

How is Willpower Possible? The Puzzle of Synchronic Self-Control and the Divided Mind

CHANDRA SEKHAR SRIPADA

University of Michigan, Ann Arbor

The exercise of willpower is puzzling because it seems to require that a person both most wants to act on a wayward desire, and most wants to resist this desire, and this seems impossible. There are two accounts that try to resolve this puzzle of synchronic self-control, Jeanette Kennett and Michael Smith's 'non-actional' account and Alfred Mele's 'ancillary action' account. I criticize these accounts because they set too strong constraints on what kinds of synchronic self-control are possible, and thus what willpower could turn out to be. I then propose a 'divided mind' account that helps make sense of particularly strong forms of willpower that cannot be accommodated on the alternative accounts. On the divided mind account, motivational architecture is divided between a deliberative motivational system and an emotional motivational system, and willpower is a *proprietary action* exclusively available to the deliberative system. I address potential objections to the divided mind account. One objection says that it is not in fact possible for a weaker desire to defeat a stronger one. A second objection says that actions that arise exclusively from parts of a mind cannot be said to belong to the whole agent.

Introduction

When temptation strikes, we sometimes exercise willpower in order to suppress the wayward desire, and if willpower is successful, we are able to pursue the course of action that we had originally set out to do. Nothing in this description seems in any way out of the ordinary. Yet on closer scrutiny, the exercise of willpower is fundamentally puzzling. The problem arises when we

reflect on an influential principle enshrined in both folk and philosophical psychology:

If a person most desires to perform some action *A*, and if she believes herself free to *A*, then she will *A*, if she does anything at all intentionally.¹

This principle, which has been dubbed the ‘Law of Desire’ (1994), generates a dilemma². Suppose Sally is on a diet, but one evening she is out having dinner with friends when the dessert cart is brought to the table. Now she finds her strongest desire³ is to have a slice of the key lime pie. If eating the pie is indeed her strongest desire, then how can Sally exercise willpower to not eat the pie? Given the Law of Desire, and given that exercising willpower is not what Sally most desires, then Sally will eat the pie and will not try to exercise willpower to not eat the pie. On the other horn of the dilemma, suppose that eating the pie is not Sally’s strongest desire. That is, Sally has a strong desire to eat the pie, but the desire to stay on her diet remains her strongest desire. In this case, there doesn’t seem to be any need for exercising willpower. Given the Law of Desire, and given that staying on the diet remains Sally’s strongest desire, then Sally will stay on her diet. There is no special need to invoke the exercise of willpower to explain why Sally stays on her diet and does not eat the pie.

The puzzle raised in this example pertains specifically to exercises of *synchronic* self-control, i.e., self-control directed at attenuating the strength of a wayward desire that is active at the very time that self-control is deployed. This contrasts with *diachronic* self-control in which one attenuates the strength of some desire that is not currently active, but which one anticipates will become active in the future. It is synchronic self-control in particular that seems puzzling because this form of control seems to require that the person most wants to pursue the wayward course of action *and* most wants to resist pursuing this course of action, and this does not seem to be possible.

Two accounts have been offered that deal with the puzzle of synchronic self-control. In my view, these accounts are mistaken because they treat our ability for synchronic self-control as something less than it really is. One account, derived from the work of Jeanette Kennett and Michael Smith (Kennett and Smith 1996, Kennett and Smith 1997), treats exercises of self-control not as *actions*, but rather as passive dispositions to have certain kinds of desire-altering thoughts. The other solution, offered by Alfred Mele (Mele 1992, Mele 2003, Mele 1997, Mele 1998), treats the exercise of self-control as a kind of *ancillary* action that cannot be undertaken by itself, but must necessarily accompany a main action that serves the wayward desire. A serious problem for these accounts is that there appear to be still stronger forms of synchronic self-control, which I dub ‘*full-blooded*’ exertions of willpower, that are both *actional* and *not ancillary* to some other action

in support of the stronger wayward desire. In this paper, I propose a divided mind account that explains how this strong form of synchronic self-control is possible.

Part I: Three Accounts of How Willpower Is Possible

Before discussing how willpower is possible, I begin by providing a basic sketch of what willpower *is*. Willpower refers to a set of mental processes that serve to regulate (i.e., attenuate, suppress, block, or otherwise modify the motivational properties of) one's own desires and other motivation-encompassing attitudes.⁴ Willpower is a form of synchronic self-control, and as such exercises of willpower regulate motivation-encompassing attitudes that are active at the very time that willpower is exercised.

The notion of willpower derives from folk theory, but it has a close correspondence with what psychologists refer to as *effortful control*, which itself falls under a broader category called *emotion regulation* (Gross 1998). It is useful to make a preliminary distinction between two ways of exercising willpower, *antecedent-focused strategies* and *response-focused strategies* (Gross 1998b). Antecedent-focused strategies alter the appraisal or interpretation of the environmental cues that elicit or sustain the wayward desire (or other motivation-encompassing attitude). An example is *reappraisal*. A man tempted to head to the casino to wager his paycheck on blackjack might initially see the casino as a source of excitement. But he might then engage in reappraisal, and instead see the casino as a smoky and dingy dungeon that traps all who are foolish enough to enter. In this way, he can reduce the strength of his desire to gamble and/or strengthen his desire to not gamble.

The second broad category of willpower strategies consists of response-focused strategies (Gross and Levenson 1993). These strategies involve either the direct inhibition of motivation or the prevention of action. The direct inhibition of motivation might be accomplished by strategies such as *distancing*.⁵ Here, one mentally steps back from one's own motivation-encompassing attitudes and experiences them from a detached, third-person perspective, thereby mitigating and perhaps even eliminating their motivational force.⁶ Action prevention involves leaving the motivation to perform some action *A* intact, but mentally *blocking* the normal translation of one's motivation to *A* into one's *A*-ing.⁷ For example, a person with a powerful itch may mentally block his motivation to scratch the itch from being translated into action. Similarly a person confronted with a threat may mentally inhibit his fleeing.⁸ One can engage in response-focused strategies even without engaging in antecedent-focused strategies, though in practice, one typically deploys both kinds of strategies in combination. For example, psychologists often use the *cold pressor* task to probe willpower (Muraven and Baumeister 2000). In this task, subjects are asked to immerse one arm in a tank of extremely cold water for as long as they can. Here the primary willpower strategy

deployed involves response prevention—subjects mentally inhibit their urge to remove their arm from the water or block this urge from being translated into action. But in addition, subjects likely *also* attempt antecedent-focused strategies, such as reappraising the situation (one might imagine that the cold sensation is part of one's enjoying a frosty winter's day).

Much more could be said about how various strategies for willpower work. But with these basic distinctions in hand, let us now turn to the main question addressed in this paper: How can a person exercise strategies of willpower of the kinds just discussed that attenuate, suppress, or block one's motivation to *A*, at the very time that the desire to *A* is one's strongest desire?

1.1 Kennett and Smith's 'Non-Actional' Account

The 'non-actional' account of willpower is based on the work of Jeanette Kennett and Michael Smith (Kennett and Smith 1996, Kennett and Smith 1997). It should be made clear from the outset that Kennett and Smith present their theory as an account of *synchronic self-control* more generally rather than as an account of *willpower* specifically. However, as we shall see, the solution that Kennett and Smith propose for making sense of synchronic self-control sets up quite strong constraints on what kinds of synchronic self-control are possible, and thus what willpower could turn out to be. It is these constraints on synchronic self-control suggested by their account, and the implications for an account of willpower, that I am interested in exploring.

Kennett and Smith ask us to reflect on the case of Frog and Toad who are both munching on cookies (Kennett and Smith 1996). They have a strong intrinsic desire to be healthy and a weaker intrinsic desire to have immediate pleasure. They know that eating the cookies will give them immediate pleasure but will not promote their health. If they were instrumentally rational, they would not eat the cookies. But the sight and smell of freshly baked cookies are more vivid and salient than the prospect of future good health. So Frog and Toad find themselves with a desire to eat cookies that is currently stronger than their desire to not eat cookies. If this is so, then how can they do anything to prevent themselves from eating cookies? It seems, at first blush at least, they can't. But perhaps this is too quick.

Certainly Frog and Toad cannot, at one and the same time, both want most to eat cookies and want even more to prevent themselves from eating cookies. That is a contradiction. However it does not follow that there is nothing that they can do to control themselves at the moment of vulnerability, at least not if we adopt a more relaxed, commonsensical attitude to what we can properly be said to 'do'. When they look at the cookies, imagine that they find themselves thinking of them not as causes of pleasure, but rather as lumps of fat, and that when they think about eating them, they imagine the fat curdling in their stomachs.

Furthermore let's suppose that the effect of having these thoughts is that Frog and Toad find themselves desiring . . . not to eat any more cookies (Kennett and Smith 1996, pg. 68–69)

According to Kennett & Smith, the exercise of synchronic self-control consists, in part, in experiencing certain thoughts—thoughts that highlight certain considerations while deemphasizing others, thus overall enhancing or blunting the strengths of one's various desires. In the case of Frog and Toad, thoughts such as these could weaken their desire to eat the cookies and allow their desire to be healthy to prevail. Now the having of such thoughts only *partly* constitutes the exercise of self-control, because there might be *other* things—'feelings of shame, the thought that one is stupid, a certain inward focus of attention' (pg. 70)—that *also* have the effect of modifying the strengths of one's desires.⁹

It bears emphasis that on Kennett and Smith's account, at the time that eating cookies is their strongest desire, Frog and Toad need not *intentionally* bring about these desire-modifying thoughts (or desire-modifying attitudes and feelings). Were this required, then the puzzle of synchronic self-control would arise anew and we should wonder how Frog and Toad can be motivated to summon up these desire-modifying thoughts when their strongest desire is to eat cookies. But though some of our thoughts are brought about intentionally, many of our thoughts occur passively in the sense that they are not brought about by a prior desire or intention to have just these thoughts (or to bring about the conditions that directly lead to the having of the thoughts). Instead, it might simply be 'in our nature' to have these thoughts at certain times, or perhaps at some prior time we trained ourselves to have these thoughts as a matter of habit. According to Kennett & Smith, it is strictly the passive *non-actional* occurrence of desire-modifying thoughts (and attitudes and feelings) that constitutes the exercise of synchronic self-control.

Kennett and Smith offer an ingenious approach to solving the puzzle of synchronic self-control. Their account no doubt captures at least *some* of the ways in which we manage to stay in control at the time of temptation, as surely habits and other non-actional mental routines play an important role in maintaining continence. However, there is a further question of whether the non-actional strategies proposed by Kennett and Smith represent *all* the ways that we can possibly exert synchronic self-control. Kennett and Smith argue that this is indeed the case.

Our claim, then, generalizing on the basis of this [Frog and Toad] case, is that all exercises of synchronic self-control are non-actional. They are non-actional because there is no suitable strongest desire to set up an exercise of actional synchronic self-control (Kennett and Smith 1997, pg. 128).

If Kennett and Smith are right, then this sets strong constraints on what willpower can be. Willpower must turn out to be some species of non-actional disposition, in the same family as the passive desire-modifying thoughts that help Frog and Toad maintain control. Is a non-actional account of willpower crafted along these lines plausible?

An obvious initial objection to a non-actional account of willpower is that it deviates quite fundamentally from common sense. While there surely are certain forms of synchronic self-control that are non-actional, the forms of synchronic self-control that are picked out by the folk term “willpower” appear to be conceived in distinctively actional terms. That is, we normally think that the exercise of willpower is something that we *intentionally* bring about. A person typically *chooses* to exert willpower and such choices, because they are taken to be choices, are often held to be deeply indicative of what is important to the person. Moreover, willpower is accompanied by a distinctive actional phenomenology. To see this, imagine that Frog and Toad try to exercise willpower against the desire to eat cookies, and during this interval, we are somehow privy to the stream of thoughts and experiences in Frog’s or Toad’s mind. It is plausible that the stream might go something like this. “Come on Frog. Beat back that urge to eat the cookies. Push it out of your mind. Squash that desire with all your might.” Moreover, as he has these thoughts, Frog likely experiences feelings of *actively striving to do something*. It feels like he is *effortfully* opposing the wayward desire.

A non-actional account of willpower would be hard pressed to make sense of this inner monologue and this phenomenology. On this account, exerting willpower is not something that Frog can intentionally do. Why then would Frog command himself to do something that he cannot intentionally bring about? Moreover, it is puzzling why the passive occurrence of certain thoughts would be accompanied by the experience of actively engaging in action. Consider the fact that phenomenology associated with action reliably indicates the type of action being performed (active and intentional versus passive and non-intentional) in so many other contexts. I am easily able to distinguish, on the basis of their respective phenomenologies, the passive unfolding of thoughts during say, daydreaming, from intentionally trying to act, for example mentally calculating the sum of two large numbers. The non-actional account owes us an explanation of why in the case of exercising willpower, phenomenology misleads by presenting as active and intentional what is in fact the contrary.

The preceding arguments against the non-actional account of willpower are based on common sense intuitions and subjective reports about its phenomenology. In section 2.1, I supplement these arguments with empirical lines of evidence derived from cognitive neuroscience that also support the idea that the exercise of willpower is actional. For now, I tentatively conclude that the non-actional account of willpower, crafted along the lines of Kennett and Smith’s more general non-actional approach to synchronic self-control,

meets some serious objections, and turn to other accounts that try to make sense of how willpower is possible.

1.2 Mele's 'Ancillary Action' Account

In 'Underestimating self-control,' Alfred Mele offers a critique of Kennett & Smith's non-actional account of self-control from the perspective of Mele's own alternative account, built up over the course of extensive writings on the subject (Mele 1997); see also Mele (1992, 2003, 1998). Mele's account of self-control accommodates the common-sense idea that exercises of synchronic self-control are actions, but treats them as very special kinds of actions. In presenting Mele's account, let's begin by considering a case.

Ian has just finished eating and he is thinking that he ought (all things considered) to get back to work now. However, he is enjoying the golf tournament on TV and he remains seated. He tells himself that he will watch the match until the next commercial break; but the commercial comes and goes and Ian is still in front of the set. Thinking he had better drag himself away from the television now, Ian utters a self-command: "Get off your butt, Ian, and paint that shed!" Ian turns off the set, picks up his painter's cap, and walks to his backyard (Mele 1987, pg. 69).

How is Ian's exercise of self-control in support of his all things considered judgment possible? Mele's explanation requires paying attention to pairs of desires that are *direct competitors* of each other, versus pairs of desires that do not directly compete. Ian's desire to watch TV directly competes with his desire to paint the shed, because he cannot (or at least believes he cannot) fulfill both at the same time. And because the former desire is stronger than the latter, Ian continues to watch TV. Ian's desire to utter the self-command, on the other hand, is not a direct competitor of his desire to watch TV. That is, Ian might watch TV and *at the very same time* utter the self-command. The direct competitor of Ian's desire to utter the self-command is, rather, Ian's desire to refrain from uttering the self-command. And according to Mele, the strength of the former might very well be greater than the latter, in which case Ian will both watch TV and utter the self-command to not watch TV, and he can thereby get himself to stop watching TV.

Now one might object in the following way. Ian's desire to utter the self-command is an instrumental desire that promotes the fulfillment of his desire to paint the shed, while his desire to refrain from uttering the self-command is an instrumental desire that promotes the fulfillment of his desire to watch TV. Since Ian's desire to watch TV is, by hypothesis, stronger than his desire to paint the shed, wouldn't the desire to watch TV transmit a larger force across the means-end relation so that Ian's desire to refrain from uttering the self-command would be stronger than his desire to utter the self-command? If so, then contra the story above, Ian cannot utter the self-command.

Mele acknowledges that, were the desire to watch TV to fully transmit its force across the means-end relation, then Ian could not utter the self-command (Mele 1997, pg. 123). But if we allow that Ian occasionally performs actions that hinder, rather than advance, his own strongest desires, then Ian can utter the self-command after all. Now what specifically would cause this to occur? On Mele's account, somewhat ironically, it is the same kinds of factors that drive temptation, i.e., focus of attention, vividness of representations, and the like.

Ian's attention, we may suppose, is focused much more on the TV than on the prospect of not making an attempt at self-control. Thus, we should not find it surprising that he is more motivated to continue watching TV than he is to refrain from making an attempt at self-control. Moreover, when the thought of uttering the self-command occurs to him, he may be much more attentive to the prospect of surprising his wife with a freshly painted shed than to his reasons for refraining from making the attempt at self-control. This would give [the desire to utter the self-command] the advantage of vividness of representation over any desire that Ian may have not to utter the self-command (Mele 1987, pg.71).

Let us call the kinds of factors Mele appeals to here—focus of attention, vividness of representations, etc.—‘attractiveness factors’. These factors *narrow* the scope of a person's reasoning by causing the person to focus on certain considerations while largely ignoring others. Attractiveness-driven narrowed reasoning is presumably how Ian ended up lingering on the couch and watching TV rather than painting the shed in the first place. But, according to Mele, attractiveness-driven narrowed reasoning also provides the key to explaining how Ian can exercise willpower to stop watching TV. On his view, the desire to exert self-control can be stronger than the desire to not exert self-control so long as Ian's attention is focused on the attractiveness of the former, and he pays relatively less attention to the costs of the latter. I believe the explanation of exercises of self-control in terms of attractiveness-driven narrowed reasoning is problematic for several reasons, and I will return to this point a bit later (see sections 1.3 and 2.4). But I do not want to press this point right now.

An advantage of Mele's view is that, unlike the non-actional account, Mele's view makes sense of exercises of synchronic self-control as intentional actions that the agent actively and deliberately initiates. Mele writes,

... once we recognize that we can do more intentionally at a time than what we want most to do, the door is open to intentional, active exercises of synchronic self-control. These exercises, being intentional actions and being subject to the agent's intentional control, go well beyond the exercises that Kennett and Smith urge are possible. In stopping where they do, they underestimate our prospects for synchronic self-control (Mele 1997, pg. 123).

Yet an objection might be raised that Mele's own view stops too soon and fails to recognize still stronger ways in which we exercise synchronic self-control. Like Kennett and Smith's view, Mele's account also appears to set strong constraints on what forms of synchronic self-control are possible. The crux of the problem is that on Mele's view, exercising self-control must always be an *ancillary* action. It is a hanger-on that is done only in addition to the main action(s) that one does on behalf of one's wayward desire. To be sure, self-control conceived in this way can achieve a lot. It can get an agent like Ian to desist from doing something he is already doing (i.e., watching TV). And it can even get a person to not do something she is about to do. For example, suppose Sally's strongest desire is to eat pie. While she is walking to the refrigerator to get a large slice of pie, there may be enough time for her to utter a self-command that changes the balance of her motivation, so that her strongest desire no longer is to eat pie, and is instead to stay on her diet. Here exercising willpower is still an ancillary action because it accompanies Sally's doing something (i.e., walking to the refrigerator) to satisfy her wayward desire, even though Sally's wayward desire is never actually satisfied because she never eats any pie [see Mele (1998, 2003, pg. 190) for other examples that illustrate this idea].

But there may be still stronger forms of synchronic self-control that Mele's ancillary action view precludes, especially if we reflect on Mele's view in light of the common sense understanding of willpower. We often think of willpower as a *barrier* that we impose on ourselves over some extended interval of time that blocks us from ever doing (or starting to do) anything in the service of a wayward desire, even though during that entire interval, our wayward desire remains our strongest. Here the exercise of willpower is not an ancillary action because the person never does, and never even *starts to do*, anything that promotes her wayward desire. We might call this *full-blooded willpower* to reflect the fact that exerting willpower over time in this way often feels particularly difficult and effortful. To illustrate the idea of full-blooded willpower, here is an example.

Mike was bitten by a blond Labrador as young child and ever since, he has had a deep-seated fear of dogs. His boss Helen has invited him over for dinner, and he welcomes the chance to get to know her and her family better, and perhaps also convince her that he is the right man to head the firm's new project. Mike, Helen, and Helen's family sit on the sofa and enjoy a friendly conversation. Unexpectedly, a large and spunky blond Labrador enters the living room, walks briskly straight at Mike and begins to paw at his legs. Mike has a powerful desire to flee the room, and this is his strongest desire. At that very time, he judges that fleeing the room would create a scene and compromise his relationship with his boss. Though it is intensely difficult, Mike exercises willpower to prevent himself from fleeing, and he thereby stays on the sofa. The dog continues to paw at Mike's legs, Mike's strongest desire remains to flee, and Mike continues

to exert willpower to prevent himself from fleeing. Sensing Mike's discomfort, Helen removes the dog from the room a few minutes later.

On Mele's ancillary action view of synchronic self-control, the preceding story is impossible as stated; it is not possible that during the extended interval where the dog is pawing at his leg, Mike's strongest desire is to flee the room, he does *not* act on this desire, and *instead* he intentionally exercises willpower to prevent himself from fleeing, and in this way he stays on the sofa. In order to make sense on Mele's ancillary action view, the story would need to be modified in one of two ways. First, Mike might flee the room when the dog comes near. As he is fleeing, Mike performs an ancillary act of self-control such as issuing a self-command: "Mike, get back onto that sofa!" The self-command then attenuates his desire to flee so that his strongest desire is to return to the sofa. Alternatively, at the time he sees the dog and his desire to flee is triggered, Mike might begin the process of fleeing. But there will surely be a small gap in time between when the act of fleeing is initiated and when it is fully underway (for example, Mike has to ready his muscles to stand up, get off the couch, and ready his muscles to run away, all before he can finally run away). During this gap, Mike might issue a self-command: "Mike, stay on this sofa!" The self-command attenuates his desire to flee so that his desire to stay on the couch is stronger than his desire to flee. Thus Mike only *starts to flee*, and does not complete this action after all (Mele 1998, Mele 2003, pg. 190).

The story of Mike highlights that the following three conditions seem co-possible:

Conditions for *full-blooded* willpower:

1. During some (arbitrarily long) interval I extending over $t_1 \dots t_n$, S s all things considered judgment about what to do is opposed by his wayward desire D to A straightaway, and throughout I , D is S s strongest desire.
2. At t_2 , even while D remains S s strongest desire, S undertakes an active intentional exercise of willpower that prevents him from acting on D .
3. At no time during I does S perform any actions that promote D , or start to perform any actions that promote D (in other words, the exercise of willpower is not an ancillary action that accompanies S s acting to promote D).

Mele's view cannot make sense of an agent's satisfying these three conditions, and instead must redescribe the story of Mike in one of the two aforementioned ways. In my view, the original story of Mike seems perfectly plausible, and it is a problem for Mele's view that it requires these redescriptions. Rather than spending time to point out in detail problems with Mele's redescriptions, my strategy will be to present an alternative view of willpower that makes sense of how the preceding three conditions can be jointly satisfied.

1.3 The Divided Mind Account

This brings us to a third candidate solution to the puzzle of synchronic self-control—a divided mind approach. I will introduce this account in the context of one plausible view of how the human mind is in fact divided into two motivational compartments that correspond, roughly, to the folk notions of Reason and Passion. It bears emphasis however that the more important point being made in what follows is that dividing the mind into motivational compartments *in some way or other* is a necessary condition for the existence of full-blooded exertions of willpower. Thus even if the particulars for my own favored account of the divided motivational structure of the mind turn out to be partially or even *entirely* wrong, the more general point that motivational division is required for full-blooded willpower should nonetheless survive.

The Deliberative and Emotional Motivational Systems

Let me briefly sketch an empirically driven view that the human mind is divided into two motivational compartments, which I will call the deliberative motivational system and the emotional motivational system.

The deliberative motivational system refers to the collection of processes that implement practical reasoning. In the decision-theoretic school in philosophy and psychology, this kind of reasoning is often given a Bayesian gloss: The agent has a set of candidate actions and outcomes contingent on taking those actions. The agent next assigns numbers to the outcomes that represent their *desirability*, and numbers to the actions that represent the *probability* of bringing about the outcomes. The action selected is that with the highest expected desirability. Alternatively, one might understand practical reasoning not in terms of weighing expected desirability but rather in terms of *weighing reasons*. On this view, an agent has a set of candidate actions and weighs the reasons favoring and opposing these actions. The relevant reasons may encompass information about probabilities and desirabilities, but may also include considerations that arise from a broader set of moral or prudential concerns. The action chosen is that which is favored by the overall weight of reasons.

Practical reasoning terminates in a judgment about what, all things considered, is the thing to do, which I refer to as one's *practical judgment*. Practical judgments motivate action. They do this either directly through formation of a new action-desire (i.e., occurrent desires to act in certain ways, typically at a determinate time),¹⁰ or indirectly, via a mediating intention. I'll call the desires produced by the deliberative motivational system one's *practical desires*.

In addition the deliberative motivation system, there is a second motivation system capable of producing action-desires. The emotional system includes not only *emotions*, such as anger, fear and disgust, but also *drives*

such as hunger and thirst, and closely related states such as *cravings*. While these processes differ in many important respects, contemporary theory views them as having certain core underlying commonalities.¹¹ For ease of explication, I focus on emotions. But similar remarks apply to drives and cravings as well.

Emotions can be understood as triggered responses to certain prototypical sorts of events of concern to the person. They produce a suite of coordinated cognitive and physiological changes, and they eventuate in motivations to pursue certain characteristic sorts of behavior. The motivational upshots of an emotion involve the formation of new action-desires, *emotional* action-desires. For example, if you suddenly see a snake in the woods, fear is rapidly and automatically triggered, and the bout of fear is associated with the formation of new action-desires to flee the situation.¹²

I want to highlight certain features of the triggering process by which an emotion episode is initiated and the action-desires that an emotion produces. Emotions are sometimes characterized as exhibiting two distinctive features, *passivity* and *recalcitrance*.¹³ The passivity of emotions pertains primarily to the triggering process. The appraisal process by which an emotion is triggered typically unfolds automatically, and does not require conscious attention or supervision. Indeed, the triggering of emotions is usually *mandatory*. If confronted with the appropriate situational elicitors, the emotion episode will be triggered whether one likes it or not. Of course one might try to disrupt the triggering of emotion by impeding processes of perception (for example, one might close one's eyes at the sight of something disgusting). But the point is that the appropriate *awareness* of a situation's being the case typically mandates the triggering of an emotion episode regardless of whether or not one wants the episode to occur.¹⁴

Recalcitrance refers to the endurance of a bout of emotion even when one has made a sincere judgment that the emotion is not apt or warranted (see Justin D'Arms and Jacobson 2003). In other words, given that an episode of emotion has been triggered, one's conscious judgments cannot terminate the episode by fiat. The aspect of recalcitrance I am most interested in is the recalcitrance of the action-desires produced during the emotion episode. These action-desires typically persist and continue to exert their motivational force even in the face of a conscious practical judgment that the actions promoted by these action-desires are *not* the thing to do.

I highlight the passivity/mandatoriness of the triggering of emotion and recalcitrance of the motivational upshots of emotion to underscore that emotions are relatively *independent* of practical judgments specifically, and one's conscious evaluative judgments more generally. The overall picture being sketched, then, is one of two motivational systems, where the two systems are relatively independent of each other. In particular, the two systems can diverge from each other in terms of their motivational upshots, i.e., in terms of the action-desires that arise from the two systems.

Willpower as a Proprietary Action

We are now in a position to locate willpower within this architecture. On the account of willpower that I am proposing, the distinctive feature of willpower is that it is an action that is *exclusively* available to only one of the two motivational compartments, specifically the deliberative compartment. The exercise of willpower has the effect of attenuating, blocking or in some other way modifying the motivational properties of an action-desire produced by the emotional motivation compartment. Thus, willpower is a *proprietary action*; it is motivationally initiated and maintained exclusively by one part of the divided mind.

Mele's helpful notion of the *motivational base* for an attitude (Mele 1992, Mele 2003) helps clarify the idea that willpower is a proprietary action. I start by defining the motivational base for an attitude, and then extend the idea to an action. Call the positive motivational base of an attitude *A* the set of motivation-encompassing attitudes that contribute to the motivational strength of *A*. Similarly, call the negative motivational base of *A* the set of motivation-encompassing attitudes that weaken the motivational strength of *A*. For example, Jimmy's desire to go to veterinary school is strengthened by his desire to be financially secure and his evaluative judgment that animals ought not needlessly suffer, and is weakened by his desire to avoid going into debt, and his concern that he might be embarrassed if he is not admitted. The first two motivation-encompassing attitudes are part of the desire's positive motivational base, while the latter two motivation-encompassing attitudes are part of the desire's negative motivational base. Given the notion of a motivational base for an attitude, we can readily extend the idea to the motivational base for an action in the following way. The motivational base for *Ss* ϕ -ing at *t* consists of the motivational base for *Ss* action-desire to ϕ at *t*.¹⁵

On a standard unified mind account, the total motivational base for exercising willpower, i.e., the positive and negative motivational bases taken together, could potentially include *any* of an agent's motivation-encompassing attitudes. The key feature of the divided mind account is that it is committed to the following thesis:

- (M) The motivational base, both positive and negative, for the exercise of willpower consists exclusively of motivation-encompassing attitudes within the deliberative motivational system.

M says that the question of whether an agent exercises willpower depends on the balance of motivations that favor or oppose exercising willpower within the deliberative motivational compartment. Motivation-encompassing attitudes that oppose willpower *outside* the deliberative motivational compartment—in particular, the wayward desire that deviates from

a person's practical judgment—are *not* part of the motivational base, either positive or negative, that is relevant to the question of whether willpower is exercised.

In assessing the plausibility of **M**, it is important to get clear on what it does and does not say. It is possible to read **M** as saying something much stronger than it does, and on this basis reject it. In particular, one might read **M** as saying that motivation-encompassing attitudes in the emotional compartment (I'll refer to these generically as one's emotional desires) are *irrelevant* to the exercise of willpower. But this reading is incorrect. **M** is compatible with emotional desires playing a wide-ranging role in whether and how willpower unfolds. What **M** disallows is, quite specifically, that these emotional desires are part of the motivational base for willpower, and thus they '*weigh in on*' (qua attitudes that *contribute motivation*) the question of whether willpower is exercised. Examination of some cases helps to clarify this point.

Every single day Jill has strong desires for salty snacks in the late afternoon. She is usually able to exercise willpower to control these desires. Today, Jill missed breakfast and lunch, and her desire for salty snacks is much stronger than usual. Just as she does every other day, Jill exerts willpower against her late afternoon cravings. But this time, the wayward desire is too strong. The exercise of willpower fails and Jill ends up devouring a large bag of salty potato chips.

It is plausible that if we hold fixed the fact that a person initiates an exercise of willpower, the exercise is more likely to succeed if the wayward desire is relatively weak than if it is strong. So this is one way in which emotional desires influence willpower—they influence whether the exercise of willpower is likely to *succeed*. This kind of influence, however, is clearly compatible with **M**, which concerns the conditions for *undertaking* an exercise of willpower rather than the conditions for whether the exercise succeeds.

Consider another case.

Based on past experience, Jack knows well that if he tries to exercise willpower against his desire to eat the pie, he will struggle mightily for a while but will eventually fail. To save himself the effort and distress, Jack decides not to exercise willpower at all and eats the whole pie.¹⁶

In this case, an emotional desire (i.e., the pie-directed desire) influences Jack's decision to not exercise willpower. Is this a counterexample to **M**? It is not. The pie-directed desire (let us call it *P*) does not itself *contribute motivation* towards Jack's desisting from an exercise of willpower. The attitude that does contribute motivation is Jack's desire to avoid effort and distress (let us call this *E*), and *this* desire, by hypothesis, operates in the deliberative motivational compartment. Desire *E* is connected to *P* in the following

way: Effort and distress associated with resisting *P* is an *object* of *E*. This is the sense in which *P* is *relevant* to Jack's exercising willpower. Despite this connection, only *E* is part of the motivational base for the exercise of willpower—*P* is merely an object of an attitude that is part of that base. **M** is thus unthreatened by the kinds of influence at work in cases like Jack's.

Consider a final case.

Like Mike [see section 1.2], Mack is terribly afraid of blond Labradors. One day, Mack is walking down the street when he sees a large blond Labrador walking his way. Without even thinking, Mack simply runs away. Were Mack to have deliberated about the matter carefully, however, he would judge that exercising willpower against his desire to flee is the thing to do, and he would have tried to prevent himself from fleeing.

Mack's emotional desire to flee from the dog is relevant to the fact that he fails to exercise willpower. Is this a counter-example to **M**? Again the answer is no. Mack's emotional desire to flee does not *motivate* his not exercising willpower. Rather, it *precludes* his having the opportunity to exercise willpower. Because the emotional desire does not contribute *motivationally* to the answer to the question "Do I exercise willpower?" (indeed the emotional desire prevents the question from being posed at all), the case of Mack is compatible with **M**.

These cases help to clarify that **M** is compatible with a number of rich and varied ways in which emotional desires influence the exercise of willpower. Emotional desires can influence *whether* willpower is exercised, how it *unfolds*, and whether it *succeeds*. Though **M** allows these varied roles, the restriction it imposes—emotional desires are not part of the motivational base for the exercise of willpower—is nonetheless critical. As we shall see, it is this restriction that holds the key for explaining how willpower, and in particular *full-blooded willpower*, is possible.

Explaining the Willpower Cases

Now that we have gotten clear on what the divided mind view says, it will be helpful to return to the cases of Mike and Ian discussed earlier, and see how the divided mind view understands these cases.

Recall that Mike is terrified of the blond Labrador and his emotional desire to flee is his overall strongest. But he also judges that it would be all things considered best not to run away and create a scene. Mike actionally exercises willpower and in doing so completely prevents himself from doing *anything* in support of his strongest desire. The two accounts of synchronic self-control discussed earlier, the non-actional account and ancillary action account, could not make sense of full-blooded exertions of willpower such as this.

According to the divided mind view, full-blooded willpower is possible because the exercise of willpower is a proprietary action. Its motivational base consists exclusively of motivation-encompassing attitudes that play a functional role within the deliberative motivational system. By hypothesis, Mike's strongest practical desire is to exert willpower in support of his not fleeing. According to the divided mind view, this alone suffices for there to be sufficient motivational support for his exercise of willpower. The fact that Mike's strongest *overall* desire is to flee is not relevant to whether sufficient motivation exists for his exerting willpower, since *this* desire lies outside the deliberative motivational compartment. It is hence not part of the motivational base for willpower.

Earlier we saw the example of Ian who, as he is watching TV, utters a self-command that gets him to stop watching. Mele's ancillary action account *can* make sense of how *non*-full-blooded exercises of self-control such as Ian's are possible. The divided mind account, however, provides a quite different explanation than Mele's for how Ian manages to exercise self-control, and I believe the divided mind explanation is overall more plausible.

Recall that according to the ancillary action account, the motivation to exercise willpower is to be explained in terms of *attractiveness-driven narrowed reasoning*. Ian's self-command hinders the satisfaction of Ian's strongest desire to watch TV. Thus, to make sense of why Ian utters the self-command, the account proposes that Ian is influenced by the attractiveness of certain prospects and largely ignores others. As a result he performs an action that hinders, rather than promotes, his own strongest desire. This, of course, is quite counterintuitive because we normally think decisions to exert willpower depend on a *broadening* of one's reasoning *beyond* what seems most attractive at the moment.

The divided mind account provides an explanation for how willpower unfolds in cases like Ian's that fits better with common sense. On the divided mind view, the deliberative motivational compartment provides the motivational support for Ian's exerting willpower, and the exercise of willpower has the effect of suppressing or blocking desires arising from the emotional motivational compartment. If this picture is correct, then we can respect the common sense idea that exerting willpower does not depend on attractiveness-driven narrowed reasoning, but rather involves a broadening of one's reasoning. The deliberative motivational system specializes in integrating a broad range of considerations in arriving at an all things considered judgment about the thing to do. The emotional system, in contrast, often produces fast and automatic responses to very specific features of the situation (e.g., the presence of danger in the case of fear or the availability of a highly attractive food item in the case of a food craving). Thus if the divided mind account of willpower is correct, exercises of willpower will typically involve motivation arising from one's responsiveness to a broader set of considerations (mediated by the deliberative compartment) overriding

motivation arising from a narrower set of considerations (mediated by the emotional compartment), thus vindicating the common sense view.

1.4 Problems and Objections

Two potential problems about the divided mind account of willpower need to be addressed. The first concerns whether, contra the Law of Desire, sense can in fact be made of the idea that a weaker desire defeats a stronger one. The second is about the ownership of action, and whether a *person* can be said to perform actions that are initiated and executed exclusively by *parts* of the person's psychology.

A Challenge from the Law of Desire

The puzzle of synchronic self-control is generated by the Law of Desire, which is restated below:

If a person most desires to perform some action *A*, and if she believes herself free to *A*, then she will *A*, if she does anything at all intentionally.

Consider the following objection to the divided mind view, which is based on this principle. On the divided mind view, when temptation strikes, motivations arising from the two motivational systems compete for control of action. In particular, one's practical desire arising from the deliberative motivational system competes with an opposing wayward desire arising from the emotional motivational system. Given the Law of Desire, however, one's final action will be in accordance with whichever of these two desires is stronger. The puzzle of synchronic self-control thus arises once again, but now at a higher level, i.e., at the level of desires arising from entire motivational systems. So the divided mind account fails to make sense of synchronic self-control after all.¹⁷

Addressing this objection requires drawing a distinction between two different ways that desires might compete. The standard model of desire competition, which is tacitly assumed for the purposes of most philosophical discussions, involves a *motivational contest*. In this sort of contest, very roughly, the desire with greater motivational strength wins out (other things being equal). On the divided mind view, however, there is a second way that desires can compete, which we can call a *regulation-mediated contest*. In this kind of contest, one's practical desire provides motivation to engage the regulatory systems that execute willpower. Once triggered, regulatory systems operate to suppress the wayward desire. The key difference between the two kinds of contests comes down to which causal powers are relevant to determining the winner of the contest. In a motivational contest, it is the causal powers of the respective desires *alone* that determine the winner. In a regulation-mediated

contest, there is an additional actor: The causal powers of regulatory systems to suppress the wayward desire play an additional independent role in deciding the winner.¹⁸

The preceding characterizations of a motivational contest and regulation-mediated contest are intended as a rough first pass. If we are to construct a more detailed and precise formulation of these two models of desire competition, we must first get clear on a cluster of interrelated notions, including desires, their causal powers, and their motivational strengths.

On a standard account, desires are *dispositions to act* (Smith 1987). A person's desire for *A* disposes the person to perform actions that (she believes) bring about *A*. Given this account of desires, the notion of the *motivational strength* of a desire arises as a natural extension. Dispositions often come in degrees. For example, one vase may be more fragile than another, or one type of wood might be more flammable than another. Similarly, the dispositions to act associated with desires come in degrees—one desire may incline action more than another. When this is so, the first desire is said to be motivationally stronger than the second.¹⁹

How in particular does a desire for *A* manage to pull off the feat of disposing a person to act in *A*-promoting ways? This is a question about the *causal powers* of desires. Standard accounts in philosophy and psychology credit desires with a number of specific causal powers that help to explain their characteristic effects on action.²⁰ For example, desires draw one's attention towards desire-relevant situations and prospects (Scanlon 1998). They also enhance certain hedonic attributes of prospects, such as their expected pleasure, pleasantness, or appeal (Strawson 2009). That is, they cast their objects in a favorable light. In addition to the preceding causal powers through which desires *indirectly* affect action, desires also exert relatively *direct* effects on action selection/production systems. For example, on one plausible and empirically well-supported view, desires are associated with number-like representations that induce an ordering over the relevant set of active desires, and action systems are configured to favor as inputs desires associated with higher numbers (Shizgal 1997).

Let us use the term *D-powers* to refer to the set of causal powers possessed by desires that explain how desires have their characteristic effects on action.²¹ The notion of D-powers helps clarify the idea of motivational strength. Earlier it was said that desire *A* is motivationally stronger than desire *B* if *A* inclines action more than *B*. The notion of D-powers helps fill out the underlying basis for attributions of motivational strength. If it is true that *A* is motivationally stronger than *B*, then this is in virtue of the fact that some or other specific D-powers associated with *A* are more potent than those associated with *B* (e.g., relative to *B*, *A* grabs attention, enhances the appeal of prospects, and biases action systems to a greater degree).

We are now in a position to state more precisely the meaning of a motivational contest.

Motivational Contest: Desire *A* competes with desire *B* in a motivational contest if *A* competes with *B* for control over action, and whether *A* or *B* wins out is (non-deviantly) explained exclusively by manifestations of their respective D-powers.

An important feature of this account is that it allows that there might be at least some ways in which desire *A* beats desire *B* that do *not* count as *A*'s beating *B* in a motivational contest, because manifestations of *other* powers, not just *A* and *B*'s D-powers, play a critical role in explaining *A*'s beating *B*. This is exactly what happens in cases involving exercises of willpower, or so I will argue.

Let us now turn to the task of characterizing a regulation-mediated contest. In this sort of contest, in addition to the causal powers of the desires involved in the contest, there is another relatively independent set of causal powers that we must account for—the causal powers of the regulatory systems that modify the strengths of desires.

On the divided mind account, the exercise of willpower is a proprietary action available to the deliberative motivational system that functions to prevent an emotional desire from winning out in controlling action. In the same way that desires are associated with a set of causal powers that explain their characteristic effects in *promoting actions*, the regulatory systems that implement willpower are associated with a set of causal powers that explain their characteristic effects in *regulating desires*. These powers are manifest in the execution of a set of strategies, discussed in section 1, that are broadly divided into antecedent-focused strategies and response-focused strategies. To recap a bit, the former strategies involve such things as the redirection of attention and distancing oneself from emotionally-salient situations. The latter strategies involve the attenuation of the motivational force of emotional desires and blocking their translation into action.

Let us use the term *R-powers* to refer to the set of causal powers possessed by the regulatory systems implementing willpower, and that explain how these regulatory systems have their characteristic effects in regulating desires.²² Once we place R-powers into the picture, then it becomes clearer how a motivationally weaker desire could beat a motivationally stronger one. For this to happen, a sequence of the following kind must occur.

1. *A* is the person's motivationally strongest practical desire, and *A* is opposed by *B*, an emotional desire that is the person's overall strongest.
2. *A* provides the motivation for exercising willpower against *B*, and an exercise of willpower ensues.

3. The R-powers of regulatory systems that implement willpower are sufficient to defeat *B*.
4. *B* is defeated.

Let us review these steps to ensure their cogency. Step 1 presents the standard account of what happens during temptation and is largely uncontroversial. Step 2 is possible because the divided mind view treats willpower as a proprietary action, and its motivational base is exclusively the motivation-encompassing attitudes in the deliberative compartment. Even if one's practical desire *A* is motivationally weaker than one's wayward desire *B*, *A* can still initiate an exercise of willpower against *B*. Step 3 is granted by hypothesis. Step 4 follows naturally from steps 1–3.

At this point, a defender of the Law of Desire might object that I am drawing an artificial distinction between the D-powers of one's practical desire that motivates an exercise of willpower and the R-powers of regulatory systems that reign in the wayward desire. A desire's D-powers, the objection continues, encompass *all* of the ways that a desire brings about the state of affairs that is the object of the desire. Suppose a practical desire motivates an exercise of willpower to defeat a wayward desire. Then, says the objector, we should take the R-powers of regulatory systems manifested during the exercise of willpower and incorporate them into the set of D-powers possessed by this practical desire. When this *augmented* set of D-powers is considered, the practical desire is not in fact motivationally weaker. So the sequence described in 1–4 does not after all present a problem for the Law of Desire.

The problem with this objection is that it confuses the causal powers of desires to motivate an action with the causal powers of the processes that unfold as a consequence of this action. An example helps to bring home this point. Suppose a man operates a forklift and he desires that a heavy load that sits on the forks be lifted high off the ground. The man forms an action-desire to press the button that activates the lift mechanism. As the man continues to depress the button, the mechanism operates and the load is lifted higher and higher off the ground just as the man desires.

In this example, it would be clearly incorrect to say that the desire to raise the heavy load not only has the power to motivate certain actions (e.g., the pressing of the lift button), it also has the power to lift the heavy load. The power to lift the load, obviously enough, belongs to the forklift; it does not belong to the man's desire that the load be lifted. To be sure, there is a causal chain that links the two. The forklift's power to lift the load was manifested due to the man's action (i.e., pressing the button), and this action in turn was caused by his desire that the load be lifted. Despite the presence of this causal chain that traces back to the man's desire, the following principle seems correct:

- (C) The causal powers of processes that operate as a consequence of one's action cannot be credited 'backwards' to the desire that initiated the action.

Let us apply **C** to the case of the exercise of willpower. On the divided mind view, the exercise of willpower is actional. Of course, exercising willpower is a special kind of action; it is a *mental* action that is *proprietary* initiated and sustained by the deliberative motivational compartment. But these additional features do not in any way change or diminish the status of exercising willpower as an action on par with any other. Once engaged, the regulatory systems that implement willpower manifest various R-powers that suppress, block, attenuate, etc. the wayward desire. The defender of the Law of Desire referred to earlier wants us to assimilate these R-powers of regulatory systems into the D-powers of the practical desire that supports the exercise of willpower. But **C** forbids us from doing just this. Just as we cannot credit the powers of the forklift to the man's desire that activated and sustained the operation of the lift, we cannot credit the powers of regulatory systems to the practical desire that activated and sustained the operation of these systems. Instead, regulatory systems should be regarded as a separate and independent locus of causal powers for suppressing the wayward desire.

The importance of respecting **C** can be further illustrated by considering another example. Suppose that on Monday Sally's strongest practical desire is to stay on her diet, but her strongest overall desire is to have a large slice of key lime pie. Sally actionally exercises willpower and she successfully blocks her desire for pie from ever manifesting in action. On Tuesday the exact same situation ensues, with the strengths of Sally's competing desires exactly the same as Monday. The difference, however, is that someone has slipped Sally a psychoactive pill that specifically targets the neural circuits that comprise her regulatory systems. As a result, the relevant R-powers of her regulatory systems to suppress the wayward desire are diminished. Sally actionally engages these regulatory systems exactly as she did on Monday but this time willpower fails and her desire for pie wins out.

If we do not respect **C**, and instead assimilate the R-powers of Sally's regulatory systems into the D-powers of her practical desires, then we are forced to say that Sally's desire to stay on her diet is weaker (motivationally) on Tuesday than Monday. But this seems plain wrong. Her desire to stay on her diet is *stipulated* to be identical on both days. In particular, this desire inclines Sally—to the exact same degree on both days—to engage in an exercise of willpower. It is the regulatory systems *themselves* that are weaker on Tuesday, and not the practical desire that initiates and sustains the operation of these systems. But if this is so, then R-powers are importantly distinct from D-powers and the former cannot simply be collapsed into the latter.²³

If R-powers are distinct from D-powers, as the preceding arguments suggest must be the case, then we need to make room for another type of contest between desires. Unlike a motivational contest, which is decided exclusively by the respective desires' D-powers, in this type of contest,

regulatory systems' R-powers make an *independent* contribution to deciding the winner.

Regulation-mediated Contest: Desire *A* competes with desire *B* in a regulation-mediated contest if *A* competes with *B* for control over action, motivation arising from *A* fully engages regulatory systems that implement willpower (directed at suppressing *B*), and, given that these regulatory systems are fully engaged, whether *A* wins out over *B* is (non-deviantly) explained by manifestations of the regulatory systems' R-powers.

Notice that in a regulation-mediated contest, it is entirely possible for a motivationally weaker desire to defeat a stronger one. If the weaker desire engages regulatory systems, then so long as the regulatory powers of these systems are sufficient, then the stronger desire can be defeated.

In philosophical psychology, it is often simply assumed that motivational contests exhaust all the ways in which one desire beats another to gain control of action. The Law of Desire *does* apply to this type of contest, and so it is assumed that the Law of Desire is a perfectly general law that applies to all contests between desires. But once we recognize the possibility that desires can compete via other sorts of contests, then the unrestricted application of the Law of Desire is challenged. According to the divided mind account, the contest during temptation between one's practical desire and one's wayward desire is not a motivational contest. Rather it is a regulation-mediated contest, and this type of contest lies outside the jurisdiction of the Law of Desire. Overall then, the Law of Desire is not shown to be *false*, but rather its domain is restricted. In the type of contest between desires that interests us, i.e., contests involving the exercise of willpower, the Law of Desire fails to apply.

Ownership of Action

On the divided mind account, willpower is a proprietary action that is initiated and sustained exclusively by one part of the divided mind. Actions, however, are normally thought to be produced by *persons*, not *parts of a person's psychology*. So one might raise a concern that on the present account, we cannot attribute willpower to the whole person, since only a strict subset of the person's psychology ever participated in its execution (and indeed the remainder of her psychology opposed the action). Inability to attribute willpower to the person, the objection goes, is a counterintuitive consequence that ought to be avoided.

The problem with this objection is that it sets unrealistically high standards for when it is that an action can be attributed to a person. When a person acts, it is *invariably* the case that only a subset of her motives and evaluative commitments are called to mind prior to the action and are brought to bear

on the choice. To exhaustively bring the entirety of one's evaluative repertoire (one's entire set of desires, cares, hopes, commitments, etc.) to bear during the selection of every single one of our actions is impossible for finite creatures like us. Some of what is not brought to mind, to be sure, is not *relevant* to the selection of the action. But, no doubt, some of the material we fail to bring to mind *is* relevant, and perhaps would have counted strongly against performing the action. In this sense, most every action has its source only in parts of an agent's psychology. Unless we are willing to say that most all actions are in fact *not* attributable to the person, we need to identify some other, more lenient, criterion for attributability.

In trying to formulate the correct criterion of attributability, it is useful to consider a family of cases that has attracted much attention in the literature on moral responsibility. This family of cases pertains not to actions at all, but rather to relatively simple mental doings (or failings to do) that clearly have their source in only a part of the mind. Suppose a selfish and callous man forgets his wife's birthday.²⁴ Most of us would say the forgetting is attributable to the man. We do not attribute the forgetting to a psychological component within the man's memory system, but rather to the person. This is despite the fact that the forgetting is caused by a failure of a discrete psychological part. The information about his wife's birth date is of course recorded in the man's mind in some or other subsystem (he is not *ignorant* of his wife's birth date). But this subsystem did not provide the required information to the other parts of the man's mind that needed this information. Yet despite the fact that a failure of only a subsystem is implicated in the forgetting, it is the *person* to whom the forgetting is attributed and who is thus morally responsible. Why?

According to an influential account, the forgetting is attributable to the man because it *reflects* the man's underlying cares, concerns, and evaluative commitments (Smith 2005, Smith 2008, Scanlon 1998). The man is selfish and callous, and the forgetting reflects these particular facts about who he truly is. Now we needn't suppose that the forgetting reflects *every single element* of the man's evaluative repertoire. People sometimes have cares and commitments that are in tension in intricate and subtle ways, so it is plausible that the forgetting reflects certain aspects of the man's evaluative self without reflecting, or perhaps even being in tension with, other aspects. Thus all that seems to be needed for attributability is that the forgetting is anchored in *some or other* aspect or element of the man's overall evaluative repertoire. When an action or attitude reflects a person's evaluative self in this way, it is correct to say the action or attitude belongs to the *person*.

Applying this account of attributability to the case of willpower, it is immediately clear that exertions of willpower, though produced by just one part of the divided mind, are almost always attributable to the person. Exercises of willpower are initiated and sustained by the deliberative motivational system, which as noted earlier, implements practical reasoning processes. It is

almost always the case that one's cool and reflective practical deliberations yield decisions about what is the thing to do that reflect at least some or other element(s) of one's underlying evaluative repertoire. Now one might make an even stronger claim that judgments made from the practical point of view have an *essential* tie to one's evaluative self (Watson 1975, c.f. Watson 1987, pg. 150). But even without this stronger claim, it is certainly plausible that in most all cases when we exert willpower, we are responding to our evaluative selves. Given the criterion for attributability under discussion, this is enough to say that exertions of willpower, though they are initiated and sustained by only one part of a divided mind, can nonetheless be properly attributed to the whole person.

1.5 Holton's 'Resolution' Account

Before moving on, it is worth spending a bit of time examining one additional account of willpower, Richard Holton's resolution account (Holton 2009). I want to highlight underlying connections between Holton's account and my own divided mind account.

Holton's account is motivated by an appreciation for the importance of *intentions* in our mental lives. Intentions are a particular species of action-guiding mental state that play a characteristic role in organizing our deliberations. There is often a need to deliberate about actions well before they are to be performed, in part because conditions favorable for deliberation exist now and will not exist later, and in part because the actions to be undertaken in the future require a suite of preparatory auxiliary actions to be undertaken. For example, I might deliberate in the summer about where to go for winter vacation knowing that the fall will be a busy time with teaching, and that I had better have obtained my tickets and packed accordingly well before I travel. Here is how the ability to form intentions helps. I deliberate in the summer and form the intention to go to Aspen for winter vacation. The intention then organizes my performing the requisite auxiliary preparatory actions in the interim (such as buying ski gear), and implementing the action when the appropriate time comes. All this occurs without needing to repeatedly re-engage in costly deliberations about where to go for winter vacation prior to performing each of these actions.

In order for intentions to play this role in organizing our deliberations, they need to have certain features. Holton follows the influential work of Michael Bratman (Bratman 1987) in arguing that intentions need to exhibit, *inter alia*, *stability*—they should not be too readily revised, and *control*—they need to be directly action-guiding without one's needing to engage in further deliberation. Holton next identifies one species of intention that is of particular interest, a *resolution*. These are intentions that are meant to withstand the force of future contrary inclinations, perhaps by being

associated with higher-order intentions to deter the contrary inclination, or perhaps by some other means. Holton then provides an account of willpower as a distinct mental faculty whose purpose is to support one's resolutions.²⁵ Willpower helps maintain resolutions over time and prevents them from being too readily revised due to the pressure of a contrary inclination.

Holton's picture is overall an attractive one, and there is much in it with which I agree. But I want to draw attention to the fact that Holton's account of willpower is like my own divided mind account of willpower in terms of the key strategy employed by both accounts to explain the motivation to exert willpower. This strategy consists in demarcating a motivational base for willpower that is a subset of one's overall set of motivation-encompassing attitudes. The version of a divided mind view I developed draws a distinction between motivation-encompassing attitudes located in the deliberative compartment versus the emotional compartment, and says that the motivational base for willpower is exclusively the former. Holton's version of a divided mind view draws a distinction between one's resolutions and one's wayward inclinations, and says, in effect, that the motivational base for willpower is exclusively one's resolutions. Holton does not directly discuss the puzzle of synchronic self-control. Were he to discuss it, however, he would largely be able to help himself to the arguments already adduced in favor of my version of the divided mind view (e.g., the argument based on the existence of full-blooded exercises of willpower), as well as a number of arguments still to come, in order to support his own version of the divided mind view.

I prefer my own version of the divided mind view over Holton's because I take there to be a tight connection between one's practical judgments on the one hand and one's intentions and resolutions on the other. In particular, I believe that a necessary condition for one's forming an intention or resolution to *A* is one's practically judging that *A*-ing is the thing to do.²⁶ Thus I did not try to emphasize the role of intentions and resolutions in supporting willpower, since, on my view, these motivation-encompassing attitudes bear an essential connection to one's practical judgments. Holton rejects the view that intentions and resolutions are essentially tied to one's practical judgments in this way. Indeed, the denial of this connection forms the basis of his highly influential view that *akrasia* (understood as acting contrary to one's best judgment) is not equivalent to weakness of will (understood as acting contrary to one's resolutions) (Holton 1999). The question of whether intentions and resolutions are essentially tied to practical judgments as I believe, or whether they lack this tie as Holton believes, is a complex one that introduces a range of considerations in action theory and moral psychology. Hence I cannot resolve this issue here. Instead, I leave it that Holton's account of willpower is itself a kind of divided mind view, and indeed a quite plausible one. I will, however, be assuming my own version of the divided mind account of willpower in what follows.

Part 2: Which Account of Willpower is Correct?

In response to the puzzle of synchronic self-control, three accounts have been presented that propose to explain how willpower is possible, the non-actional account, the ancillary action account, and the divided mind account. In this final section of the paper, I present three lines of evidence that, I believe, favor the divided mind account of willpower over the alternatives.

2.1 The Existence of Full-Blooded Willpower

The first argument in favor of the divided mind account has already been discussed throughout the first part of the paper. Of the three accounts of willpower, only the divided mind account makes sense of exercises of *full-blooded* willpower, i.e., willpower that is both actional and non-ancillary. Common sense understands willpower to be both actional and non-ancillary, as illustrated by the case of Mike discussed earlier. It is an advantage of the divided mind account that it makes sense of this especially strong kind of self-control to which common sense is committed.

Nonetheless, if the success of a philosophical theory depends substantially on whether or not some phenomenon occurs in the world, then arguments rooted in common sense may not be enough. These arguments will often be stronger if there is *also* empirical evidence for the relevant phenomenon. In this spirit, I present two lines of empirical evidence that support the existence of full-blooded willpower. The first supports the existence of actional willpower, while the second supports the existence of non-ancillary willpower.

Emotion Regulation and Actional Willpower

Emotion regulation refers to the collection of mental processes by which a person alters the intensity, duration, content, or quality of an occurrent emotion episode (Gross 1998). Emotion regulation is a broad category that encompasses a variety of distinct strategies, and the folk term ‘willpower’ likely picks out a subset of emotion regulation strategies. In this section, I want to focus on how the emerging body of research surrounding emotion regulation provides evidence for actional willpower.

Over the last decade, a number of studies have used functional neuroimaging to better understand the brain basis of emotion regulation (see Ochsner and Gross 2005, Phan and Sripada In Press for reviews). The typical set-up of these studies goes as follows. Subjects in the experiments are presented with pictures that are known from prior research to reliably evoke robust emotional responses, typically fear and disgust. They are then taught how to use specific emotion regulation strategies to attenuate their emotional responses to the pictures. One of the most common strategies used in these

studies is *reappraisal*, which involves reinterpreting the meaning of a situation. This strategy resembles that described by Kennett & Smith in which Frog and Toad picture the cookies not as sources of pleasure but rather as lumps of fat sitting their stomach, in order to attenuate their desire to eat the cookies. Another common strategy is *distancing*. Subjects are asked to view the picture from the perspective of a detached observer, e.g., as if one is an anthropologist merely observing, but not in any way involved in, the pictured scene.²⁷

After training, subjects are then placed in an MRI scanner that records patterns of brain activation, and are presented with a new set of pictures that they have never before seen. On some trials they are asked to regulate their emotional responses to the pictures in accordance with the regulation strategies previously learned. This is usually done by presenting a cue immediately prior to the start of each trial that signals whether the impending trial is a 'regulate' trial. On other trials, they are asked to just passively attend to the emotional pictures without attempting to regulate their responses. Because the 'regulate' and 'attend' trials are matched in all respects but for the presence of regulation, one can contrast measured brain responses to the two kinds of trials against each other to gain insight into brain circuits specifically involved in emotion regulation.

There are now several dozen studies that have used the preceding experimental design, or something quite similar to it. Results from these studies show that emotion regulation relies on a set of interconnected neural circuits located primarily in the prefrontal cortex. Greater activation in this prefrontal network has been found to predict 1) reduced activation in regions known to subservise emotion (such as the amygdala and insula); 2) reduced ratings by subjects of their post-picture levels of "negative affect" and 3) reduced physiological manifestations of emotion such as galvanic skin responses, a measure of emotional arousal. These findings suggest that the prefrontal network plays a role in *inhibiting* regions that underwrite the production and maintenance of an emotion episode [see Phan and Sripada (In Press) for detailed descriptions of individual studies].

For our purposes, the point I want to highlight is that these emotion regulation studies provide strong evidence for the existence of *actional willpower*. The defense of this claim can be broken into two parts. First, there is good reason to believe that these studies are assessing *willpower* because a) the strategies taught to subjects in these studies are plausible as ones that capture what is meant when the folk talk about 'exercises of willpower'; and b) subjects' deployment of these strategies produces objective evidence of doing precisely what willpower is supposed to do, that is attenuating the strength of one's occurrent motivation-encompassing attitudes.²⁸

Second, subjects' use of these emotion regulation strategies appears to be *actional*. The experiments require that subjects deploy emotion regulation

only on certain trials while on other trials they passively attend to the pictures without regulating. Subjects themselves come to know whether a given trial will require the deployment of regulation only a few moments before the trial is about to begin. The only plausible explanation for how subjects manage to engage in emotion regulation only on the trials in which the appropriate signal is given, and not on the other trials, is that once the appropriate signal is given, subjects *consciously* and *intentionally* engage in emotion regulation to attenuate the strength of their emotional responses. In short then, these studies provide clear evidence for the actional exercise of willpower.

Exposure Therapy and Non-Ancillary Willpower

Exposure therapy is a widely used and highly efficacious treatment for anxiety disorders (Beck 1976, Foa, et al. 2007, Powers, et al. 2010). This therapy relies on the idea that repeated, prolonged presentation of a fear-inducing stimulus will eventually desensitize the brain mechanisms responsible for triggering a fear response to that stimulus. In order to see how exposure therapy supports the existence of non-ancillary willpower, it is useful to present a specific case.

Paul is a 35 year-old man who has acrophobia, i.e., fear of heights. His fear is so severe that he avoids most any situation that requires being exposed to heights. This includes driving on elevated overpasses, looking out of the window of tall buildings, standing on balconies, climbing on ladders, and much else. Paul recently passed up his dream position with a consulting firm because their penthouse office suite could only be reached with a glass elevator.

Paul enters exposure therapy. He sees his therapist every week for 1-hour long exposure sessions. During each session, he is confronted with fear-inducing height-related stimuli. Paul is instructed that desensitization only occurs when he experiences the distressing emotions, so he must not try to regulate the experience of emotion. But self-control is still very much required. He must suppress his natural tendency to *withdraw* from the fear-inducing stimuli and instead *approach* them.

Paul's level of distress is recorded on a standardized measure, called the Subjective Units of Distress Scale (SUDS), that ranges from 0 to 100, where 0 is no distress at all, and 100 is the worst distress one can imagine. His therapist chooses the height-related stimuli carefully so that Paul's distress rating hits around 70–80 and it is maintained at that level for most of the 1-hour session. After several sessions, Paul finds he can tolerate progressively more powerful height-related scenarios. After a dozen sessions, Paul is placed on the fire escape on the top floor of a 20-story building. He has a harness for safety. Though it is intensely difficult, Paul manages to climb down to the ground on the fire escape.

For an exercise of willpower to count as genuinely non-ancillary, then two conditions should hold. First, the exercise of willpower unfolds over

some extended interval of time throughout which the agent's wayward desire is his overall strongest. Second, the exercise of willpower fully blocks or prevents the person from ever acting on his strongest desire (i.e., the exercise of willpower is *instead of* rather than *concurrent with* the person's acting on the wayward desire). Are these conditions met in the case of exposure therapy?

Several lines of evidence support the claim that throughout each exposure session, Paul's strongest desire is to withdraw from the height-related stimuli. First, Paul has a long history of avoidance of heights that severely interferes with his daily functioning. His passing up his dream job at a consulting firm illustrates how powerful is his fear of heights (and the associated desire to avoid heights). Second, Paul reports his level of subjective distress throughout each session, and it is not uncommon for his SUDS score to rise above 80. This means that Paul is approaching the worst fear he can possibly imagine. Third, some of the exposure scenarios are quite extreme. It is typical in exposure therapy to *overshoot the target* by using highly exaggerated stimuli, such as Paul's climbing down the fire escape. This generates additional desensitization of fear mechanisms, which serves as a buffer in anticipation that there will be some regress after the exposure treatment is completed. For example, a person with contamination fears who cannot even touch a doorknob may eventually be asked to place his hands in a toilet bowl for an extended period of time. It is plausible that these extreme scenarios generate intense desires to withdraw that are indeed the agent's strongest.

Turning now to the second condition for non-ancillary willpower, it does appear that Paul's exercise of willpower fully prevents him from ever acting on behalf of his strongest desire. When confronted with a height-related stimulus, such as being at the very edge of a rooftop, Paul experiences fear and feels a powerful urge to *withdraw* and keep his distance from the edge. Keep in mind Paul does *not* engage in antecedent-focused emotion regulation strategies such as reappraisal. He allows the emotion experience to occur unimpeded. Instead, Paul's exercise of willpower is directed at *suppressing* the motivational strength of the desire to withdraw from the edge or *preventing* the ability of this desire from ever being translated into action, or both. As noted in section 1, these are response-focused strategies and might involve his *distancing* himself from the desire to withdraw or *blocking* the motivational efficacy of the desire to withdraw. Because Paul's exercise of willpower is successful, he does not act on his desire to withdraw from the edge. Indeed, though it is intensely difficult, he is actually able to instead *approach* the edge. Thus Paul's exercise of willpower is non-ancillary because it does *not* co-occur with Paul's acting on behalf of his strongest desire, i.e., to withdraw from the edge of the roof, and indeed it co-occurs with Paul's acting *contrary* to this wayward desire.

To sum up, neuroimaging studies of emotion regulation support the existence of *actional* willpower, and the workings of exposure therapy support the

existence of *non-ancillary* willpower. The divided mind account of willpower is unique in making sense of actional, non-ancillary willpower, so the preceding lines of evidence support the account.

2.2 *The Fundamental Asymmetry of Willpower*

Let us turn now to a second argument in support of the divided mind account of willpower. The operation of willpower involves a basic *asymmetry*. Willpower is exercised in the service of one's practical judgment and for the purposes of suppressing one's wayward desire. The reverse does not occur; one's wayward desire does not exercise willpower in order to suppress the motivations that arise from one's practical judgment. This observation is perhaps so obvious, it might strike one as strange that anyone should take time to draw attention to it. Though obvious, it is important to highlight this asymmetry because it raises a question that has not been widely discussed. In virtue of what does this asymmetry arise?

The divided mind account of willpower has an exceedingly simple explanation. Willpower is a proprietary action and the 'proprietor' is, *exclusively*, the deliberative motivational system. The observation that willpower is always exercised in the service of, and never against, one's practical judgments is thus explained by the fact that it is exclusively the motivation-encompassing attitudes within the deliberative motivational compartment that constitute the motivational base for willpower.²⁹

The non-actional account and ancillary action account of willpower do not have an obvious explanation for the asymmetry of willpower. On these accounts, willpower is something that helps one's weaker desire to prevail in a motivational contest over one's stronger desire, *regardless of the source or type of either desire*. Thus *symmetry* appears to be built into these accounts in a fairly fundamental way, and this in turn leads to some strange results. Suppose a man has a very strong desire to eat a slice of cake, but a still stronger desire to stay on his diet. On Mele's ancillary action view, the desire to eat the cake can still win out over the desire to stay on the diet in the following way. The man exerts willpower against his desire to stay on the diet by uttering to himself a self-command: "Eat the cake!" In this way, he attenuates his desire to stay on the diet and gets himself to eat the cake.

The preceding story seems bizarre, perhaps even amusingly so, because temptation-directed desires are simply *not the kinds of things* that exercise willpower to suppress the desires that arise from one's practical judgments. The symmetry built into the non-actional and ancillary action accounts of willpower is thus a problem.³⁰ Willpower is fundamentally asymmetric, and it is an advantage of the divided mind account that it quite naturally explains why the asymmetry arises.

2.3 Dual-Process Models in Cognitive Science

An additional argument in favor of the divided mind account of willpower is that the divided motivational architecture being proposed is an instance of a more general kind of architecture, called a dual-process architecture, that is now well accepted across many domains in cognitive science. It thus helps the divided mind approach that it has empirical support that arises *independently* of its being adduced as a solution to the puzzle of synchronic self-control.³¹

In general form, a dual-process model postulates that information processing in some psychological domain is subserved by two distinct systems. One system [often referred to as ‘system 1’, the terminology is from Stanovich and West (2000)] is relatively fast, automatic, performs relatively simple associative operations, and has access to only narrow stores of domain-specific information. The other system (often referred to as ‘system 2’) is slow, consciously controlled, uses linguistic/logical representations, and has access to much larger stores of domain-general information. Inputs are processed by both system 1 and 2 *simultaneously*, and it is possible for the systems to produce discrepant or divergent outputs. In these situations, at least some kinds of dual-process models propose that system 2 exerts regulatory control over system 1 (Gilbert 1999). Here regulatory control is a *proprietary* function of system 2, and it overrides, suppresses, or modulates outputs from system 1.

Dual-process models are now ubiquitous in cognitive and social psychology. They are well-accepted explanations for how information processing in the mind/brain is organized in domains such as judgment under uncertainty (Gilovich, et al. 2002, Kahneman 2003), theoretical reasoning (Sloman 1996, Stanovich 2005, Stanovich and West 2000), and social perception and decision-making (Chaiken, et al. 1989, Strack and Deutsch 2004). Most important for the present purposes, theorists have developed dual-process models of *motivation* in which, very roughly, emotional mechanisms occupy the role of system 1, while practical reasoning systems occupy the role of system 2 (Hofmann, et al. 2009, Loewenstein 1996, Metcalfe and Mischel 1999, Friese, et al. 2009).³² Willpower is incorporated in these models as the regulatory control process by which system 2 suppresses or modulates system 1. Neuroscientific evidence for a dual-process architecture of motivation is accumulating and includes evidence from animal models, neuropsychological studies of brain lesions, and human cognitive neuroscience (see Bechara 2005, Bechara and Van Der Linden 2005, Sanfey, et al. 2006 for reviews).

It is useful to reflect on the question of why dual-process architectures are so prevalent across so many domains of cognition. Here I offer what I take to be a plausible answer for the case of motivation, though the logic clearly generalizes for the other domains in which dual-process architectures are present.

Suppose a creature had *unlimited time* to select actions. Then in this case, it makes sense for the creature to have mental processes structured along the lines of system 2. These processes are slow and place substantial demands on cognitive resources. However, they have the advantage of integrating large amounts of information, keeping track of complex trade-offs, and being able to represent consequences that unfold over much longer intervals of time. If, however, a creature had *almost no time at all* to select actions, then it makes sense for the creature to have mental processes structured along the lines of system 1. These processes are fast, automatic, and they can be designed to produce characteristic action tendencies (e.g., fleeing in the face of a threat) that have proven to be advantageous over evolutionary time.

What happens, however, if a creature needs to select actions in *both* types of contexts, and especially if it often *does not know* which of these contexts it is presently in (i.e., it often does not know whether there will turn out to be enough time to select actions the ‘system 2 way’)? In this case, it makes sense to outfit the creature with *both* types of mental processes, i.e., both system 2-like processes and system 1-like processes, with external information processed by both systems *at the same time*. As the slow serial operations of system 2 churn along, system 1 has already readied potential action plans for immediate deployment. Now, for this to work, some arrangement needs to be in place to handle situations in which the two systems reach divergent conclusions about which action to undertake. One attractive arrangement is to equip system 2 with mechanisms of regulatory control that override, suppress, or otherwise modulate outputs from system 1. Regulatory control of course has to be *asymmetric*—it makes little sense to allow both systems to override each other. Additionally, it makes sense to place the regulatory control in the hands of system 2, rather than system 1, because system 2 tends to have integrated more information regarding one’s long-term interests in rendering its motivational verdicts.

If these functional speculations are correct, then there are excellent *performance* reasons for creatures like us to have been outfitted with a dual-process motivational architecture. Moreover, it makes sense for this architecture to be structured so that the system 2-like compartment exercises *asymmetric, propriety* regulatory control over the system 1-like compartment. The divided mind account of willpower of course proposes a picture of how motivation is divided and how regulatory control is implemented that is very much along these lines.

Let me bring together the preceding strands of evidence. First, dual-process architectures are common across multiple domains of cognition. Second, there is already significant evidence from psychology and neuroscience for a dual-process structure in the domain of motivation specifically. Third, there are excellent functional/performance-based reasons for motivation-related processing to be structured along dual-process lines. Since the divided mind account of willpower postulates a division in motivation

that closely resembles the standard dual-process structure, the preceding lines of evidence support the divided mind account of willpower.

It would be *brash* to propose an entire theory about the structure of motivation based on nothing more than the need to address a single philosophical puzzle, i.e., the puzzle of synchronic self-control. The fact that the dual-process architecture embodied in the divided mind account fits with a substantial body of findings from cognitive science thus not only provides extra evidence for the view, it also helps address this charge of brashness.

To sum up, three lines of evidence support the divided mind account of willpower in comparison to the alternatives: the existence of full-blooded willpower, the asymmetry of willpower, and the connections between the divided mind account of willpower and dual-process architectures more generally. These lines of evidence provide convergent support for the divided mind view.

3. Conclusion

Philosophical puzzles are often born when we take the time to scrutinize the ordinary. This is certainly true in the case of the puzzle of synchronic self-control. Nothing could be more ordinary than the exercise of willpower. And yet, willpower is fundamentally puzzling because it seems to require that the person both most wants to pursue a wayward course of action and most wants to resist pursuing this course of action, and this seems impossible.

I reviewed and criticized two accounts of willpower that try to make sense of how willpower is possible, a non-actional account derived from the work of Jeanette Kennett and Michael Smith and an ancillary action account from Alfred Mele. I then proposed the divided mind view that holds that motivational architecture is divided between two motivational systems, a deliberative system and an emotional system, and willpower is a *proprietary action* exclusively available to the deliberative system. I argued that that this account is unique in being able to make sense of the existence of full-blooded willpower, i.e., exercises of willpower that are both actional and not ancillary to some other main action that serves the wayward desire.

The divided mind account raises problems of its own, in terms of how one can make sense of a motivationally weaker desire defeating a stronger one, and how actions that arise from a part of a person's psychology can be said to belong to whole person. These problems, however, can be readily addressed. The divided mind account is, I've argued, best supported by the overall evidence, including philosophical and empirical lines of evidence.³³

Notes

¹ This formulation of the principle is from Davidson (1980). It is false as stated for a number of different reasons [see Mele (2003)]. Some of the reasons why the principle is false are relevant

to the topic of the current paper, though many of the reasons are not so relevant. Even if false in multiple ways, the principle provides a useful starting point for discussion. In section 1.3, I propose a divided mind account of willpower, and in 1.4, I show specifically why this account renders the principle false.

² The dilemma represents the conjunction of two puzzles about self-control posed in Mele (1987).

³ Here I follow most philosophers in understanding the locution ‘the person most desires to *A*’ to mean ‘the desire to *A* is the person’s (motivationally) strongest.’

⁴ *Ss essentially motivation-encompassing A attitude* is, at a first approximation, an attitude that represents *Ss A-ing* as part of its content, and that non-deviantly contributes to *Ss A-ing* in all possible scenarios in which *S* has this attitude [see Mele (1995, 2008) for a detailed account]. It is controversial which states are motivationally-encompassing in this sense. For the purposes of this paper, I will assume the following count: action-desires, intentions, certain emotions such as fear, and practical judgments about what, all things considered, is the thing to do.

⁵ Distancing can also be used as an antecedent-focused strategy (see section 2.1).

⁶ Eastern meditation techniques frequently incorporate distancing-based methods for attenuation or suppression of desires and other motivation-encompassing attitudes. These methods inform relatively new psychotherapeutic practices, such as mindfulness-based psychotherapies.

⁷ Mele calls a willpower strategy broadly similar to this *brute resistance*. See Mele (1987, pg. 26).

⁸ One might raise the following challenge: There is a conceptual tie between motivation to *A* and *A-ing*, so that it is not possible for one’s motivation to *A* to be unchanged while one’s actually *A-ing* is foreclosed. The problem with this view is that being motivated to *A* is conceptually tied not to *actually A-ing*, but rather to the *disposition* to bring about *A* (see section 1.4). The dispositional properties of an object can be retained even if the manifestations of the disposition are systematically blocked. For example, a porcelain vase retains its fragility, even if it is wrapped in Styrofoam packaging. In a similar way, the motivation to *A* can remain intact, even if willpower blocks the translation of this motivation into action.

⁹ Using the distinction introduced at the start of this section, Kennett and Smith’s account of willpower primarily seeks to account for certain kinds of antecedent-focused strategies.

¹⁰ The terminology is from Alfred Mele (e.g., Mele 2003, pg. 16). Action-desires contrast with standing desires that certain states of affairs obtain, such as one’s desire for world peace.

¹¹ The behavioral economist George Loewenstein calls these processes ‘visceral factors’ (Loewenstein 1996) while the neuroscientist Antoine Bechara call them ‘impulses’ (Bechara 2005).

¹² The preceding characterization of emotions is based on points of consensus in the literature, including views as otherwise different as Solomon (1976) and Griffiths (1998).

¹³ Gary Watson puts the point elegantly: “It is an essential feature of the appetites and passions that they engender (or consist in) desires whose existence and persistence are independent of the person’s judgment of the good.” (Watson 1975).

¹⁴ See Fodor (1983, pg. 52–55) for more on mandatoriness of cognitive processes.

¹⁵ Alternatively, one might say the motivational base for *Ss ϕ -ing* at *t* is the conjunction of the motivational bases for all of *Ss* motivation-encompassing attitudes to ϕ at *t*.

¹⁶ Thanks to an anonymous reviewer for suggesting that I discuss a case along these lines.

¹⁷ Thanks to an anonymous reviewer for this way of formulating the objection.

¹⁸ In an insightful article on the Law of Desire (Clarke 1994), which also gave the ‘Law’ this name, Randolph Clarke offers a helpful analogy. Suppose two air-jets are facing each other with a ping-pong ball in between, and there are no other air-jets in the picture. Must the ping-pong ball move away from the jet that blows more strongly? Clarke notes this ‘Law of Air-Jets’ will not hold if the table is inclined sufficiently steeply towards the stronger air-jet. The point of the example is that there may be *other* influences that, like the incline, aid a weaker desire that are not themselves desires. In this way, a weaker desire could defeat a stronger one. Clarke offers

some speculations about what specifically in the mind/brain plays the role of the incline. The divided mind account of willpower says it is regulatory systems implementing willpower that play this role.

¹⁹ See Mele (1998) for an account of motivational strength along these lines.

²⁰ The philosophical debate about desires tends to center around the question of which features are *necessary* and *sufficient* for a mental state to be a desire. My interest in this section is in a somewhat different topic. I simply assume that the standard theory of desire is correct—desires are mental states that dispose the person to act in desire promoting ways. (I leave it open that additional qualifiers are needed to avoid certain kinds of counter-examples.) The issue I am probing here concerns the specific causal powers possessed, perhaps only *contingently*, by desires. By means of what specific mental mechanisms and processes do desires get their bearers (and in particular, creatures like *us*) to act in desire promoting ways?

²¹ A full and accurate inventory of D-powers may need to await the arrival of a true and complete scientific theory of desire.

²² Here again, a full and accurate inventory of R-powers may need to await the arrival of a complete and true scientific theory of willpower.

²³ In Sripada (2010), I provide another line of argument in favor of the view that the powers of regulatory systems are relatively independent of the motivational powers of the desire that supports willpower's exercise. This line of argument focuses on failures of willpower due to lack of motivation versus failures due to 'fixed ceilings' in the efficacy of regulatory systems.

²⁴ Cases somewhat similar to this are discussed in Smith (2005, 2008).

²⁵ Neil Levy criticizes Holton's talk about willpower as a 'distinct mental faculty' in Levy (2011). According to Levy, willpower is simply the operation of (what I have called) the deliberative motivational system, and talk of a distinct mental faculty of willpower is misleading. Holton (2009) partly accepts Levy's criticism, and clarifies that willpower is not so much a distinct faculty, but rather it is something more like a specific skill that agents 'actively employ' (pg. 134). Thus, says Holton, the exercise of willpower cannot be reduced to the having of a belief, desire, or intention. My own position has affinities with both views. Like Levy, I hold that the exercise of willpower is intimately connected with the operation of the deliberative motivational system. (Though on my view, the connection is more nuanced: Willpower is not *identical* to the operation of the deliberative motivational compartment. Rather, this compartment serves as the *exclusive motivational base* for the exercise of willpower.) Like Holton, I hold that the exercise of willpower is a specific mental action (or a family of closely related actions, see Section 1) that agents actively employ.

²⁶ This rough characterization requires a number of qualifications to deal with a host of immediate counter-examples. But I believe that suitably qualified, the principle is indeed defensible.

²⁷ Here distancing is presented as an *antecedent-focused* strategy. In section 1, I presented an example of how distancing could also be used as a *response-focused* strategy.

²⁸ These studies actually produced evidence that regulation decreases subjective ratings of "negative affect" and levels of physiological arousal. We would need to assume there is a link between what is measured by these methods and the motivational strength of action-desires elicited during the emotion episode, which seems like a plausible assumption.

²⁹ It may strike some as *ad hoc* that willpower turns out to be a proprietary action exclusively available to the deliberative motivational system, rather than being exclusively available to the emotional motivational system. However, a bit later, I note the connections between the divided mind account of motivation being developed here and dual-process theories of cognitive architecture more generally. There are, as it turns out, excellent *performance* reasons for information processing to be divided between two systems. Performance demands also appear to favor that the 'slow cognitive' system exercise asymmetric top-down control over the 'fast reactive' system. These observations might help dispel the sense that the deliberative system's having exclusive access to willpower is *ad hoc*.

³⁰ Mele defines self-control as the ability to “master motivation that is contrary to one’s best judgment” (Mele 1987, pg. 54). Does this not account for the asymmetry of willpower being discussed here? No. Mele’s definition of self-control only *stipulates* that exercises of self-control are the subset of cases of ‘mastering motivation’ that are used to support one’s practical judgment against a wayward desire. One might then ask why the other subset (in which abilities to master motivation are used to support a weaker wayward desire against all things considered practical judgment) is empty. Mele’s account does not itself provide an explanation for why. Put another way, Mele’s account of how self-control works does not itself preclude, and indeed fully allows, that there should be a complementary notion of self-control (call it self-control*) in which a person uses abilities to master motivation to support her weaker wayward desire against a stronger practical judgment. The present question then becomes why do we never see exercises of self-control*?

³¹ Some of this evidence for dual-process architectures in various domains of cognition is also discussed in Sripada (2010).

³² See also Levy (2011) for a philosophically-informed discussion of dual systems approaches to motivation.

³³ Versions of parts of this paper were presented at: University of Michigan Decision Consortium Group, the Moral Psychology Research Group (Pittsburgh, PA), and the ‘Philosophy and Science of Weakness of Will’ series at Amherst College. I am grateful to these audiences for their helpful comments. Also, thanks very much to the following for their comments on drafts of this paper, or earlier versions of parts of this paper: Alfred Mele, Josh May, Sarah Buss, and Jacob Ross.

References

- Bechara, A. (2005). “Decision Making, Impulse Control and Loss of Willpower to Resist Drugs: A Neurocognitive Perspective,” *Nat Neurosci* vol. 8, no. 11: 1458–63.
- Bechara, A., and Van Der Linden, M. (2005). “Decision-Making and Impulse Control after Frontal Lobe Injuries,” *Curr Opin Neurol* vol. 18, no. 6: 734–9.
- Beck, Aaron T. (1976). *Cognitive Therapy and the Emotional Disorders*. (New York: International University Press).
- Bratman, Michael. (1987). *Intentions, Plans, and Practical Reason*. (Cambridge, MA: Harvard University Press).
- Chaiken, S., Liberman, A. and Eagly, A. H. (1989). “Heuristics and Systematic Processing within and Beyond the Persuasion Context.” in Uleman and Bargh, eds. *Unintended Thought* (New York: Guilford Press) (212–52).
- Clarke, Randolph. (1994). “Doing What One Wants Less: A Reappraisal of the Law of Desire,” *Pacific Philosophical Quarterly* vol. 75, no. 1: 1–11.
- Davidson, Donald. (1980). “How Is Weakness of the Will Possible?” in Davidson, ed. *Actions and Events* (Oxford: Clarendon Press) (21–42).
- Foa, Edna B., Hembree, Elizabeth A. and Rothbaum, Barbara Olasov. (2007). *Prolonged Exposure Therapy for Ptsd: Emotional Processing of Traumatic Experiences: Therapist Guide*, Prolonged Exposure Therapy for Ptsd: Emotional Processing of Traumatic Experiences: Therapist Guide. (New York, NY US: Oxford University Press).
- Fodor, J. A. (1983). *The Modularity of the Mind*. (Cambridge, MA: MIT Press).
- Friese, Malte, Wänke, Michaela, and Hofmann, Wilhelm. (2009). “Unscrambling Self-Regulatory Behavior Determination: The Interplay of Impulse Strength, Reflective Processes, and Control Resources.” in Forgas, Baumeister and Tice, eds. *Psychology of Self-Regulation: Cognitive, Affective, and Motivational Processes*. (New York, NY US: Psychology Press) (53–71).
- Gilbert, Daniel. (1999). “What the Mind’s Not.” in Chaiken and Trope, eds. *Dual-Process Theories in Social Psychology* (New York: Guilford Press) (3–11).

- Gilovich, T., Griffin, D., and Kahneman, D. (2002). "Heuristics and Biases: The Psychology of Intuitive Judgment." (Cambridge, MA: Cambridge University Press).
- Griffiths, Paul. (1998). *What Emotions Really Are*. (Chicago: University of Chicago Press).
- Gross, J. J. (1998). "Antecedent- and Response-Focused Emotion Regulation: Divergent Consequences for Experience, Expression, and Physiology," *J Pers Soc Psychol* vol. 74, no. 1: 224–37.
- Gross, J. J., and Levenson, R. W. (1993). "Emotional Suppression: Physiology, Self-Report, and Expressive Behavior," *J Pers Soc Psychol* vol. 64, no. 6: 970–86.
- Gross, James J. (1998). "The Emerging Field of Emotion Regulation: An Integrative Review," *Review of General Psychology* vol. 2, no. 3: 271–99.
- Hofmann, Wilhelm, Friese, Malte, and Strack, Fritz. (2009). "Impulse and Self-Control from a Dual-Systems Perspective," *Perspectives on Psychological Science* vol. 4, no. 2: 162–76.
- Holton, Richard. (1999). "Intention and Weakness of Will," *The Journal of Philosophy* vol. 96, no. 5: 241–62.
- . (2009). *Willing, Wanting, Waiting*. (New York: Oxford University Press).
- D'Arms, Justin, and Jacobson, Daniel. (2003). "The Significance of Recalcitrant Emotions (or Anti-Quasi-judgmentalism)." in Hatzimoysis, ed. *Philosophy and the Emotions* (Cambridge: Cambridge University Press).
- Kahneman, D. (2003). "A Perspective on Judgment and Choice: Mapping Bounded Rationality," *American Psychologist* vol. 58: 697–720.
- Kennett, Jeanette, and Smith, Michael. (1996). "Frog and Toad Lose Control," *Analysis* vol. 56, no. 2: 63–73.
- . (1997). "Synchronic Self-Control Is Always Non-Actional," *Analysis* vol. 57, no. 2: 123–31.
- Levy, Neil. (2011). "Resisting 'Weakness of the Will,'" *Philosophy and Phenomenological Research* vol. LXXXII, no. 1: 134–55.
- Loewenstein, George. (1996). "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes* vol. 21: 272–92.
- Mele, A.R. (1997). "Underestimating Self-Control: Kennett and Smith on Frog and Toad," *Analysis* vol. 57, no. 2: 119–23.
- Mele, Alfred. (1992). *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. (New York: Oxford Univ Press).
- . (1987). *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. (New York: Oxford University Press).
- . (2003). *Motivation and Agency*. (New York: Oxford University Press).
- Mele, Alfred R. (1998). "Motivational Strength," *Nous* vol. 32, no. 1: 23–36.
- Mele, R. (1998). "Synchronic Self-Control Revisited: Frog and Toad Shape Up," *Analysis*, vol. 58, no. 4: 305–10.
- Metcalf, J., and Mischel, W. (1999). "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower," *Psychol Rev* vol. 106, no. 1: 3–19.
- Muraven, Mark, and Baumeister, Roy F. (2000). "Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?," *Psychological Bulletin* vol. 126, no. 2: 247–52.
- Ochsner, Kevin N., and Gross, James J. (2005). "The Cognitive Control of Emotion," *Trends in Cognitive Sciences* vol. 9, no. 5: 242–49.
- Phan, K. Luan, and Sripada, Chandra Sekhar. (In Press). "Emotion Regulation." in Vuilleumier and Armony, eds. *Handbook of Human Affective Neuroscience* (Cambridge, UK: Cambridge University Press).
- Powers, Mark B., Halpern, Jacqueline M., Ferenschak, Michael P., Gillihan, Seth J., and Foa, Edna B. (2010). "A Meta-Analytic Review of Prolonged Exposure for Posttraumatic Stress Disorder," *Clinical Psychology Review*.

- Sanfey, Alan G., Loewenstein, George, McClure, Samuel M., and Cohen, Jonathan D. (2006). "Neuroeconomics: Cross-Currents in Research on Decision-Making," *Trends in Cognitive Sciences* vol. 10, no. 3: 108–16.
- Scanlon, T. M. (1998). *What We Owe Each Other*. (Cambridge, MA: Harvard University Press).
- Shizgal, Peter. (1997). "Neural Basis of Utility Estimation," *Current Opinion in Neurobiology* vol. 7, no. 2: 198–208.
- Sloman, S. A. (1996). "The Empirical Case for Two Systems of Reasoning," *Psychological Bulletin* vol. 119: 2–22.
- Smith, Angela. (2008). "Control, Responsibility, and Moral Assessment," *Philosophical Studies* vol. 138: 367–92.
- . (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* vol. 115: 236–71.
- Smith, M. (1987). "The Humean Theory of Motivation," *Mind* vol. 96: 36–61.
- Solomon, Robert. (1976). *The Passions*. (Notre Dame, IN: University of Notre Dame Press).
- Sripada, Chandra Sekhar. (2010). "Philosophical Questions About the Nature of Willpower," *Philosophy Compass* vol. 5: 793–805.
- Stanovich, K. E., and West, R. F. (2000). "Individual Differences in Reasoning: Implications for the Rationality Debate," *Behavioral and Brain Sciences* vol. 23: 645–65.
- Stanovich, K.E. (2005). *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. (Chicago: The University of Chicago Press).
- Strack, Fritz, and Deutsch, Roland. (2004). "Reflective and Impulsive Determinants of Social Behavior," *Personality and Social Psychology Review* vol. 8, no. 3: 220–47.
- Strawson, Galen. (2009). *Mental Reality*. (A Bradford Book).
- Watson, Gary. (1987). "Free Action and Free Will," *Mind* vol. 96: 145–72.
- . (1975). "Free Agency," *The Journal of Philosophy* vol. 72, no. 8: 205–20.