of moral codes and practices actually in place across cultures (and their commonalities, if any).

In Baumeister and Vonasch 2012, the descriptive focus is on the substance, or essence, of morality. One major theory regarding the substance of morality is the view that making moral judgments – and the capacities required to do so – is the essence of morality (see, e.g., Gray et al. 2012). In a commentary of such views, Baumeister and Vonasch argue that such a unitary focus in research leaves out other vital features of human morality, such as action selection and control (e.g., deciding how to act and controlling one's action given a moral code). If so, they note, then capacities such as self-regulation (e.g., the kinds of self-control capacities in section V.1) are just as important as capacities like perception and understanding of other minds. Likewise, adequate focus should be drawn to the cultural context as framing the moral rules and norms governing social interactions (Baumeister and Vonasch 2012). Indeed, in a commentary on the evolution of fairness, Ainsworth and Baumeister (2013) note that self-regulation capacities are critical to enable social cooperation and acting on norms of fairness, as dictated by cultural norms, in the face of competing selfish desires.

V.4 What is the role of consciousness in agency?

As we have seen, Libet and others inspired by Libet's work have argued for the epiphenomenalist challenge to free will: Conscious intentions to act never cause our actions, so no one exercises free will. This argument isn't accepted wholesale by researchers of agency, as we have seen with Mele (2009)'s argument that conscious intentions to do some task later can increase the likelihood of successful completion. However, there's at least a kernel of truth to this epiphenomenalist challenge; consciousness does not seem to be required for the completion of all tasks. For instance, work conducted on blindsight

patients demonstrates how conscious awareness isn't required for accuracy in task completion (for recent work regarding, see Ko and Lau 2012). More broadly, even for neurotypical individuals, there is a complex relationship between agents' confidence in their detect of some stimuli (e.g., object), task performance accuracy, awareness, and cognitive control. This relationship has been explored extensively for visual discrimination tasks in the lab (see, e.g., Maniscalco et al. 2012, Maniscalco and Lau 2012, Rehnev et al. 2012a, Rehnev et al. 2012b, deLange et al. 2013, Maniscalco and Lau 2014, Solovey et al. 2014, Grimaldi et al. 2015, Koizumi et al. 2015, Maniscalco and Lau 2015, Morales et al. 2015a, Morales et al. 2015b, Maniscalco and Lau 2016, Maniscalco et al. 2016, Peters et al. 2017a, and Peters et al. 2017b).

Even if we accept the stronger claim that conscious mental life isn't among the causes of or guide our immediate actions, there's evidence that consciousness plays a distal or indirect role in our agency, particularly for tasks that require flexibility and planning. For instance, Baumeister and colleagues (Masicampo and Baumeister 2013; Baumeister 2014a; Baumeister and Bargh 2014; Baumeister, Vohs, and Masicampo 2014) propose that consciousness plays important social and cultural functions: Specifically, they propose that consciousness is an evolutionarily adaptive capacity that allows humans to generate novel and beneficial socially and culturally appropriate behaviors, allowing for social coordination and communication. Here, part of the story is the theory that conscious thought serves as a hub for mental simulation – or thinking through nonpresent events. This kind of capacity makes possible the consideration of alternative actions and coordination of one's actions across time and with social others in one's cultural community.

In a similar vein, Baumeister, Schmeichel, and DeWall (2014) outline how the exercise of certain creative skills in the lab, such as drawing or jazz improvisation, can be diminished when participants are under cognitive load. Cognitive load is a well-worn laboratory paradigm for taxing or occupying participants' cognitive resources. Participants are given two tasks to complete simultaneously, with one of the tasks requiring extensive cognitive resources, such as counting backwards by six. The thought is that if task performance is hindered on another simultaneous task – here the creative task of drawing or musical improvisation – compared to the performance of participants not under cognitive load, then that task too must draw on cognitive resources. Baumeister and colleagues (Baumeister, Schmeichel, and DeWall 2014; Baumeister, Vohs, and Masicampo 2014) argue that these findings fit with the view of the conscious mind as integrating and transforming key products of unconscious processes to produce complex behavior. In this way, the conscious mind isn't so much competing with the unconscious processes, but is rather *complementing* the functions of those unconscious processes.

V.5 Developmental psychology: How do our agential capacities develop?

The previous psychology research surveyed here in section V has been conducted with adult participants. Here we have discussed how adult humans exercise self-control, act in social contexts, as well as whether they believe in free will and how their beliefs impact their social behavior. We've also looked at accounts of free will, proposed by psychologists, which are grounded in the psychological research on agential capacities. But one may well wonder how adults develop these free will capacities. After all, infants and

young children aren't held fully accountable, morally or legally, for their behavior and this is due in part to the fact that young children aren't assumed to be fully in control of their behavior.

These questions are fascinating ones and could occupy a review in itself. For the purposes of the present review, we will briefly consider a selection of recent research on belief in free will among children. We've noted that in surveys and interviews adults tend to focus on choice and constraint as the main elements of exercising free will (e.g., Monroe and Malle 2014). In their research with young children aged 4 to 6, Gopnik and colleagues (Gopnik and Kushnir 2014; Kushnir et al. 2015) similarly focused on ability to choose and constraints on action. Specifically, Gopnik and Kushnir told the children stories about agents performing various kinds of actions (e.g., physically possible, physically impossible) from either a first-person or third-person perspective. Here judgments about the ability to do otherwise and lack of constraint are used as a proxy or operationalization for children's free will judgments. The children were then asked for each kind of action performed: "Did he (you) have to do [fill in the action] or could he (you) have chosen to [alternative action]?" Here the alternative action was either a physically possible or impossible one – one physically constrained by force or the laws of nature. For example, children either drew a dot on a sheet of paper unaided or were made to draw a dot (with their hand held firm). They were then asked "Could you have chosen to draw a line rather than a dot, or did you have to have to draw a dot?" The majority of children responded that one could have chosen to do alternative physically possible actions (e.g., 71% for the dot/line question in the unaided condition), but in comparison only a small percentage said that one could have chosen to perform a physically impossible act (e.g., 19% for the dot/line question in the constrained condition). This effect was demonstrated cross culturally in a sample of children in the same age range in China (Wente et al. 2016).

One aspect of choice and constraint as elements of free will is its connection to inhibition, or acting contrary to one's desire. Section V.1 examined the psychological capacities that underlie exercise of inhibition in adults via work on self-control. Gopnik and Kushnir (2014; Kushnir et al. 2015) tested whether, and how, young children understand the motivational force of our desires in action and the ability for self-control. In particular, the researchers were interested in whether young children view desires as a mental constraint on action. Both 4 year-olds and 6 year-olds were told stories about a choice to not do something desirable (e.g., desire to eat a cookie) or a choice to do something undesirable (e.g., look in a scary closet) in either third-person or first-person cases. They were then asked whether the agent (or you) can choose to do the undesirable action or if she (you) have to do the desirable action. While 6 year-old children evidenced an understanding that agents can act contrary to their desires, 4 year-old children were less likely to respond in this way. Interestingly, children in the US and China show differences in these patterns of judgments, with 6 year-olds in the US more frequently endorsing this ability for self-control than in China (Wente et al. 2016).

The results suggest, then, that understanding self-control as a free will capacity emerges during the course of childhood, perhaps as the children themselves develop these capacities. The researchers suggest that these intuitions and conceptions of free will develop in conjunction with understanding of agency and causality, as part of theory of mind capacities. Gopnik and Kushnir sketch how intuitions of agency

demonstrated in the studies connect to the broader theory of how children come to exercise and understand intentional agency.

VI. THEOLOGICAL ISSUES: CAN THE NATURE OF THE DIVINE LIMIT OR RULE OUT FREE WILL?

Thus far, with the preceding discussion of metaphysical and neurological determinism, the focus has largely been on nonagential threats to the exercise of free will in the form of processes, events, and laws of nature that may limit or rule out free agents. But we may well worry – as philosophers, theologians, and folk everywhere have – about a very different kind of determinism: divine foreknowledge and divine intervention over the lives of humans. We'll understand this kind of determinism as the following two theses:

Divine Foreknowledge: God has infallible knowledge of all truths about the world (including all human actions).

Divine Power: God is omnipotent and has ultimate determination over all events in the world (including all human actions).

We can then ask a corresponding *compatibility question* about Divine Foreknowledge (or Divine Power) and free will:

Can humans exercise free will if Divine Foreknowledge (or Divine Power) is true?

Recall the Consequence Argument; a similar line has been run historically about the existence of human freedom given divine foreknowledge or divine power: Assume divine foreknowledge: Then God *infallibly* believes at some earlier time, $t1$, that a human agent will perform a specific action, A , at a later time, $t2$. Or, alternatively, assume divine power: that God as a creator brings it about via his agency at some earlier time $t1$ that a human agent will perform a specific action, A at a later time, $t2$. But now, in this way, all human actions are fixed given God's infallible knowledge and the consequences of God's earlier agency. For, there is nothing human agents can now do about God's foreknowledge or agency (prior to human existence). As such, which actions human agents perform are too out of the hands of human control (For a good introduction to the argument, see Haskers 2011). Hence, divine power and divine foreknowledge appears, at first glance, to rule out deep openness – there is only one possible continuation of the past given God's providence. Therefore, divine power and divine foreknowledge preclude the ability to decide (or act) differently and ultimate sourcehood. As may be apparent, this kind of determinism is only a threat to incompatibilist accounts of free will and not compatibilists accounts which understand ability to decide (or act) differently and ultimacy in ways that don't require deep openness. But now if humans don't exercise incompatibilist free will, there is a worry that humans don't deserve divine judgment and punishment for their good and bad actions. This section focuses on theological-driven issues of free will, outlining new work on the relationship between human freedom and divine determinism, often termed *theological fatalism*.

The prima facie incompatibility of divine power and human freedom turns on the details of God’s nature as an agent. One such doctrine concerning the nature of God, Divine Universal Causality, holds that “necessarily God directly causes all entities distinct from himself for as long as such entities exist” (Grant 2016, p. 214). On this doctrine, God causes all human agents and all human actions. Hence, one may well worry that human actions, given this necessary cause, cannot be free in an incompatibilist sense (e.g., the agent cannot act otherwise than he does or is not the source of his actions). Grant (2016) argues that God’s causal relation to human actions is compatible with both Divine Universal Causality and with human libertarian free will.

Another philosophical theologian takes a decidedly more historical approach to issues of God’s nature and human freedom: Couenhoven is an Augustine scholar (2013b; 2016a; 2016b; 2016c) who extensively examines the intersection of original sin, predestination, and human freedom. Original sin refers, roughly, to the idea that humans are born, or are inherently, sinful. Augustine is a major proponent of original sin. Predestination is often discussed as a narrower version of theological fatalism: here the focus is on human salvation: if God causes all being distinct from himself, then it is predestined, or predetermined, whether a human agent will obtain eternal life, or salvation (see, e.g., Couenhoven 2014; 2017). Still, Augustine held that human agents can act freely and be responsible for their conduct. In his book, *Stricken by Sin* (2013c), Couenhoven reconstructs and critically engages the doctrines of original sin and predestination of medieval theologian Augustine. Couenhoven defends a compatibilist account of responsibility, drawing from Augustine, even in the face of original sin and predestination. This Augustine-influenced account is then situated in our contemporary social lives.

Coenhoven’s work bridges traditional philosophical theology with more analytic philosophy of religion (see, e.g., Coenhoven 2013a). For a user-friendly guide to the terms used in the discussion of predestination, see Couenhoven’s “Predestination” in the *Vocabulary for the Study of Religion*. For an accessible overview of Couenhoven’s research questions and the applicability for our everyday decisions, interactions, and experiences, see a recent online interview with Couenhoven (2015a). Couenhoven has also been interviewed on the topics of theological fatalism and the problem of evil on PBS’s program *Closer to the Truth* in Season 14 (episodes 2 and 12).

As Couenhoven’s work exemplifies, the proposed incompatibility between divine foreknowledge and human freedom has a long intellectual history. Recently, Hunt (2016) has proposed that the apparent conflict between the nature of God and human agency should be reframed as an aporetic problem. Instead of thinking of either divine foreknowledge or human freedom as being at risk, philosophical theologians ought to approach the issue as a puzzle that recommends revisiting of theological assumptions and conceptions. He argues that on this approach, we have good reason to think that there is something wrong with the argument for theological fatalism, while also taking seriously that we may need to revise our implicit or background assumptions given the argument. Hunt draws on work of theological fatalism from Augustine and to motivate this approach to the argument and proposes a solution to the problem.

The dominant understanding of divine knowledge is that God is all-knowing. Hunt (2013) explores a related intriguing asymmetry between divine power and divine knowledge. He notes that although God's divine power, or omnipotence, is understood to mean that God is unrestrained in what he can do, not understood to mean that God does in fact do everything he can. However, God's knowledge, or omniscience, is understood as actively exercised ability – that is, God doesn't just have the capability to know everything, but he does in fact know everything. This conception of God's nature is well-established. However, as Hunt puts it, "Why is God thought of as omniscient rather than 'omni-knowledgeable,' where the latter means roughly 'being able to know whatever one wants to know whenever one wants to know it'" (2013, p. 1). Hunt canvasses the justifications for this asymmetry and argues that an omni-knowledgeable being ought to be considered as omniscience. Acceptance of omni-knowledgeability for theological theory construction is explored.

We've mentioned divine power and divine foreknowledge as proposed attributes of God. One further aspect of the nature of God in philosophical theology is perfect goodness. The philosophical literature includes discussion of, for example, whether divine perfection is possible (or perhaps necessary) and further, assuming the possibility, the features of a divinely perfect being are one point of discussion (see, e.g., Murphy 2014 for a good overview of the topic). For our purposes, though, the status of God as a perfectly good being intersects with issues of freedom and moral responsibility. First, to say that God is perfectly good is to attribute, on one reading, moral qualities to God. But if one attributes moral qualities to God, then God is subject to our moral assessment and may be obligated to act, morally, in certain ways. If moral assessment requires free will, then God is a free agent (divine freedom). Leftow (2013) argues that moral perfection is possible, and that if God exhibits moral perfection, he must be subject to and satisfy moral obligations.

Consideration of God's nature as a being on the whole— divine power, divine foreknowledge, and divine perfection – yield further questions regarding agency. God is, on the view noted above, a free agent. Can God satisfy conditions of free agency that have been proposed for human agents? To what extent, does God's omnipotence (form of divine power) and omniscience (form of divine knowledge) constrain his agency? For instance, one point of discussion is whether God has the ability to do otherwise, one of the free will capacities considered in section III. Some theists argue that God provides for humans in ways that he need not have done (e.g., human free agency, sending Christ to die for them). Leftow (2016) explores the extent of and limits on God's ability to do otherwise, given his omniscience and so perfect rationality. He accepts a version of voluntarism (theory of God's nature as will) and argues against a rationalist picture of God's agency. He argues that God, on this conception, is bound to the morally best options, of which there may be more than one. However, God may have preferences regarding nonmoral matters (e.g., aesthetic) and in the faces of options with differing nonmoral value, God is not bound to choose among the option(s) that optimize nonmoral value and may choose in line with preferences. (In contrast, Couenhoven 2016d defends non voluntarism regarding God's freedom and responsibility.) In a similar vein, Leftow (2015) defends an account of robust divine freedom and argues that a certain established doctrine of God's nature, divine simplicity, is incompatible with robust divine freedom and so ought to be rejected. Other puzzles of divine agency assuming divine perfection arise, such as weakness for temptation. If God is perfectly good, then he presumably is not tempted to do evil nor does evil. By

‘doing evil’ here, we mean one species of morally bad actions. However, Biblical accounts assert that Christ, the Son of God, was tempted as human agents are. Human agents are, at times, tempted to do evil. Hence, Christ was at times tempted to do evil. Leftow (2014) takes up this puzzle. He argues the given Christ’s embodied nature, temptation to do evil is consistent with divine goodness.

One well-known theological problem at the intersection of divine power and human freedom is the problem of evil: roughly, how could God, a perfectly good and omnipotent being, create a world with evil. This quandary may lead us to be skeptical of the existence of such a God. Grant (2015) tackles a related problem tied to human freedom: Consider again Divine Universal Causality – the assumption that God is the cause of all being distinct from God, and humans are part of that being. But now humans perform sinful actions. Hence it seems that God causes sin. This conclusion, however, is at odds with the view that God does not cause human sin. Grant (2015) defends a historical response to this argument, articulated by Anselm and Aquinas. Grant argues, in line with these theists, that God does not cause sin as it is the sinner, the human agent, who causes the privation, or lack, of virtue of sinful acts.

VII. CONCLUSION

We take it that when we act in the world it is, at least sometimes, up to us how we act; it seems as though we decide, ourselves, what to do and control our movements so that we are agents of change in the world around us. The disciplines of philosophy and the sciences have taken up this issue of agential control to reveal aspects of our lives as agents: *Over what* do we exercise control worthy of being called free agency? Here we’ve seen proposals that our free agency may extend to our decisions, our intentional actions, and perhaps our attitudes and failures to act. We have also canvassed more skeptical claims about free agency. We have considered questions of the *kind of control* that agents do exercise or would require to exercise free will - here proposals ranging from more modest event-causal pictures (free actions caused by our intentions and other mental states) to more demanding agent causal pictures have been reviewed. We have also considered challenges to free will, from the nature of the world to the nature of the Divine. How one answers these questions has prime bearing for our view of ourselves as self-governing and responsible agents, morally, legally, and (from the view of theology) in the eyes of God.

What’s striking about this collection of research on free will – albeit undertaken from different approaches, descriptive and normative– is the rich interdisciplinary discussion fostering new understanding. Of course there are limits to what psychology and neuroscience can tell us about metaphysical matters such as determinism and dualism. To this end, scientists and philosophers would do well to continue to foster interdisciplinary dialogue on what, exactly, we mean when we speak of determinism (and varieties thereof) as well as how nonexperts conceive of free will and moral responsibility and the impact of that conception on their social behavior. Still, descriptive work from the sciences, outlined here and in ongoing work, on how we exercise agency in terms of the underlying psychological and neurological processes and the role of consciousness and social context aids in addressing whether we measure up to the normative philosophical accounts of what free agency requires and, perhaps, ideally would be. We’ve seen, for example, work that suggests the powers and limitations of our capacity for self-control as well as the neural markers of intentional action. Likewise, the normative fields of philosophy and theology are rich with conceptual resources and distinctions, such as enriched understanding of

intentional action and the abilities of free agents, that are increasingly taken up in the theoretical framework of neuroscientists. Continued collaboration of the humanities and sciences promises new insights regarding our lives as free and responsible agents.

VIII. REFERENCES

- Ainsworth, S.E., & Baumeister, R. F. (2013). Cooperation and fairness depend on self-regulation. *Behavioral and Brain Sciences*, *36(1)*, 59-122. DOI: <http://doi.org/10.1017/S0140525X12000696>
- Alexander, P., Schlegel, A., Sinnott-Armstrong, W., Roskies, A., Tse, P.U., & Wheatley, T., (2015). Dissecting the Readiness Potential: An investigation of the relationship between readiness potentials, conscious willing, and action. In A. Mele (ed.) *Surrounding Free Will: Philosophy, Psychology, Neuroscience*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780199333950.003.0011>
- Alexander, P., Schlegel, A., Sinnott-Armstrong, W., Roskies, A., Wheatley, T. & Tse, P. (2016). Readiness potentials driven by non-motoric processes. *Consciousness and Cognition* *39*, 38-47. DOI: <http://doi.org/10.1016/j.concog.2015.11.011>
- Alquist, J. L., Ainsworth, S. E., & Baumeister, R. F. (2012). Determined to conform: Disbelief in free will increases conformity. *Journal of Experimental Social Psychology* *49(1)*, 80-86. DOI: <http://doi.org/10.1016/j.jesp.2012.08.015>
- Alquist, J. L., Ainsworth, S. E., Baumeister, R. F., Daly, M., & Stillman, T. F. (2014). The Making of Might-Have-Beens: Effects of Free Will Belief on Counterfactual Thinking. *Personality and Social Psychology Bulletin* *41(2)*, 268-283. DOI: <http://doi.org/10.1177/0146167214563673>
- Alquist, J. L., & Baumeister, R. F. (2010). Free to be Different: Disbelief in Free will Increases Conformity. Unpublished manuscript.
- Alquist, J., & Baumeister, R. F. (2012). Self-control: Limited resources and extensive benefits. *Wiley Interdisciplinary Reviews: Cognitive Science* *3(3)*, 419-423.
- Balauer, M. (2004). A coherent, naturalistic, and plausible formulation of libertarian free will. *Noûs*, *38(3)*, 379-406.
- Banks, William P., & Eve A. Isham. (2009). We Infer Rather Than Perceive the Moment We Decided to Act. *Psychological Science*, *20*: 17-21. DOI: [10.1111/j.1467-9280.2008.02254](https://doi.org/10.1111/j.1467-9280.2008.02254).
- Barutchu, A., Becker, S. I., Carter, O., Hester, R., & Levy, N. (2013a). The Role of Task-Related Learned Representations in Explaining Asymmetries in Task Switching. *PLoS ONE* *8(4)*, e61729. DOI: <http://doi.org/10.1371/journal.pone.0061729>
- Barutchu, A., Carter, O., Hester, R., & Levy, N. (2013b). Strength in cognitive self-regulation. *Frontiers in Psychology* *4*, 174. DOI: <http://doi.org/10.3389/fpsyg.2013.00174>
- Baumeister, R. F. (2008). Free will in scientific psychology. *Perspectives on psychological science*, *3(1)*, 14-19.
- Baumeister, R. F. (2013a). Self-control, fluctuating willpower, and forensic practice. *Journal of Forensic Practice* *15(2)*, 85-96.

- Baumeister, R.F.. (2013b). Do you really have free will?. Slate. URL: http://www.slate.com/articles/health_and_science/science/2013/09/free_will_debate_what_does_free_will_mean_and_how_did_it_evolve.html.
- Baumeister, R. F. (2014a). Constructing a scientific theory of free will. In W. Sinnott-Armstrong (ed.) *Free Will and Responsibility*. Cambridge: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262026680>
- Baumeister, R. F. (2014b). Grateful responses to thoughtful comments by Holton, Payne, and Cameron. In W. Sinnott-Armstrong (ed.) *Free Will and Responsibility*. Cambridge: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262026680>
- Baumeister, R. F. & Bargh, J. A. (2014). Conscious and unconscious: Toward an integrative understanding of human mental life and action. In Sherman, J. W., Gawronski, B., & Trope, Y. (eds.) *Dual process theories of the social mind*. Guildford Publications. ISBN: <http://www.isbnsearch.org/search?s=9781462514397>
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Self-control depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252-1265. Doi: 10.2037/0022-3514.74.5.1252
- Baumeister, R. F., & Brewer, L. E. (2012). Believing versus Disbelieving in Free Will: Correlates and Consequences. *Social and Personality Psychology Compass* *6*(10), 736-745. DOI: <http://doi.org/10.1111/j.1751-9004.2012.00458.x>
- Baumeister, R.F., Clark, C., & Luguri, J. (2014). Free will: Belief and reality. In A. Mele (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Baumeister, R. F., Crescioni, A. W., & Alquist, J. L. (2011). Free will as advanced action control for human social life and culture. *Neuroethics*, *4*, 1-11. doi:10.1007/s12152-010-9058-4
- Baumeister, R. F., Masicampo, E. J., & DeWall, C. N. (2009). Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, *35*(2), 260-268.
- Baumeister, R. F., & Monroe, A. (2014). Recent research on free will: Conceptualizations, beliefs, and processes. *Advances in Experimental Social Psychology* *50*, 1-52. DOI: <http://doi.org/10.1016/B978-0-12-800284-1.00001-1>
- Baumeister, R.F., Schmeichel, B.J., & DeWall, C.N. (2014). Creativity, Consciousness, and Free Will: Evidence from Psychology Experiments. In Paul, E.S., & Kaufman, S.B. (eds.) *The Philosophy of Creativity*. ISBN: <http://www.isbnsearch.org/search?s=9780199836963>
- Baumeister, R. F., Vohs, K. D., & Masicampo, E. J. (2014). Maybe it helps to be conscious, after all. *Behavioral and Brain Sciences* *37*(1), 20-21. DOI: <http://doi.org/10.1017/S0140525X13000630>
- Baumeister, R. F., & Vonasch, A. (2011). Self-organization as conceptual key to understanding free will. *AJOB Neuroscience* *2*(3), 44-46. DOI: <http://doi.org/10.1080/21507740.2011.584949>
- Baumeister, R. F., & Vonasch, A. J. (2012). Is the essence of morality mind perception, self-regulation, free will, or culture? *Psychological Inquiry*, *23*, 134-136. doi: 10.1080/1047840X.2012.667758
- Bertrams, A., Englert, C., Dickhäuser, O., & Baumeister, R. F. (2013). Role of self-control strength in the relation between anxiety and cognitive performance. *Emotion* *13*(4), 668-680.

- Björnsson, G. (2015). Incompatibilism and 'Bypassed' Agency. In A. Mele (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Björnsson (2016). Outsourcing the deep self: Deep self discordance does not explain away intuitions in manipulation arguments. *Philosophical Psychology* 29(5), 637-653. DOI: <http://doi.org/10.1080/09515089.2016.1150448>
- Björnsson, G., & Pereboom, D. (2014). Free Will Skepticism and Bypassing. In W. Sinnott-Armstrong (ed.) *Moral Psychology: Free Will and Moral Responsibility Vol 4*. ISBN: <http://www.isbnsearch.org/search?s=9780262525473>
- Björnsson, G., & Pereboom, D. (2016). Traditional and Experimental Approaches to Free Will and Moral Responsibility. In Sytma, J., & Buckwalter, W. (eds.) *A Companion to Experimental Philosophy*. Chichester: John Wiley and Sons. DOI: <http://doi.org/10.1002/9781118661666.ch9>
- Björnsson, G., & Persson, K. (2012). The explanatory component of moral responsibility. *Noûs*, 46(2), 326-354.
- Brass, M., & Haggard, P. (2007). To do or not to do: the neural signature of self-control. *Journal of Neuroscience*, 27(34), 9141-9145.
- Bratman, M. (1984). Two faces of intention. *The Philosophical Review*, 93(3), 375-405.
- Carey, J., & Paulhus, D. L. (2011, January). The Independence of Free will and Determinism in Judgments of Moral Responsibility. Poster presented at the 12th Annual meeting of the Society for Personality and Social Psychology, San Antonio, TX.
- Chater, N., & Oaksford, M. (2013). Programs as causal models: Speculations on mental programs and mental representation. *Cognitive Science* 37(6), 1171-1191. DOI: <http://doi.org/10.1111/cogs.12062>
- Clark, C. L., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology* 106(4), 501-513. DOI: <http://doi.org/10.1037/a0035880>
- Clarke, R. (2003). *Libertarian accounts of free will*. Oxford: Oxford University Press.
- Clarke, R. (2014a). Freedom, Responsibility, and Omitting to Act. In Palmer, D. (ed.) *Libertarian Free Will: Contemporary Debates*. New York: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199860081>
- Clarke, R. (2014b). *Omissions: Agency, Metaphysics, and Responsibility*. New York: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199347520>
- Clarke, R. (2015). Negligent Action and Unwitting Omission. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Clarke, R. (2017). Ignorance, Revision, and Commonsense. In Robichaud, P., & Wieland, J. W. (eds.) *Responsibility: The Epistemic Condition*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780198779667>
- Couenhoven, J. Predestination. Vocabulary for the Study of Religion. URL: http://referenceworks.brillonline.com/entries/vocabulary-for-the-study-of-religion/predestination-COM_00000419.
- Couenhoven, J. (2013a). Fodge-ogs and HedgeOxes. *Journal of the American Academy of Religion* 81(3), 586-591. DOI: <http://doi.org/10.1093/jaarel/lft042> (Cites: 1)

-
- Couenhoven, J. (2013b). Saint Augustine. *International Encyclopedia of Ethics*. ISBN: <http://www.isbnsearch.org/search?s=9781405186414>
- Couenhoven, J. (2013c). *Stricken by Sin, Cured by Christ*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199948697>
- Couenhoven, J. (2014). The Augustinian Background of Luther's Bondage of the Will. *On the Apocalyptic and Human Agency: Conversations With Augustine of Hippo and Martin Luther*. ISBN: <http://www.isbnsearch.org/search?s=9781443866279>
- Couenhoven, J., *Stricken by Sin, Cured by Christ*, text interview published at <http://symposiumethics.org/2015/10/13/stricken-by-sin-cured-by-christ/>, October 2015a.
- Couenhoven, J. (2015b). Can Free Will Survive God's Foreknowledge? on PBS's *Closer to Truth*, Season 14, Episode 12 (2015).
- Couenhoven, J. (2015c). Is God Responsible for Evil? on PBS's *Closer to Truth*, Season 14, Episode 2 (2015).
- Couenhoven, J. (2016a). Augustine of Hippo. In Timpe, K., Griffith, M., & Levy, N. (eds.) *The Routledge Companion to Free Will*. New York: Routledge. ISBN: <http://www.isbnsearch.org/search?s=9781138795815>
- Couenhoven, J. (2016b). Augustine. In Johnson, K. L., & Lauber, D. (eds.) *T & T Clark Companion to the Doctrine of Sin*. T & T Clark. ISBN: <http://www.isbnsearch.org/search?s=9780567451156>
- Couenhoven, J. (2016c). The Indicative in the Imperative: On Augustinian Oughts and Cans. McCann, H. J. (ed.) *Free Will and Classical Theism*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780190611200.001.0001>
- Couenhoven, J. (2016d) The Problem of God's Immutable Freedom. In Timpe, K., & Speak, D. (eds.) *Free Will and Theism - Connections, Contingencies, and Concerns*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780198743958>
- Couenhoven, J. (2017). The Protestant Reformation. In Golob, S., & Timmermann, J. (eds.) *Cambridge History of Moral Philosophy*. Cambridge: Cambridge University Press. ISBN: <http://www.isbnsearch.org/search?s=9781107033054>
- Crane, T. (2014). *The Mechanical Mind. 3rd edition*. New York: Routledge.
- Crescioni, A. W., Baumeister, R. F., Ainsworth, S. E., Ent, M., & Lambert, N. M. (2015). Subjective correlates and consequences of belief in free will. *Philosophical Psychology* 29(1), 41-63. DOI: <http://doi.org/10.1080/09515089.2014.996285>
- Daly, M., Baumeister, R. F., Delaney, L., & MacLachlan, M. (2014). Self-control and its relation to emotions and psychobiology: Evidence from a Day Reconstruction Method study. *Journal of Behavioral Medicine* 37(1). 81-93. DOI: <http://doi.org/10.1007/s10865-012-9470-9>
- de Lange, F. P., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *Journal of Neuroscience* 33(4), 1400-1410. DOI: <http://doi.org/10.1523/JNEUROSCI.1094-12.2013>
- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge: Cambridge University Press.
- Ent, M. R., & Baumeister, R. F. (2014). Embodied free will beliefs: Some effects of physical states on metaphysical opinions. *Consciousness and Cognition* 27, 147-154. DOI: <http://doi.org/10.1016/j.concog.2014.05.001>

- Feldman, G., Baumeister, R. F., & Wong, K. F. E. (2014). Free will is about choosing: The link between choice and the belief in free will. *Journal of Experimental Social Psychology* 55, 239-245. DOI: <http://doi.org/10.1016/j.jesp.2014.07.012>
- Filevich, E., & Haggard, P. (2013). Persistence of internal representations of alternative voluntary actions. *Frontiers in Psychology* 4, 202. DOI: <http://doi.org/10.3389/fpsyg.2013.00202>
- Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. New York: Cambridge University Press.
- Fischer, J. M., & Tognazzini, N. (2011). The physiognomy of responsibility. *Philosophy and Phenomenological Research*, 82(2), 381-417.
- Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The journal of philosophy*, 66(23), 829-839.
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 5-20.
- Franklin, C. E. (2011). The problem of enhanced control. *Australasian Journal of Philosophy*, 89(4), 687-706.
- Franklin, C. E. (2013). The scientific plausibility of libertarianism. In I. Haji & J. Caouette (eds.), *Free will and moral responsibility* (pp. 123-141). Cambridge: Cambridge Scholars Publishing.
- Fried, I., Mukamel, R., & Kreiman, G. (2011). "Internally Generated Preactivation of Single Neurons in Human Medial Frontal Cortex Predicts Volition." *Neuron*, 69, no. 3: 548-562. DOI: [10.1016/j.neuron.2010.11.045](http://doi.org/10.1016/j.neuron.2010.11.045)
- Furstenberg (2014). Proximal Intentions, Non-executed Proximal Intentions and Change of Intentions. *Topoi* 33(1), 13-22. DOI: <http://doi.org/10.1007/s11245-013-9191-x>
- Furstenberg, A., Breska, A., Deouell, L. Y., & Sompolinsky, H. (2013). Neural signature of change of intentions. *Journal of Molecular Neuroscience* 51(1), 1-135. DOI: <http://doi.org/10.1007/s12031-012-9923-1>
- Furstenberg, A., Deouell, L. Y., & Sompolinsky, H. (2014). Abstracts of the 22nd Annual Meeting of the Israel Society for Neuroscience (ISFN) & The 2nd Bi national Italy-Israel Neuroscience Meeting, December 14-17, 2013, Eilat, Israel. *Journal of molecular neuroscience* : MN. DOI: <http://doi.org/10.1007/s12031-014-0252-4>
- Furstenberg, A., Breska, A., Sompolinsky, H., & Deouell, L. Y. (2015a). Evidence of change of intention in picking situations. *Journal of Cognitive Neuroscience* 27(11), 2133-2146. DOI: http://doi.org/10.1162/jocn_a_00842
- Furstenberg, A., Deouell, L. Y., & Sompolinsky, H. (2015b). Change of Intention in 'Picking' Situations. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., ... & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *Journal of personality and social psychology*, 92(2), 325.
- Gailliot, M. T., Gitter, S. A., Baker, M. D., & Baumeister, R. F. (2012). Breaking the rules: Low trait or state self-control increases social norm violations. *Psychology*, 3(12), 1074.
- Glannon, W. (2012a). Obsessions, compulsions, and free will. *Philosophy, Psychiatry, & Psychology*, 19(4), 333-337.

- Glannon, W. (2012b). Neurophysiological Aspects of Enhancing the Will. *The Monist* 95(3), 378-398. DOI: <http://doi.org/10.5840/monist201295320>
- Glannon, W. (2012c). Review of Sean A. Spence, *The Actor's Brain: Exploring the Cognitive Neuroscience of Free Will*. *AJOB Neuroscience* 3(4), 93-95. DOI: <http://doi.org/10.1080/21507740.2012.713073>.
- Glannon, W. (2013). Neuromodulation, agency and autonomy. *Brain Topography* 27(1), 46-54. DOI: <http://doi.org/10.1007/s10548-012-0269-3>
- Glannon (2014a). Intervening in the psychopath's brain. *Theoretical Medicine and Bioethics* 35(1), 43-57. DOI: <http://doi.org/10.1007/s11017-013-9275-z>
- Glannon (2014b). *The Free Will Fix: How New Brain Implants Can Restore and Enhance Free Will*. Aeon. URL: <https://aeon.co/essays/how-new-brain-implants-can-restore-and-enhance-free-will>.
- Glannon, W. (2015). (ed.), *Free Will and the Brain: Neuroscientific, Philosophical and Legal Perspectives*. Cambridge, UK: Cambridge University Press.
- Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. *European review of social psychology*, 4(1), 141-185.
- Gopnik, A., & Kushnir, T. (2014). The Origins and Development of Our Conception of Free Will. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological inquiry*, 23(2), 101-124.
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1451), 1775-1785.
- Grimaldi, M.A., Lau, H., & Basso, P. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. *Neuroscience and Biobehavioral Reviews* 55, 88-97. DOI: <http://doi.org/10.1016/j.neubiorev.2015.04.006>
- Gröpel, J., Baumeister, R. F., & Beckmann, P. (2014). Action Versus State Orientation and Self-Control Performance After Depletion. *Personality and Social Psychology Bulletin* 40(4), 476-487. DOI: <http://doi.org/10.1177/0146167213516636>
- Grant, W. M. (2015). The privation account of moral evil: A defense. *International Philosophical Quarterly* 55(3), 271-286. DOI: <http://doi.org/10.5840/ipq201572139>
- Grant, W. M. (2016). Divine Universal Causality without Occasionalism (and with Agent Causality). In Timpe, K., & Speak, D. (eds.) *Free Will and Theism*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780198743958>
- Haggard, P., Mele, A., O'Connor, T., & Vohs, K. (2015). Free Will Lexicon. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Haggard, P., & Eimer, M. (1999). On the Relation between Brain Potentials and the Awareness of Voluntary Movements. *Experimental Brain Research* 126, no. 1: 128-133.
- Hallett, M. (2011). Volition: How Physiology Speaks to the Issue of Responsibility. In Sinnott-Armstrong, W., & Nadel, L. (eds.) *Conscious Will and Responsibility*. Oxford: Oxford University Press.

- Hasker, W. (2011). Divine Knowledge and Human Freedom. In R. Kane (ed.). *The Oxford Handbook of Free Will*. Oxford: Oxford University Press.
- Hofmann, W., Baumeister, R. F., Forester, G., & Vohs, K. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology* 102(6), 1318-1335.
- Hofmann, W., Luhmann, M., Fisher, R. R., Vohs, K. & Baumeister, R.F (2013). Yes, But Are They Happy? Effects of Trait Self-Control on Affective Well-Being and Life Satisfaction. *Journal of Personality* 82(4), 265-277. DOI: <http://doi.org/10.1111/jopy.12050>
- Hofmann, W., Vohs, K. D., & Baumeister, R. F. (2011). What People Desire, Feel Conflicted About, and Try to Resist in Everyday Life. *Psychological Science* 23(6), 582-588. DOI: <http://doi.org/10.1177/0956797612437426>
- Hunt, D. (2013). God's Extended Mind. *European Journal for Philosophy of Religion*. URL: <http://www.philosophy-of-religion.eu/contents11.html>
- Hunt, D. (2016). Theological Fatalism as an Aporetic Problem. In McCann, H. J. (ed.) *Free Will and Classical Theism: The Significance of Freedom in Perfect Being Theology*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780190611200>
- Hunt, D., & Shabo, S. (2013). Frankfurt cases and the (in) significance of timing: A defense of the buffering strategy. *Philosophical Studies*, 164(3), 599-622.
- Ismael, J. (2013). Causation, Free Will, and Naturalism1. In Ross, D., Ladyman, J., & Kincaid, H. (eds.) *Scientific Metaphysics*. Oxford; Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780199696499.003.0010>
- Ismael, J. (2015). On Being Some-One. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Ismael, J. (2016a). How do causes depend on us? The many faces of perspectivalism. *Synthese* 193(1), 245-267. DOI: <http://doi.org/10.1007/s11229-015-0757-6>
- Ismael, J. (2016b). From Physical Time to Human Time. In Delev, Y., & Roubach, M. (eds.) *Cosmological and Psychological Time*. New York: Springer. ISBN: <http://www.isbnsearch.org/search?s=9783319225890>
- Ismael, J. (2016c). *How Physics Makes Us Free*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780190269449>
- Ismael, J. (2017a). An Empiricist's Guide to Objective Modality. In Slater, M. H., & Yodell, Z. *Metaphysics and the Philosophy of Science*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199363209>
- Ismael, J. (2017b). Passage, Flow, and the Logic of Temporal Perspectives. *Time of Nature and the Nature of Time*. ISBN: <http://www.isbnsearch.org/search?s=9783319537252>
- Ismael, J (2017c). On Whether the Atemporal Conception of the World Is Also Amodal. *Analytic Philosophy*. DOI: <http://doi.org/10.1111/phib.12062>
- Jo, H-G, Hinterberger, T., Wittmann, M., and Schmidt, S. (2015). Do Meditators have Higher Awareness of their Intentions to Act?. *Cortex*, 65: 149-158. DOI: 10.1016/j.cortex.2014.12.015
- Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego Depletion: Is it All in Your Head? Implicit Theories about Willpower Affect Regulation. *Psychological Science* 21(11), 1686-1693.
- Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.

- Kane, R. (2014). New arguments in debates on libertarian free will: Responses to contributors. In Palmer, D. (ed.) *Libertarian free will: Contemporary debates*. Oxford: Oxford University Press.
- Knobe, J. (2013). Free will and the scientific vision. In Machery, E., & O'Neill, E. (eds.) *Current Controversies in Experimental Philosophy*. New York: Routledge. ISBN: <http://www.isbnsearch.org/search?s=9780415519663>
- Ko, Y., & Lau, H. (2012). A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1594), . DOI: <http://doi.org/10.1098/rstb.2011.0380>
- Koch, C. (2012). *Consciousness: Confessions of a romantic reductionist*. MIT press.
- Kohler, P. J., Cavanagh, P., & Tse, P. (2015). Motion-induced position shifts are influenced by global motion, but dominated by component motion. *Vision Research* 110(A), 93-99. DOI: <http://doi.org/10.1016/j.visres.2015.03.003>
- Kohler, P.J., Cavanagh, P., & Tse, P. (2017). Motion-induced position shifts activate early visual cortex. *Frontiers in Neuroscience* 11, 168. DOI: <http://doi.org/10.3389/fnins.2017.00168>
- Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 284(1), 1-17.
- Koizumi, A., Maniscalco, B., & Lau, H. (2015). Does perceptual confidence facilitate cognitive control?. *Attention, Perception, and Psychophysics* 77(4), 1295-1306. DOI: <http://doi.org/10.3758/s13414-015-0843-3>
- Kushnir, T., Gopnik, A., Chernyak, N., Seiver, E., & Wellman, H. M. (2015). Developing intuitions about free will between ages four and six. *Cognition* 138, 79-101. DOI: <http://doi.org/10.1016/j.cognition.2015.01.003>
- Leftow, B. (2013). God's deontic perfection. *Res Philosophica* 90(1), 69-95. DOI: <http://doi.org/10.11612/resphil.2013.90.1.5>
- Leftow, B. (2014). Tempting god. *Faith and Philosophy* 31(1), 3-23. DOI: <http://doi.org/10.5840/faithphil201312301>
- Leftow, B. (2015). Divine Simplicity and Divine Freedom. *Proceedings and Addresses of the ACPA* 89, 45-56. DOI: <http://doi.org/10.5840/acpaprocc20172661>
- Leftow, B. (2016). Two Pictures of Divine Choice. In McCann, H. J. (ed.) *Free Will and Classical Theism*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780190611200.003.0009>
- Lau, H. C, Rogers, R. D., & Passingham, R. E. (2007). Manipulating the Experienced Onset of Intention after Action Execution. *Journal of Cognitive Neuroscience*, 19: 81-90. DOI: [10.1162/jocn.2007.19.1.81](http://doi.org/10.1162/jocn.2007.19.1.81)
- Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences* 15(8), 365-373. DOI: <http://doi.org/10.1016/j.tics.2011.05.009>
- Levy, N. (2011). *Hard luck: How luck undermines free will and moral responsibility*. Oxford: Oxford University Press.
- Levy, N. (2014). Consciousness, implicit attitudes and moral responsibility. *Noûs*, 48(1), 21-40.
- Levy, N. (2016a). Addiction, Autonomy, and Informed Consent: On and off the Garden Path. *Journal of Medicine and Philosophy* 41(1), 56-73. DOI: <http://doi.org/10.1093/jmp/jhv033>

- Levy (2016b). The sweetness of surrender: Glucose enhances self-control by signaling environmental richness. *Philosophical Psychology* 29(6), 813-825. DOI: <http://doi.org/10.1080/09515089.2016.1173202>
- Levy, N. (2017). Of Marshmallows and Moderation. In Sinnott-Armstrong, W., and Miller, C. (eds.) *Moral Psychology: Virtue and Happiness*. Cambridge: MIT Press. DOI: <http://doi.org/10.2307/j.ctt1n2tvzm>
- Libet, B. (1985). "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action." *The Behavioral and Brain Sciences*, 8: 529-566.
- Libet, B. (1994). "A Testable Field Theory of Mind-Brain Interaction." *Journal of Consciousness Studies*, 1: 119-126.
- Libet, B. (1999). "Do We have Free Will?." *Journal of Consciousness Studies*, 6, (8-9): 47-57.
- Libet, B. (2005). "Do We Have Free Will?" In *The Oxford Handbook of Free Will*, edited by Robert Kane, 551-564. Oxford: Oxford University Press.
- Libet, B., E. W. Wright, & C. A. Gleason. (1982). "Readiness Potentials Preceding Unrestricted Spontaneous Pre-Planned Voluntary Acts." *Electroencephalography and Clinical Neurophysiology*, 54: 322-335.
- Libet, Benjamin, Curtis A. Gleason, Elwood W. Wright, & Dennis K. Pearl. (1983). "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential)." *Brain* 106 (3): 623-642.
- Lipsman, N. & Glannon, W. (2013). Brain, mind and machine: What are the implications of deep brain stimulation for perceptions of personal identity, agency and free will?. *Bioethics* 27(9), 465-470. DOI: <http://doi.org/10.1111/j.1467-8519.2012.01978.x>
- Maoz, U., Mudrik, L., Rivlin, R., Ross, I. B., Mamelak, A. N., & Yaffe, G. (2014). On reporting the onset of the intention to move. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Maoz, U., Rutishauser, U., Kim, S., Cai, X., Lee, D., & Koch, C. (2013). Predeliberation activity in prefrontal cortex and striatum and the prediction of subsequent value judgment. *Frontiers in Neuroscience* 7, 225. DOI: <http://doi.org/10.3389/fnins.2013.00225>
- Maoz, U. & Yaffe, G. (2014). Cognitive Neuroscience and Criminal Responsibility. In Gazzaniga, M. (ed.) *The Cognitive Neurosciences*. Cambridge: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262027779>
- Maoz, U., Ye, S., Ross, I. B., Mamelak, A. N., & Koch, C. (2012). Predicting Action Content On-Line and in Real Time before Action Onset - an Intracranial Human Study. *Advances in Neural Information Processing Systems (NIPS)*. URL: <http://papers.nips.cc/paper/4513-predicting-action-content-on-line-and-in-real-time-before-action-onset-an-intracranial-human-study.pdf>
- Maniscalco, B., Bang, J.W., Iravani, L., Camps-Febrer, F., & Lau, H. (2012). Does response interference depend on the subjective visibility of flanker distractors?. *Attention, Perception, and Psychophysics* 74(5), 841-851. DOI: <http://doi.org/10.3758/s13414-012-0291-2>
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition* 21(1), 422-430. DOI: <http://doi.org/10.1016/j.concog.2011.09.021>

- Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: Meta-d', response-specific meta-d', and the unequal variance SDT model. In Fleming, S., & Frith, C. (ed.) *The Cognitive Neuroscience of Metacognition*. Berlin: Springer. DOI: http://doi.org/10.1007/978-3-642-45190-4_3
- Maniscalco, B., & Lau, H. (2015). Manipulation of working memory contents selectively impairs metacognitive sensitivity in a concurrent visual discrimination task. *Neuroscience of Consciousness 2015(1)*, niv002. DOI: <http://doi.org/10.1093/nc/niv002>
- Maniscalco, B., & Lau, H. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience and Consciousness 2016(1)*, niw 002. DOI: <http://doi.org/10.1093/nc/niw002>
- Maniscalco, B., Peters, M. A. K., & Lau, H. (2016). Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Attention, Perception, and Psychophysics 78(3)*, 923-937. DOI: <http://doi.org/10.3758/s13414-016-1059-x>
- Martin, N. D., Rigoni, D., & Vohs, K. D. (2017). Free will beliefs predict attitudes toward unethical behavior and criminal punishment. *Proceedings of the National Academy of Sciences, 114(28)*, 7325-7330.
- Masicampo, E. J., & Baumeister, R. F. (2012). Committed but closed-minded: When making a specific plan for a goal hinders success. *Social Cognition 30(1)*, 37-55. DOI: <http://doi.org/10.1521/soco.2012.30.1.37>
- Masicampo, E. J., & Baumeister, R. F. (2013). Conscious thought does not guide moment-to-moment actions-it serves social and cultural functions. *Frontiers in Psychology 4*, 478. DOI: <http://doi.org/10.3389/fpsyg.2013.00478>
- McCann, H. (1998). *The Works of Agency*. Ithaca: Cornell University Press.
- McCann, H. J. (2012). Making decisions. *Philosophical Issues, 22(1)*, 246-263.
- McCann, H. J. (2013a). Action individuation. In E. Lapore, and Ludwig, K. (eds.) *A Companion to Donald Davidson*, Oxford: Wiley.
- McCann, H. J. (2013b). Action. *International Encyclopedia of Ethics*.
- MacKenzie, M. J., Vohs, K. D., & Baumeister, R. F. (2014). You Didn't Have to Do That: Belief in Free Will Promotes Gratitude. *Personality and Social Psychology Bulletin 40(11)*, 1423-1434. DOI: <http://doi.org/10.1177/0146167214549322>
- Mele, A. R. (2006). *Free Will and Luck*. New York: Oxford University Press.
- Mele, A. R. (2009). *Effective Intentions: The Power of Conscious Will*. Oxford: Oxford University Press.
- Mele, A. R. (2010a). Conscious deciding and the science of free will. In Baumeister, R. F., Mele, A., & Vohs, K. (eds.) *Free Will and Consciousness, How Might they Work*. Oxford: Oxford University Press.
- Mele, A. (2010b). Scientific Skepticism about Free Will. In Nadelhoffer, T., Nahmias, E., & Nichols, S. (eds.) *Moral Psychology: Classical and Contemporary Readings*. Oxford: Wiley-Blackwell. ISBN: <http://www.isbnsearch.org/search?s=9781405190206>
- Mele, A. R. (2010c). Libet on free will: readiness potentials, decisions, and awareness. In Sinnott-Armstrong, W., & Nadel, L. (eds.) *Conscious Will and Responsibility: A Tribute to Benjamin Libet*, 23. Oxford: Oxford University Press.

- Mele, A. (2010d). Moral responsibility for actions: epistemic and freedom conditions, *Philosophical Explorations* 13(2), 101-111. DOI: 10.1080/13869790903494556
- Mele, A. (2010e). Testing free will. *Neuroethics* 3(2), 161-172. DOI: <http://doi.org/10.1007/s12152-008-9027-3>
- Mele, A. (2010f). Moral responsibility and the continuation problem. *Philosophical Studies* 162(2), 237-255. DOI: <http://doi.org/10.1007/s11098-011-9757-7>
- Mele, A. (2010g). Conscious Intentions. In Campbell, J. K., O'Rourke, M., & Silverstein, H. (eds.) *Action, Ethics, and Responsibility*. Cambridge: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9781443847889>
- Mele, A. (2011a). Surrounding free will: A response to Baumeister, Crescioni, and Alquist. *Neuroethics* 4(1), 25-29. DOI: <http://doi.org/10.1007/s12152-010-9094-0>
- Mele, A. (2011b). Chance, choice and freedom. *The Philosophers' Magazine* 55(4), 61-65. DOI: <http://doi.org/10.5840/tpm201155107>
- Mele, A. (2011c). Moral Psychology. In Miller, C. (ed.) *The Continuum Companion to Ethics*. London: Bloomsbury. ISBN: <http://www.isbnsearch.org/search?s=9781441121752>
- Mele, A. (2011d). Self-Control in Action. In Gallagher, S. (ed.) *The Oxford Handbook of the Self*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199679546>
- Mele, A. R. (2012a). Crimes of Negligence: Attempting and Succeeding. *Criminal Law and Philosophy*, 6(3), 387-398.
- Mele, A. R. (2012b). Folk Conceptions of Intentional Action. *Philosophical Issues*, 22(1), 281-297.
- Mele (2012c). Free Will and Science. In Kane, R. (ed.) *The Oxford Handbook of Free Will: Second Edition*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/oxfordhb/9780195399691.003.0026>
- Mele, A. (2012d). Autonomy and Neuroscience. In Radoilska, L. (ed.) *Autonomy and Mental Health*. New York: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199595426>
- Mele, A. (2012e). Consciousness in Action: Free Will, Moral Responsibility, Data, and Inferences. In *Cognition, Reasoning, Emotion and Action*. ISBN: <http://www.isbnsearch.org/search?s=9788498607468>
- Mele, A. (2012f). When Are We Self-Deceived?. *Humana.Mente* 20. URL: http://www.humanamente.eu/PDF/Issue20_Paper_Mele.pdf
- Mele, Alfred, (2012g). Another Scientific Threat to Free Will? *The Monist* 95(3), 422-440. DOI: <http://doi.org/10.5840/monist201295322>
- Mele, A. R. (2012h). Intentional, unintentional, or neither? Middle ground in theory and practice. *American Philosophical Quarterly*, 49(4), 369-379.
- Mele, A. (2012i). Free Will and Neuroscience: Revisiting Libet's Studies. In Suarez, A., & Adams, P. (eds.) *Is Science Compatible with Free Will?*. London: Springer. ISBN: <http://www.isbnsearch.org/search?s=9781461452119>
- Mele, A. (2012j). Premium, Mid-Grade and Regular Free Will Fuel Accountability. *Phi Kappa Phi Forum*. URL: <https://www.questia.com/magazine/1P3-2602337391/premium-mid-grade-and-regular-free-will-fuel-accountability>.

- Mele, A. (2012k). Review: Responsibility without Free Will (Review of Reviewed Work: Who's in Charge: Free Will and the Science of the Brain. by Michael S. Gazzaniga). *BioScience* 62(3), 313-314. DOI: <http://doi.org/10.1525/bio.2012.62.3.15>
- Mele, A. (2012l). *Backsliding: Understanding Weakness of Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199896134>
- Mele, A. (2012m). The Case Against the Case Against Free Will. *The Chronicle of Higher Education*. URL: <http://chronicle.com/article/Alfred-R-Mele-The-Case/131166/>.
- Mele, A. (2012n). Autonomie, moralische Verantwortung und das Fortsetzungsproblem. In Nisa-Rumelin, J., & Ozman, E. (eds.) *Welt der Gründe*. Hamburg: Felix Meiner Verlag. ISBN: <http://www.isbnsearch.org/search?s=9783787322695>
- Mele, A. (2013a). Free Will. In *International Encyclopedia of Ethics*, H. LaFollette (Ed.). DOI:10.1002/9781444367072.wbiee159
- Mele (2013b). Libertarianism and Human Agency. *Philosophy and Phenomenological Research* 87(1), 72-92. DOI: <http://doi.org/10.1111/j.1933-1592.2011.00529.x>
- Mele (2013c). Manipulation, Moral Responsibility, and Bullet Biting. *The Journal of Ethics* 17(3), 167-184. DOI: <http://doi.org/10.1007/s10892-013-9147-9>
- Mele, A. (2013d). Moral Responsibility, Manipulation, and Minutelings. *The Journal of Ethics* 17(3), 153-166. DOI: <http://doi.org/10.1007/s10892-013-9148-8>
- Mele, A. (2013e). Self-control, motivational strength, and exposure therapy. *Philosophical Studies*. DOI: <http://doi.org/10.1007/s11098-013-0224-5>
- Mele, A. (2013f). Unconscious decisions and free will. *Philosophical Psychology* 170(2), 359-375. DOI: <http://doi.org/10.1080/09515089.2012.724395>
- Mele, A. (2013g). Actions, Explanations, and Causes. In G. D'Oro (ed.), *Reasons and Causes: Causalism and Non-causalism in the Philosophy of Action*. Palgrave-Macmillan.
- Mele, A. (2013h). Agency. In B. Kaldis (ed.) *Encyclopedia of Philosophy and Social Sciences*. Oxford: Sage Publications.
- Mele, A. (2013i). Free Will and Neuroscience. *Philosophic Exchange*. URL: http://digitalcommons.brockport.edu/phil_ex/vol43/iss1/3/
- Mele, A. (2013j). Free Will, Science, and Punishment. In Nadelhoffer, N. (ed.) *The Future of Punishment*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199779208>
- Mele, A. (2013k). Is What You Decide Ever up to You?. In Caouette, J. (ed.) *Free Will and Moral Responsibility*. Cambridge: Cambridge Scholars Publishing. ISBN: <http://www.isbnsearch.org/search?s=9781443850520>
- Mele, A. (2013l). Vetoing and Consciousness. In Clark, A., Kiverstein, J., & Vertkant, K. (eds.) *Decomposing the Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199746996>
- Mele, A. (2013m). Author Q & A: Alfred Mele Explains How We Act Against Our Better Judgements. *The Philosophers' Magazine* 60(1), 125-126. DOI: <http://doi.org/10.5840/tpm20136033>.
- Mele, A. R. (2014a). The Single Phenomenon view and experimental philosophy. In M. Vargas, and G. Yaffe (eds.). *Rational and Social Agency: The Philosophy of Michael Bratman*, New York: Oxford University Press.

- Mele, A. R. (2014b). *A Dialogue on Free Will and Science*. New York: Oxford University Press.
- Mele, A. (2014c). *Free: Why Science Hasn't Disproved Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199371624>
- Mele, A. (2014d). Kane, Luck, and Control. In Palmer, D. (ed.) *Libertarian Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199860081>
- Mele, A. (2014e). Reply to Nadelhoffer and Vargas. In Sinnott-Armstrong, W. (ed.) *Moral Psychology*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262026680>
- Mele, A. R. (2014f). Free Will and Substance Dualism: The Real Scientific Threat to Free Will?. In Sinnott-Armstrong, W. (ed.) *Moral Psychology, Volume 4*. Cambridge, MA: MIT Press. DOI: <http://doi.org/10.7551/mitpress/9780262026680.003.0006>
- Mele, A. (2015a). Luck and Free Will. In Pritchard, D., & Whittington, L. J. (eds.) *The Philosophy of Luck*. Oxford: Wiley-Blackwell. ISBN: <http://www.isbnsearch.org/search?s=9781119030577>
- Mele, A. (2015b). Free will and moral responsibility: does either require the other?. *Philosophical Explorations*, 18(3), 297-309.
- Mele (2015c). Libertarianism, compatibilism, and luck. *Journal of Ethics* 19(1), 1-21. DOI: <http://doi.org/10.1007/s10892-014-9188-8>
- Mele (2015d). Luck, control, and free will: Answering Berofsky. *Journal of Philosophy* 112(7), 337-355. DOI: <http://doi.org/10.5840/jphil2015112721>
- Mele, A. (2016a). Moral Responsibility: Radical Reversals and Original Designs. *Journal of Ethics* 20(1-3), 69-82. DOI: <http://doi.org/10.1007/s10892-016-9222-0>
- Mele, A. (2016b). Character in Action. In I. Fileva (ed.). *Questions of Character*, Oxford: Oxford University Press.
- Mele, A. (2016c). Libet and the Timing of the Will. In Clausen, J., & Levy, N. (eds.) *Handbook of Neuroethics*. London: Springer. ISBN: <http://www.isbnsearch.org/search?s=9789400747081>
- Mele (2017a). Direct control. *Philosophical Studies* 174(2), 275-290. DOI: <http://doi.org/10.1007/s11098-016-0681-8>
- Mele, A. (2017b). On Pereboom's Disappearing Agent Argument. *Criminal Law and Philosophy* 11(3), 561-574. DOI: <http://doi.org/10.1007/s11572-015-9374-1>
- Mele, A. (2017c). Two Libertarian Theories: or Why Event-causal Libertarians Should Prefer My Daring Libertarian View to Robert Kane's View. *Philosophy* 80, 49-68 . DOI: <http://doi.org/10.1017/S1358246117000108>
- Mele, A. (2017d) *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780190659974>
- Mele, A. (2018a). Free Will and Consciousness. In Jacqueline, D. (ed.) *The Bloomsbury Companion to the Philosophy of Consciousness*. London: Bloomsbury. ISBN: <http://www.isbnsearch.org/search?s=9781474229036>
- Mele, A. (2018b). Free Will, Moral Responsibility, and Scientific Epiphenomenalism. *Frontiers in Psychology* 9, 2536.
- Mele, A. (2019). *Manipulated Agents: A Window to Moral Responsibility*. Oxford: Oxford University Press.
- Mele, A., & Kearns, S. (2014). Have Compatibilists Solved the Luck Problem for Libertarians?. *Philosophical Inquiries*. URL: <https://www.philing.it/index.php/philing/article/view/97>

- Mele, A. R., & Shepherd, J. (2013). Situationism and agency. *Journal of Practical Ethics*, 1(1).
- Mele, A., Vohs, K., & Baumeister, R. (2010). Free Will and Consciousness: An Introduction and Overview of Perspectives. In Baumeister, R., Mele, A., & Vohs, K. (eds.) *Free Will and Consciousness: How Might They Work?*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1080/20403313.2016.1144745>
- Miles, J. B. (2011). 'Irresponsible and a Disservice': The integrity of social psychology turns on the free-will dilemma. *British Journal of Social Psychology* 52(2), 205-218. Advance online publication. DOI: 10.1111/j.2044-8309.2011.02077.x
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological review*, 80(4), 252.
- Monroe, A. E. (2014). Free will without metaphysics. *In-Mind*. URL: <http://www.in-mind.org/blog/post/free-will-without-metaphysics>.
- Monroe, A. E., Dillon, K. D. & Malle, B. F. (2014). Bringing free will down to Earth: People's psychological concept of free will and its role in moral judgment. *Consciousness and Cognition* 27, 100-108. DOI: <http://doi.org/10.1016/j.concog.2014.04.011>
- Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about people's folk concept of free will. *Review of Philosophy and Psychology*, 1(2), 211-224.
- Monroe, A. E. and Malle, B. F. (2014). Free will without metaphysics. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Monroe, A. E., Vohs, K. D., & Baumeister, R. F. (2016). Free will evolved for morality and culture. In Miller, A. G. (ed.) *The Social Psychology of Good and Evil (2nd ed)*. The Guilford Press. ISBN: <http://www.isbnsearch.org/search?s=9781462525393>
- Morales, J., Chang, J., & Lau, H. (2015a). Controlling for performance capacity confounds in neuroimaging studies of conscious awareness. *Neuroscience of Consciousness* 1, niv008. DOI: <http://doi.org/10.1093/nc/niv008>
- Morales, J., Solovey, G., Maniscalco, B., Rahnev, D., de Lange, F. P., & Lau, H. (2015b). Low attention impairs optimal incorporation of prior knowledge in perceptual decisions. *Attention, Perception, and Psychophysics* 77(6), 2021-2036. DOI: <http://doi.org/10.3758/s13414-015-0897-2>
- Mudrik, L., & Maoz, U. (2014). "Me & My Brain": Exposing Neuroscience's Closet Dualism. *Journal of Cognitive Neuroscience* 27(2), 211-221. DOI: http://doi.org/10.1162/jocn_a_00723
- Murphy, M. (2014). Perfect Goodness. In the *Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2014/entries/perfect-goodness/>.
- Nadelhoffer, T. A. (Ed.). (2013). *The Future of Punishment*. Oxford: Oxford University Press.
- Nadelhoffer, T. (2014). Dualism, Libertarianism, and Scientific Skepticism about Free Will. In Sinnott-Armstrong, W. (ed.) *Moral Psychology: Free Will and Moral Responsibility*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262525473>
- Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C. & Ross, L. T. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition* 25, 27-41. DOI: <http://doi.org/10.1016/j.concog.2014.01.006>

- Nahmias, E. (2011). Is Neuroscience the Death of Free Will?. *New York Times*. URL: <http://opinionator.blogs.nytimes.com/2011/11/13/is-neuroscience-the-death-of-free-will>.
- Nahmias, E. (2012). Does contemporary neuroscience support or challenge the reality of free will?. Templeton Big Questions Online. URL: <https://www.bigquestionsonline.com/content/does-contemporary-neuroscience-support-or-challenge-reality-free-will>.
- Nahmias, E. (2014). Is Free Will an Illusion? Confronting Challenges from the Modern Mind Sciences. In Sinnott-Armstrong, W. (ed.) *Moral Psychology*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262525473>
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, 18(5), 561-584.
- Nahmias, E., Morris, S. G., Nadelhoffer, T., & Turner, J. (2006). Is incompatibilism intuitive?. *Philosophy and Phenomenological Research*, 73(1), 28-53.
- Nahmias, E., Coates, D. J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: experiments on folk intuitions. *Midwest Studies in Philosophy*, 31, 214-232.
- Nahmias, E. & Thompson, M. (2014). A Naturalistic Vision of Free Will. In Machery, E., & O'Neill, E. (eds.) *Current Controversies in Experimental Philosophy*. New York: Routledge. ISBN: <http://www.isbnsearch.org/search?s=9780415519670>
- Olivola, C. Y. & Shafir, E. (2011), The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *J. Behav. Dec. Making*, 26: 91-105. DOI:10.1002/bdm.767
- O'Connor, T. (2000). *Persons and Causes*. New York: Oxford University Press.
- Pacherie, E. (2006). Towards a dynamic theory of intentions. In Pockett, S., Banks, W. P., and Gallagher, S. (eds.) *Does Consciousness Cause Behavior*. Cambridge, MA: MIT Press.
- Parkinson, J., & Haggard, P. (2014). Subliminal priming of intentional inhibition. *Cognition* 130(2), 255-265. DOI: <http://doi.org/10.1016/j.cognition.2013.11.005>
- Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Pereboom, D. (2015). Omissions and different senses of responsibility. In Buckareff, A., Moya, C., and Rosell, S. (eds.) *Agency, Freedom, and Moral Responsibility*. London: Palgrave Macmillan.
- Pérez de Calleja, M. (2014). Cross-world luck at the time of decision is a problem for compatibilists as well. *Philosophical Explorations*, 17(2), 112-125.
- Peters, MAK, Fesi, J., Amendi, N., Knotts, J. D., Lau, H., & Ro, T. (2017a). Transcranial magnetic stimulation to visual cortex induces suboptimal introspection. *Cortex* 93, 119-132. DOI: <http://doi.org/10.1016/j.cortex.2017.05.017>
- Peters, MAK, Thesen, T., Ko, Y. D., Maniscalco, B., Carlson, C., Davidson, M., Doyle, W., Kuzniecky, R., Devinsky, O. Halgren, E., & Lau, H. (2017b). Perceptual confidence neglects decision-incongruent evidence in the brain. *Nature Human Behaviour*. DOI: <http://doi.org/10.1038/s41562-017-0139>
- Phillips, J., & Shaw, A. (2014). Manipulating Morality: Third-Party Intentions Alter Moral Judgments by Changing Causal Reasoning. *Cognitive Science* 39(6), 1320-1347. DOI: <http://doi.org/10.1111/cogs.12194>

- Rahnev, D. A., Huang, E., Lau, H. (2012a). Subliminal stimuli in the near absence of attention influence top-down cognitive control. *Attention, Perception, and Psychophysics*. DOI: <http://doi.org/10.3758/s13414-011-0246-z>
- Rahnev, D. A., Maniscalco, B., Lubner, B., Lau, H., & Lisanby, S. H. (2012b). Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. *Journal of Neurophysiology* 107(6), 1556-1563. DOI: <http://doi.org/10.1152/jn.00985.2011>
- Ramaswamy, V., Furstenberg, A., Breska, A., Deouell, L., & Sompolinsky, H. (2013). Single-trial decoding of intention from EEG. *Journal of Molecular Neuroscience* 51(1), 1-135. DOI: <http://doi.org/10.1007/s12031-012-9923-1>
- Reavis, E. A., Frank, S., Tse, P. U. (2015). Caudate nucleus reactivity predicts perceptual learning rate for visual feature conjunctions. *NeuroImage* 110, 171-181. DOI: <http://doi.org/10.1016/j.neuroimage.2015.01.051>
- Reavis, E. A., Frank, S., Greenlee, M. W., & Tse, P. U. (2016). Neural correlates of context-dependent feature conjunction learning in visual search tasks. *Human Brain Mapping* 37(6), 2319-2330. DOI: <http://doi.org/10.1002/hbm.23176>
- Rigoni, D., Kuhn, S., Sartori, G., & Brass, M. (2011). Inducing disbelief in free will alters brain correlates of pre-conscious motor preparation: The brain minds whether we believe in free will or not. *Psychological Science*, 22, 613-618.
- Roskies, A. (2012a). How does the neuroscience of decision making bear on our understanding of moral responsibility and free will?. *Current Opinion in Neurobiology* 22(6), 1022-1026. DOI: <http://doi.org/10.1016/j.conb.2012.05.009>
- Roskies, A.L. (2012b). Don't Panic: Self-Authorship without Obscure Metaphysics. *Nous-Supplement: Philosophical Perspectives* 26(1), 323-342. DOI: <http://doi.org/10.1111/phpe.12016>
- Roskies, A. (2013). The neuroscience of volition. In Clark, A., Kiverstein, J., & Vierkant, T. (eds.) *Decomposing the Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199746996>
- Roskies, A. (2014a). Can neuroscience resolve issues about free will. In Sinnott-Armstrong, W. (ed.) *Moral Psychology: Volume 4, Free Will and Moral Responsibility*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262525473>
- Roskies, A. (2014b). Response to commentators. In Sinnott-Armstrong, W. (ed.) *Moral Psychology: Volume 4, Free Will and Moral Responsibility*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262525473>
- Roskies, A.L. (2015). Monkey decision-making as a model for human decision-making. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Roskies, A. (2016a). How can neuroscience contribute to philosophical methodology?. In Cappelen, H., Gendler, T., Hawthorne, J. P. (eds.) *The Oxford Handbook of Philosophical Methodology*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=978019966877>
- Roskies (2016b). Decision-Making and Self-Governing Systems. *Neuroethics* 11(3), 245-257. DOI: <http://doi.org/10.1007/s12152-016-9280-9>

- Roskies, A., & Malle, B. (2013). A Strawsonian look at desert. *Philosophical Explorations* 16(2), 133-152. DOI: <http://doi.org/10.1080/13869795.2013.787439>
- Roskies, A., & Nahmias, E. (2017). "Local determination", even if we could find it, does not challenge free will: Commentary on Marcelo Fischborn. *Philosophical Psychology* 30(1-2), 185-197. DOI: <http://doi.org/10.1080/09515089.2016.1248286>
- Russell, P. (1988). Causation, compulsion, and compatibilism. *American Philosophical Quarterly*, 25(4), 313-321.
- Salvaris, M., & Haggard, P. (2014). Decoding intention at sensorimotor timescales. *PLoS ONE* 9(2), e85100. DOI: <http://doi.org/10.1371/journal.pone.0085100>
- Sartorio (2013a). Making a difference in a deterministic world. *Philosophical Review* 122(2), 189-214. DOI: <http://doi.org/10.1215/00318108-1963707>
- Sartorio, C. (2013b). Why Two Source-Incompatibilist Arguments are Question-Begging. *Flickers of Freedom*. URL: http://philosophycommons.typepad.com/flickers_of_freedom/featured-author-carolina-sartorio/
- Sartorio, C. (2015a). The Problem of Determinism and Free Will is Not the Problem of Determinism and Free Will. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>
- Sartorio, C. (2015b). Actual Sequence Freedom. In Bacchini, F., Caputo, S., and Dell'Utri M. (eds.) *New Advances in Causation, Agency and Moral Responsibility*. Newcastle upon Tyne: Cambridge Scholars Publishing. ISBN: <http://www.isbnsearch.org/search?s=9781443873567>
- Sartorio, C. (2015c). Sensitivity to Reasons and Actual Sequences. In Shoemaker, D. (ed.) *Oxford Studies in Agency and Responsibility*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780198744832>
- Sartorio, C. (2016). *Causation and Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780198746799>
- Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A. Tse, P. & Wheatley, T. (2013). Barking up the wrong free: Readiness potentials reflect processes independent of conscious will. *Experimental Brain Research* 229(3), 329-335. DOI: <http://doi.org/10.1007/s00221-013-3479-3>
- Schlegel, A., Alexander, P., Sinnott-Armstrong, W. Roskies, A. Tse, P. & Wheatley T. (2015). Hypnotizing Libet: Readiness potentials with non-conscious volition. *Consciousness and Cognition* 33, 196-203. DOI: <http://doi.org/10.1016/j.concog.2015.01.002>
- Schlegel, A., Alexander P., Sinnott-Armstrong, W., Roskies, A., Wheatley, T., Tse, P. U. (forthcoming). The suspension of will: readiness potential precede hypnotically induced movement.
- Schooler, J. W. (2013). Does belief in free will make us better people?. Templeton Big Questions Online . URL: <https://www.bigquestionsonline.com/content/does-belief-free-will-make-us-better-people>.
- Schooler, J., Nadelhoffer, T., Nahmias, E., & Vohs, K. D. (2015). Measuring and manipulating beliefs about free will and related concepts: The good, the bad, and the ugly. In Mele, A. (ed.) *Surrounding Free Will*. Oxford: Oxford University Press. ISBN: <http://www.isbnsearch.org/search?s=9780199333950>

-
- Shadlen, M. N., & Roskies, A. L. (2012). The neurobiology of decision-making and responsibility: reconciling mechanism and mindedness. *Frontiers in neuroscience*, *6*, 56.
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. L., Schooler, J. W., Baumeister, R. F., & Vohs, K. D. (2014). Free Will and Punishment: A Mechanistic View of Human Nature Reduces Retribution. *Psychological Science* *25*(8), 1563-1570. DOI: <http://doi.org/10.1177/0956797614534693>
- Shariff, A. F. & Vohs, K. D. (2014). What Happens to a Society That Does Not Believe in Free Will?. *Scientific American*. URL: <http://www.scientificamerican.com/article/what-happens-to-a-society-that-does-not-believe-in-free-will/>.
- Snoek, A., Levy, N., & Kennett, J. (2016). Strong-willed but not successful: The importance of strategies in recovery from addiction. *Addictive Behaviors Reports* *4*, 102-107. DOI: <http://doi.org/10.1016/j.abrep.2016.09.002>
- Solovey, G., Graney, G. G., & Lau, H. (2014). A decisional account of subjective inflation of visual perception at the periphery. *Attention, Perception, and Psychophysics* *77*(1), 258-271. DOI: <http://doi.org/10.3758/s13414-014-0769-1>
- 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.
- Sripada, C. S. (2012). What makes a manipulated agent unfree?. *Philosophy and Phenomenological Research* *85*(3), 563-593.
- Stillman, T.F., Baumeister, R. F., & Mele, A. R. (2011). Free will in everyday life: Autobiographical accounts of free and unfree actions. *Philosophical Psychology* *24*(3), 381-394. DOI: <http://doi.org/10.1080/09515089.2011.556607>
- Stillman, T. F., Baumeister, R. F., Vohs, K. D., Lambert, N. M., Fincham, F. D., & Brewer, L. E. (2010). Personal philosophy and personnel achievement: Belief in free will predicts better job performance. *Social Psychological and Personality Science*, *1*(1), 43-50.
- Strawson, P.F. (1962). Freedom and Resentment. *Proceedings of the British Academy* *48*, 187-211.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality*, *72*(2), 271-324.
- Tse, P. U. (2013a). *The Neural Basis of Free Will: Criterial Causation*. Cambridge, MA: MIT Press. ISBN: <http://www.isbnsearch.org/search?s=9780262019101>
- Tse, P. U. (2013b). Free Will Unleashed in *New Scientist*, 8 June 2013, pp 28-29
- Van Inwagen, P. (1983). *An Essay on Free Will*. Oxford; Oxford University Press.
- Vohs, K.D., Baumeister, R. F., & Schmeichel, B.J. (2012). Motivation, personal beliefs, and limited resources all contribute to self-control. *Journal of Experimental Social Psychology* *48*(4), 943-947. DOI: <http://doi.org/10.1016/j.jesp.2012.03.002>
- Vohs, K., Baumeister, R. F., Vonasch, A., Pocheptsova, A., & Dhar, R. (2017). Ego depletion induces mental passivity: Behavioral effects beyond impulse control. *Motivation Science* *3*(4), 321-336. DOI: <http://doi.org/10.1037/mot0000058>
- Vohs, K. D., & Schooler, J. W. (2008). The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating. *Psychological science*, *19*(1), 49-54.

-
- Vonasch, A., & Baumeister, R. F. (2013). Implications of free will beliefs for basic theory and societal benefit: Critique and implications for social psychology. *British Journal of Social Psychology* 52(2), 219-227. DOI: <http://doi.org/10.1111/j.2044-8309.2012.02102.x>
- Vonasch, A., Baumeister, R., & Mele, A. 2018. Ordinary People Think Free Will is a Lack of Constraint, Not the Presence of a Soul. *Consciousness and Cognition* 60, 133-151.
- Wallace, R. J. (1994). *Responsibility and the moral sentiments*. Cambridge: Harvard University Press.
- Waller, B. (1988). Free Will Gone Out of Control. *Behaviorism*, 16(2), 149-157. Retrieved from <http://www.jstor.org/stable/41236066>
- Waller, R. R. (2012). Beyond Button Presses: The Neuroscience of Free and Morally Appraisable Actions. *The Monist* 95(3): 441-462.
- Waller, R.R. (forthcoming). Science of Free Will: Neuroscience. In Campbell, J. (ed.) *Blackwell Guide to Free Will*. Oxford: Blackwell.
- Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American psychologist*, 54(7), 480.
- Wente, T., Bridgers, S., Zhao, X., Seiver, E., Zhu, L., & Gopnik, A. (2016). How Universal Are Free Will Beliefs? Cultural Differences in Chinese and U.S. 4- and 6-Year-Olds. *Child Development* 87(3), 666-676. DOI: <http://doi.org/10.1111/cdev.12528>
- Wiese, C. W., Tay, L., Duckworth, A. L., D'Mello, S., Kuykendall, L., Hofmann, W., Baumeister, R. F., & Vohs, K. D. (2018). Too much of a good thing? Exploring the inverted-U relationship between self-control and happiness. *Journal of Personality* 86(3), 380-396. DOI: <http://doi.org/10.1111/jopy.12322>