Philos Stud (2014) 170:359–375
DOI 10.1007/s11098-013-0224-5

Self-control, motivational strength, and exposure therapy

Alfred R. Mele

Published online: 24 October 2013

Abstract Do people sometimes exercise self-control in such a way as to bring it about that they do not act on present-directed motivation that continues to be motivationally strongest for a significant stretch of time (even though they are able to act on that motivation at the time) and intentionally act otherwise during that stretch of time? This paper explores the relative merits of two different theories about synchronic self-control that provide different answers to this question. One is due to Sripada (Noûs 1–38, 2012) and the other to Mele (Irrationality, 1987; Autonomous agents, 1995; Motivation and agency, 2003). Special attention is paid to evidence Sripada offers for an affirmative answer to the question, and some guidance is offered on the project of finding evidence for an affirmative answer.

Keywords Action · Desire · Motivation · Motivational strength · Self-control

The idea that there is a tight connection between what we are most strongly motivated to do and what we intentionally do has led to some interesting puzzles about self-control. How philosophers formulate and attempt to resolve the puzzles is influenced by their preferred formulation of the idea about motivational strength. In Sect. 1, I provide some background, including illustrations of this point about influence. In Sect. 2, I turn to a novel position on self-control defended by Sripada (2012). In Sect. 3, I assess the evidence Sripada offers for a central plank of his position and argue that the evidence does not undermine an alternative position that rejects that plank. I close with a discussion of the project of finding evidence for or against the competing positions.

A. R. Mele (✉)
Florida State University, Tallahassee, FL, USA
e-mail: almele@fsu.edu

1 Some background

On Wednesday, as you are planning your work schedule for the rest of the week, it becomes clear that you will need to work late on Friday in order to meet all your obligations. You judge it best on Wednesday to work late on Friday, but you anticipate being tempted to stop working early on Friday and join your friends at the pub. So you take steps now to decrease the likelihood that you will succumb to temptation later. You send your friends an e-mail message announcing your resolution to work late on Friday and expressing your apologies in advance for missing the weekly gathering. You know that you would be the butt of many jokes if you were to show up at the pub after sending that message, and your plan is to use your desire not to be the butt of jokes to augment your motivation for doing what you judge best. Your strategy works.

In this case, you exercise *diachronic* self-control. The puzzles to be discussed here concern a related phenomenon—*synchronic* self-control. Jeanette Kennett and Michael Smith write: "when [people] exercise synchronic self-control they identify what they most want to do, at a certain time, as something that they shouldn't do, and they try for that reason, at that very time, not to do it" (1997, p. 123). They ask:

> Is it possible to reconcile the fact that people can exercise synchronic self-control, on the one hand, with a truism in the philosophy of action, on the other? The truism is that whenever people do something intentionally at some time, and so whenever they try to do something, they want to do that thing more than they want to do anything else they believe they can do at the time. (Kennett and Smith 1997, p. 123; see Kennett and Smith 1996, p. 63)

Here we have a puzzle about self-control. Given the alleged truism, if at a time an agent wants most to *A* right then, then he *A*-s then, if he intentionally does anything at all then. This is so even if he believes that he should not *A* then, and even if he thinks that he should exercise self-control in support of his not *A*-ing. So how can an agent exercise self-control in support of his not *A*-ing when he wants most to *A* right then? As Kennett and Smith observe, the alleged truism is consistent with an exercise of synchronic self-control, provided that the exercise is not an action (1996, 1997). The alleged truism leaves it open that our agent does not intentionally do anything at the time, even though he wants most to *A* then; and what brings it about that he does not *A* might be a non-actional exercise of self-control that changes what he wants most to do before he has time to act on his strongest desire (Kennett and Smith 1996; 1997, p. 128). This is the heart of their attempted resolution of their puzzle.

Kennett and Smith contend that "all exercises of synchronic self-control are non-actional. They are non-actional because there is no suitable strongest desire to cause an exercise of actional self-control" (1997, p. 128). An example of a non-actional exercise of self-control is non-actionally picturing cookies as lumps of fat at a time at which one wants most to eat a cookie right then (Kennett and Smith 1996, 1997). The images may come to mind without one's actively bringing them to mind, and they may significantly dampen one's desire for a cookie.

Kennett and Smith report that they "would not be worried" if their statement of the alleged truism were "liable to counterexample" (1997, p. 125, n. 2). Their "statement of the truism," they assert, "only needs to be accurate enough for us to both bring out and solve the puzzle about self-control" (p. 125, n. 2). However, the alleged truism certainly restricts the range of admissible resolutions of the puzzle. And the "truism" is false.

Here is the alleged truism again: "whenever people do something intentionally at some time, and so whenever they try to do something, they want to do that thing more than they want to do anything else they believe they can do at the time" (Kennett and Smith 1997, p. 123). And here is a counterexample. What Joe wants to do most now is to dribble a basketball with his left hand while tossing another basketball into the air with his right hand—a compound action. He succeeds in dribbling the ball, but the other ball accidentally falls from his other hand. Kennett and Smith's "truism" yields the verdict that, at the time in question, Joe wanted to dribble a basketball with his left hand more than he wanted to do anything else that he believed he could do at the time. And that verdict is false.

Here is another counterexample. What Ann wants most to do now is pedal her exercycle while reading an e-mail message and listening to her radio—a compound action. Among the things that Ann is doing intentionally at the time are the following three: pedaling her exercycle, reading an e-mail message, and listening to her radio. The alleged truism yields the verdict that Ann wants to do each of these things more than she wants to do anything else she believes she can do at the time. And that verdict cannot possibly be true. (For example, while she is acting, Ann cannot want to pedal her exercycle then more than she wants to read the e-mail message then while also wanting to do the latter then more than she wants to do the former then.)

Might it be that, despite these counterexamples, the alleged truism is "accurate enough for [Kennett and Smith] to both bring out and solve the puzzle about self-control" (Kennett and Smith 1997, p. 125, n. 2)? Even if the answer is yes in the case of their particular puzzle, we may well be more interested in a puzzle about self-control that is grounded in a more plausible alleged truism about intentional action.

Consider the following claim:

> *T*. Whenever we act intentionally, we do, or try to do, what we are most strongly motivated to do at the time.[1]

Claim *T* leaves it open that we can perform more than one intentional action at a time. Suppose that you are intentionally *A*-ing, intentionally *B*-ing, and intentionally *C*-ing at the same time and that you are most strongly motivated to *B* at the time. That is consistent with *T*. Suppose that, at the same time, you are intentionally *A*-ing, intentionally *C*-ing, and trying but failing to *B* and that you are most strongly motivated to *B* at the time. That too is consistent with *T*. Return to Joe. His story is no problem for *T*. He is trying to do what he is most strongly motivated to do—namely, dribble a ball with his left hand while tossing another ball into the air with

[1] I defend qualified versions of this thesis in Mele (1992, Chap. 3, 2003, Chap. 8).

his right hand. Nor is Ann's story a problem. She is doing what she is most strongly motivated to do.

Whereas Kennett and Smith's alleged truism leads naturally to the resolution they offer to their puzzle about synchronic self-control, reflection on *T* suggests another route to an answer to the question how synchronic self-control is possible—one that allows for *actional* exercises of synchronic self-control.[2] In Mele (1998), I distinguished between two kinds of scenario in which there may be a place for synchronic self-control. In what I called *extreme* scenarios (p. 307), an agent is presently engaged in an activity that is at odds with what he judges best (e.g., he is watching TV) and, at the time, he wants most to continue this activity (see Mele 1987, pp. 63, 69–72). In *moderate* scenarios (Mele 1998, p. 307), the agent wants most to *A* (e.g., eat a cookie) straightaway, doing that is against his better judgment, and he is not yet *A*-ing. (It takes time to act.) In both kinds of scenario, I believe, an actional exercise of self-control is possible.

I start with an extreme scenario. Ian is now more strongly motivated to continue watching a golf tournament on television than he is to do anything else now, but he judges it best to turn off the television and get back to work. Can Ian mount a successful actional exercise of self-control—synchronic self-control—in support of his terminating his television viewing and getting back to work? Elsewhere, I argued that the answer is yes (Mele 1987, pp. 69–74) and I replied to various reservations about my answer (Mele 1995, pp. 43–57). My immediate aim is to provide another illustration of the point that how philosophers attempt to resolve puzzles about synchronic self-control is influenced by their view about how motivational strength is linked to intentional action. One who notices that *T* leaves it open that agents sometimes do things intentionally without being most strongly motivated to do them while also doing what they are most strongly motivated to do may be inclined to ask two questions about Ian. First, might he make an attempt at self-control while he is watching TV? Second, might that attempt be successful? In Mele (1987, 1995), I defend affirmative answers to both questions.

I turn to a moderate scenario. Fred wants most to eat a cookie straightaway, but doing so is against his better judgment. The cookie is in his hand; he has not started eating it yet. Can Fred successfully exercise self-control in the service of his better judgment? Notice that neither Kennett and Smith's alleged truism (I dub it *KS*) nor *T* applies straightforwardly to this question. Here they are again:

> *KS*. Whenever people do something intentionally at some time, and so whenever they try to do something, they want to do that thing more than they want to do anything else they believe they can do at the time.
> *T*. Whenever we act intentionally, we do, or try to do, what we are most strongly motivated to do at the time.

On straightforward, literal readings, both theses are about agents' motivational conditions *while* they are acting intentionally. But the question about Fred concerns his motivational condition *before* he does something of special interest.

[2] For problems with Kennett and Smith's attempted resolution, see Mele (1998).

Consider the following version of *T*:

*T**. Whenever we act intentionally, we do, or try to do, something that we are most strongly motivated to do at the time and that we were most strongly motivated to do before that time.

How does *T** apply to Fred? At some time, *t*, Fred was most strongly motivated to eat a cookie straightaway. But what was his motivational condition like after *t*? Was there time for Fred to bring it about that he was not most motivated to eat the cookie straightaway and that he did not eat it?

Elsewhere, I defended an affirmative answer to the last question, taking my lead in part from some neuroscientific work on the "vetoing" of proximal desires and intentions (Mele 1997; 2003, Chap. 8). Proximal desires (or intentions) to *A* include desires (or intentions) to *A* straightaway, to begin *A*-ing straightaway, and (while one is in the process of *A*-ing) to continue *A*-ing. There is evidence about how much time it takes for a proximal intention to flex a wrist to generate measurable muscle activity, and I argued that it is enough time to permit an agent who is most strongly motivated to *A* to revoke even a proximal intention to *A* and refrain from *A*-ing (Mele 1997; 2003, Chap. 8).[3] This is not nearly as strange as it may seem to some readers. Consider a driver who is most strongly motivated to start turning left now and who proximally intends to make the turn. As he is about to start turning, he sees a dog dart into his intended path and does not turn. Fortunately, he had time to adjust to what he saw. Now, if perception can trigger an adjustment of this kind, might something like a self-command be capable of such triggering? For example, if Fred were to say "Stop!" to himself, might he refrain from eating the cookie? And can Fred issue this self-command even though he is most strongly motivated to eat the cookie straightaway. In Mele (1997; 2003, Chap. 8), I defend affirmative answers to questions of this kind. I will not review the arguments here. The purpose of the present discussion is stage setting. Readers of the following sections can judge whether the stage setting was helpful.

2 A "divided mind" view of synchronic self-control

Chandra Sripada has developed a novel response to a puzzle about synchronic self-control (2012). His point of departure is the following "principle":

S. 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. (p. 2)

This principle is false, as Sripada recognizes (2012, pp. 33–34, n. 1). For example, a soccer player who most desires to score a goal on his penalty kick and believes himself free to do so may intentionally kick the ball and unintentionally kick it too far to the right. Even so, Sripada sees *S* as providing "a useful starting point for discussion" (pp. 33–34, n. 1).

[3] For a discussion of the evidence itself, see Mele (2009, Chap. 3).

Sripada understands synchronic self-control differently than Kennett and Smith do. He identifies it with "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" (2012, p. 2). Return to my story about you and the pub. Suppose that on Wednesday, even after judging that it would be best to work late on Friday, you have a desire to stop working early on Friday and join your friends at the pub—an "active" desire concerning action 2 days in the future. And suppose that on Wednesday, in order to attenuate the strength of this active desire, you send the e-mail message described earlier. Then, according to Sripada's way of carving things up, you are exercising synchronic self-control at the time. However, the difference in usage need not be a source of confusion. Sripada's focus is on cases in which synchronic self-control is exercised against *proximal* desires, and that is the locus of the debate he has joined.

To prepare the way for his view, Sripada attempts to knock resolutions proposed by Kennett and Smith and by me out of contention. He sides with me in contending, against Kennett and Smith, that there are actional exercises of synchronic self-control (2012, pp. 4–7).[4] About my own view he writes:

> [it] 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… is that there appear to be still stronger forms of synchronic self-control… that are… *not ancillary* to some other action in support of the stronger wayward desire. (pp. 2–3)

It is true that my resolution of a puzzle about synchronic self-control raised by what I called *extreme* cases features what Sripada calls an ancillary action. In these cases, *while* he is acting against his better judgment, an agent actionally exercises self-control to put an end to this behavior. So the actional exercise of self-control accompanies, as Sripada says, an action that serves a wayward desire. But I do not take this view about moderate cases of the sort discussed in Sect. 1. I contend that an agent may successfully exercise self-control against what he is most strongly motivated to do and, specifically, against his strongest proximal desire in some cases in which he is not acting—and does not act—on that desire (see Mele 2003, pp. 184–89).

I mention this in passing. My aim now is not to set the record straight about my own view. Instead it is to articulate Sripada's proposed resolution to a puzzle about synchronic self-control and then to assess its merits.

Sripada's "divided mind" approach to a puzzle about synchronic self-control features a pair of "motivational systems"—a "deliberative" and an "emotional" one (2012, p. 11). "Willpower," in his view, is "*exclusively* available to the deliberative system," and exercises of it have the "effect of attenuating, blocking or

[4] Kennett has since come around to the view that there can be actional exercises of synchronic self-control (2001, p. 146).

in some other way modifying the motivational properties of an action-desire produced by the emotional motivation compartment" (p. 13). "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" (p. 13). The notions at work here of positive and negative motivational bases come from Mele (1987). There, I defined the positive motivational base of a desire as "the collection of all occurrent motivations of the agent that make a positive contribution to the motivational strength of that desire" (p. 67) and the negative motivational base of a desire as "the collection of all occurrent motivations of the agent that make a negative contribution to the motivational strength of that desire" (p. 68). Regarding the latter, I wrote: "the attenuating motivation can prevent the strength of the desire in question from reaching a level which it would otherwise have reached, decrease the strength that the desire actually has at a given moment, or both" (p. 68). Part of what Sripada is claiming is that desires outside the deliberative system have no attenuating effect on the strength of desires to exercise willpower (see also Sripada, p. 15).

Sripada distinguishes between two different kinds of desire-involving contests. "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 independent role in deciding the winner" (2012, pp. 17–18). "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" (p. 22).

Recall Sripada's point of departure:

> *S*. 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. (2012, p. 2)

He contends that rather than showing that *S* is false, he shows that "its domain is restricted": in "contests involving the exercise of willpower, [*S*] fails to apply" (p. 22). Sripada also contends that "of the three accounts of willpower" he discusses (Kennett and Smith's, mine, and his), "only the divided mind account makes sense of exercises of *full-blooded* willpower" (p. 26), which exercises he represents as follows:

> Conditions for *full-blooded* willpower:
>
> 1. During some (arbitrarily long) interval *I* extending over *t1*… *tn*, *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 *t2*, 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*). (p. 10)

In Sect. 3, I discuss some interesting evidence Sripada offers for the claim that this set of conditions is satisfied in some cases. A comment on a certain difference between Sripada's view and mine is in order now. As I mentioned, my view allows for non-ancillary actional exercises of synchronic self-control. In my view, when these exercises work, they change the balance of the agent's motivational condition. For example, in the case of an agent whose proximal desire to *A* is stronger than any competing desire, a successful exercise of non-ancillary actional synchronic self-control in the service of his *B*-favoring better judgment results in his becoming more strongly motivated to *B* than to *A* (see Mele 2003, pp. 184–189).

Having said that, I should add that I have expressed my openness to the possibility that some mechanisms or processes linking proximal desires to *A* to attempts to *A* are themselves biased in favor of or against the occurrence of the attempts (Mele 2003, pp. 167–168). So I am open, for example, to the idea that, owing to facts about such mechanisms or processes, a weaker proximal desire may win out over a stronger one. (For principles linking motivational strength to action that accommodate this idea about bias, see Mele 2003, pp. 172, 187.) Accordingly, I am open to the idea that an approach to synchronic self-control in the neighborhood of the one Sripada defends might work. I turn now to an assessment of Sripada's defense of his position on synchronic self-control.

3 Sripada's "divided mind" view versus my "motivational shift" view

Sripada asserts that one advantage of his account of synchronic self-control is that it "makes sense of [an] especially strong kind of self-control to which common sense is committed" (2012, p. 26). Here, some care is needed. Sripada and I agree that there are actional exercises of synchronic self-control, and we agree that an agent may successfully exercise non-ancillary actional synchronic self-control. As I have mentioned, a difference between our accounts is that on mine when exercises of synchronic self-control against what one is most strongly proximally motivated to do are successful, they bring it about that the agent's wayward proximal desire is no longer strongest (call this the *motivational shift thesis*), whereas on Sripada's account there is an important sphere in which actional exercises of synchronic self-control succeed even though the agent continues to be most strongly motivated by a wayward proximal desire. As a convenient bit of shorthand, it can be said that I have defended a *motivational shift* view in the sphere in question. And Sripada rejects this view.

I am open to the idea that Sripada is right about this (more on this later). But the question just now is whether, on this issue, common sense comes down on his side, as he claims. Perhaps it is a pronouncement of common sense that an agent can successfully exercise self-control against his most affectively intense proximal action-desire while that desire remains affectively most intense. But I agree with this. Motivational strength, as Sripada says (2012, p. 18), is supposed to be a matter

of action-causing power—something that should not be confused with affective intensity.[5]

Now, perhaps it is a pronouncement of common sense that an agent can successfully exercise synchronic self-control in the service of a better judgment that conflicts with what he is most strongly motivated to do straightaway in the relevant "action-causing power" sense. But I agree with this too. The question before us is whether common sense supports the idea that we can do this and bring it about that we do not act on proximal motivation *that continues to be motivationally strongest for a significant stretch of time* (even though we are able to act on that motivation) and intentionally act otherwise during that stretch of time. For example, if Mike, who is afraid of dogs, is more strongly motivated to walk away from a bothersome dog right then than he is to do anything else then, can he, by exercising self-control, bring it about that he intentionally continues to sit for, say, 10 minutes, even though, throughout that span of time, he is more strongly motivated to walk away then than he is to do anything else then and is able to walk away (see Sripada 2012, pp. 9–10)? Until I see a powerful argument that common sense does support this idea, I will keep an open mind. Carefully constructed probes using the techniques of experimental philosophy might prove useful. Of course, they will need to direct the experimental participants to the pertinent notion of motivational strength (as opposed to the affective intensity of a desire or of an associated state of mind, for example) and to scenarios of the kind at issue.

Another advantage that Sripada claims for his view over mine is that his, but not mine, accounts for the fact that "willpower [i.e., self-control] is always exercised in the service of, and never against, one's practical judgments" (2012, p. 30). In my view, the alleged fact is actually a falsehood. Consider the following:

> Young Bruce has decided to join some wayward Cub Scouts in breaking into a neighbor's house, even though he… judges it best not to do so… Experiencing some trepidation about the house-breaking, Bruce tries to steel himself for the deed. Although he judges it best not to participate in the crime, he attempts, successfully, to master his fear, and he proceeds to pick the lock. Here, it seems, Bruce has exhibited strength of will; he has exercised self-control in conquering his fear. If that is right, he has done so even if—as we may suppose—he did not judge it best to master the fear that he experienced, nor judge it better to do so than not to do so. Some exercises of self-control apparently are not performed in the service of a better or best judgment. (Mele 1995, p. 60; see Mele 1987, p. 54)

I call exercises of self-control of the kind described in this story *unorthodox* exercises, because the conduct they support in fact clashes with the agent's better judgment. Now, one may stipulate senses of "self-control" and "willpower" according to which these things can be exercised only "in the service of, and never

[5] Here is an illustration from Mele (2003, p. 162). "Carol, a morally upright psychiatrist who experiences, to her own consternation, a phenomenologically intense urge for a sexual romp with a seductive patient, may have, in a desire that has little phenomenological kick, a stronger inclination to forego that course of action."

against, one's practical judgments" (Sripada 2012, p. 30). But the stipulated senses are narrower than what ordinary usage of the terms permits (for evidence of that, see Mele 2012, pp. 95–96).

There is a significantly deeper danger for Sripada's view here than failing to respect ordinary usage. As I have mentioned, he locates willpower exclusively in "the deliberative system" (2012, p. 13). That feature of his view is supposed to account for the alleged fact at issue (p. 30). But if, as it seems, Bruce may exercise willpower in conquering his fear and committing the crime against his better judgment, Sripada has drawn the boundaries around willpower too narrowly. Furthermore, if exercises of willpower always involve the employment of "regulatory powers" (Sripada 2012, p. 22), then it seems that these powers are not housed exclusively in—or accessible exclusively to—"the deliberative system," and one should begin to wonder whether any *actual* systems of the kind Sripada works hard to describe are as separate from one another as he represents his "deliberative" and "emotional" motivational systems as being.[6]

Recall Sripada's claim that desires outside the deliberative system have no attenuating effect on the strength of desires to exercise willpower (that is, desires of the former kind are never in the negative motivational base of desires of the latter kind). He claims that "this restriction… holds the key for explaining how willpower, and in particular *full-blooded willpower*, is possible" (2012, p. 15). Is this restriction plausible?

Some readers may think that no desires have negative motivational bases. In their view, for example, the strength of a smoker's proximal desire to light up upon exiting a building cannot be lessened by the desire not to offend that arises in him when someone calls his attention to a "no-smoking" sign. But Sripada does not share this view. (Nor do I.) Consider another smoker who has judged it best to quit and now, at a particularly stressful time, is tempted to smoke. If the former smoker's desire not to offend can have the effect that his proximal desire to smoke is weaker than it would otherwise be, why should we believe that the latter smoker's desire to smoke cannot have a similar effect on his desire not to smoke, thereby rendering the motivational base of any desire he may have to exercise willpower now in support of his better judgment weaker than it would otherwise be? Here again, the plausibility of the idea that the two systems are as separate from one other as Sripada claims is called into question.

As I have mentioned, I defended a "motivational shift" view about successful exercises of synchronic self-control, and Sripada rejects this view. According to a motivational shift view, in cases in which one successfully exercises synchronic self-control against one's strongest proximal motivation one's exercise of self-control shifts the balance of one's proximal motivation so that it favors the course of action in the service of which one exercised self-control. As far as I can see, if Sripada is right to reject this view, his being right does not depend on either of the following claims of his:

[6] If regulatory powers are accessible to "the deliberative system" but not exclusively so, are they also accessible to "the emotional system"? Not necessarily. They may be accessible to the person and usable by that person in unorthodox and orthodox exercises of self-control, even if they are not accessible to the person's emotional system itself.

(*S1*) "willpower [i.e., self-control] is always exercised in the service of, and never against, one's practical judgments" (2012, p. 30); (*S2*) desires outside the deliberative system are never in the negative motivational base of desires to exercise willpower. Both claims are expressions of the separateness Sripada sees between what the calls the "deliberative" and "emotional" motivational systems.

One might think that Sripada needs *S1* to help underwrite the idea that exercises of synchronic self-control can prevent one's strongest proximal motivation from being effective even though the wayward proximal motivation continues to be strongest for quite some time. In this connection, Sripada distinguishes between "the set of causal powers possessed by desires that explain how desires have their characteristic effects on action" ("D-powers," 2012, p. 18) and "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" ("R-powers," p. 19). A "weaker desire" might engage regulatory systems, and "so long as the regulatory powers of these systems are sufficient, then the stronger desire can be defeated" (Sripada 2012, p. 22). One might worry that if "regulatory systems" and their R-powers are accessible not only to the deliberative system but also to the emotional system, then, if anything engages them, it will always be the agent's strongest desires that do.

Can a theorist who is convinced that there are unorthodox exercises of self-control of the kind described in my story about young Bruce find a way to handle this worry? Elsewhere, I distinguished between two kinds of commitment to action, *evaluative* and *executive* (Mele 1995, pp. 71–74). Prime examples of evaluative commitments are deliberative judgments made by agents to the effect that, all things considered, their *A*-ing would be best. Executive commitments are constituted by agents' intentions and decisions to *A*, whether or not those intentions and decisions are aligned with any relevant evaluative judgments. Both kinds of commitment are distinguished from desires, including desires to do things. The worry at issue can be handled in a way that accommodates the existence of unorthodox exercises of self-control by defending the thesis that only agential commitments—whether evaluative or executive—can engage regulative systems and activate their R-powers. Clearly, if there are unorthodox exercises of self-control, this thesis is more plausible than *S1*.[7]

What about *S2*, Sripada's claim that desires outside the deliberative system are never in the negative motivational base of desires to exercise willpower? If, as Sripada claims, desires in the deliberative system that are weaker than competing emotional desires can engage regulatory systems, then even some deliberative-system desires that have emotional desires in their negative motivational base may be powerful enough to engage regulatory systems. And if the deliberative system has access to R-powers and such powers can be put to use in defeating "the stronger desire" (2012, p. 22), it would seem that even some deliberative-system desires with emotional desires in their negative motivational base can benefit from this fact.

[7] Sripada asserts that even if the details of his "account of the divided motivational structure of the mind turn out to be… *entirely* wrong, the more general point that motivational division is required for full-blooded willpower should nonetheless survive" (2012, p. 11). He can count the difference between agential commitments and desires as a motivational division.

I have argued that even if some elements of Sripada's alternative to my "motivational shift" view of synchronic self-control are implausible, they are not required as underpinnings for his very interesting thesis that there are actual cases of actional synchronic self-control in which agents bring it about that they do not act on proximal motivation *that continues to be motivationally strongest for a significant stretch of time* (even though they are able to act on that motivation) and intentionally act otherwise during that stretch of time. I dub this thesis *Block*, and I turn to the evidence Sripada offers for it.

The evidence comes from exposure therapy (Sripada 2012, pp. 28–30). Exposure therapy for a person, Paul, with a severe fear of heights involves confronting him with "fear-inducing height-related stimuli" (Sripada, p. 28). Paul is instructed to "suppress his natural tendency to *withdraw* from the fear-inducing stimuli and instead *approach* them." For example, "after a dozen sessions," he may be placed on the top of a tall building, attached to a harness, and instructed to climb down the fire escape all the way to the ground (Sripada, p. 28).

Sripada asserts that "several lines of evidence support the claim that throughout each exposure session, Paul's strongest desire is to withdraw from the height-related stimuli" (2012, p. 29). The claim in the case at hand is that throughout his climb down the fire escape, Paul is most strongly motivated not to climb down, and his strongest proximal motivation during each step down (perhaps until climbing up would be more frightening that climbing down?) is not to move downward.

The first piece of evidence Sripada offers is Paul's long history of height avoidance (2012, p. 29). This history is treated as evidence that, despite the treatment Paul received in the preceding twelve sessions, he still has a terrible fear of heights and—more to the point—as evidence that the fear is so strong as to render Paul more strongly motivated not to climb down the fire escape than to climb down it, even as he is climbing down.

Sripada's second piece of evidence concerns reports patients make of their level of distress during their sessions. A "Subjective Units of Distress Scale" (SUDS) is used. Sripada writes: "Paul reports his level of subjective distress throughout each session, and it is not uncommon for his SUDS score to rise above 80 (on a scale from 0 to 100). This means that Paul is approaching the worst fear he can possibly imagine" (2012, p. 29).

Sripada's third and final point in this connection highlights the extreme nature of some exposure scenarios. He offers the example of "a person who cannot even touch a doorknob" eventually being asked to "place his hands in a toilet bowl for an extended period of time" (2012, p. 29). Paul's climbing down the fire escape may be comparably extreme. Sripada writes: "It is plausible that these extreme scenarios generate intense desires to withdraw that are indeed the agent's strongest."

How powerful is this collection of evidence? I start with Paul's history. Evidently, he made a lot of progress in the preceding twelve sessions. Before his therapy started, "Paul passed up his dream position with a consulting firm because their penthouse office suite could only be reached with a glass elevator" (Sripada 2012, p. 28). At that time, he refused to use a completely enclosed elevator with glass walls, thereby cutting himself off from getting his dream job. And now, here he is, with only a harness to protect him, stepping off the top of a twenty storey

building and climbing down a fire escape to complete a therapy assignment. Obviously, there has been a major change in Paul. Why should we not believe that the change is such that now, by exercising self-control, Paul can and does bring it about that, as he begins climbing down the fire escape, and as he climbs down, he is more strongly motivated to do that than he is to do otherwise? (A lot of progress can be made in a dozen sessions of exposure therapy. For an interesting discussion of the effects on acrophobes of just three one-hour sessions of *virtual reality* exposure therapy, see Emmelkamp et al. 2002.)[8]

I am not suggesting that a conceptual truth yields the verdict that Paul is more strongly motivated to do what he does than to behave otherwise. As I mentioned, I am open to the possibility that some mechanisms or processes linking proximal desires to *A* to attempts to *A* are themselves biased in favor of or against the occurrence of the attempts in such a way that, owing to facts about such mechanisms or processes, a weaker proximal desire may win out over a stronger one. The present discussion is about where the evidence points. And there is more evidence to consider.

I turn to the SUDS reports. I begin with a relatively minor point. Sripada's representation of what they signify is exaggerated. He says that a score of above 80 "means that Paul is approaching the worst fear he can possibly imagine" (2012, p. 29). Well, what about climbing down the fire escape *without* a harness, walking from the top of one twenty storey building to the top of another one across the street on a tightrope with a harness, doing the same thing without a harness, jumping off a sixty storey building using a hang glider or bungee cord, and so on? Unless Paul's imagination is severely stunted, he can imagine a great many worse fears that more closely approach "the worst fear he can possibly imagine."

One route to a more important point begins with the observation that patients who are treated with exposure therapy "learn about the sympathetic nervous system's flight or fight response to fear-eliciting stimuli (actual or perceived), and how not to misinterpret these physical signs of arousal as indications of actual danger in anxiety-provoking, but safe situations" (Back et al. 2001, p. 39). So we should not be surprised if the intensity of their feelings of "distress" about *A*-ing is not matched by the strength of their motivation not to *A*. A distress scale should not be confused with a motivational strength scale. This is not to say that Sripada confuses the two. Rather, he uses SUDS reports as evidence about the strength of avoidance motivation.

The competing hypotheses at issue now are my hypothesis (*H1*) that Paul's exercises of synchronic self-control during the episode at issue affect the balance of his competing proximal motivations in such a way that it favors his starting and continuing to climb down the fire escape and Sripada's hypothesis (*H2*) that Paul successfully exercises self-control even though he is most strongly motivated not to climb down the fire escape, and even though his strongest proximal motivation during each step down (perhaps until climbing up would be more frightening that climbing down?) is not to climb down. For the purpose of comparing the two hypotheses, pick a specific point during the downward climb—say, just after Paul has taken his tenth step down—and suppose that Paul is still exercising self-control.

[8] The article also discusses *in vivo* exposure therapy for acrophobes, the sort of therapy Paul undergoes.

H1 is consistent with the truth of the statement (1) that Paul, who is exercising self-control in such a way that he is just a bit more strongly motivated to continue climbing down than not to do so, takes another step down. And *H2* is consistent with the truth of the statement (2) that Paul, who is exercising self-control, takes another step down despite being just a bit more strongly motivated not to continue his downward climb than to continue it. I do not see how Paul's SUDS reports put us in a position to be confident that statement 1 is false and that if either of the two statements is true, it is 2. Just for the purposes of illustration, imagine a pair of disputants who believe that motivational strength can be precisely measured. The proponent of *H1* says that the strength of Paul's proximal motivation to take a step down is 51 whereas that of his proximal motivation not to take a step down is 49, and the proponent of *H2* makes the opposite assignment of strengths. Given that we have no reason to believe that SUDS scores are particularly reliable guides to motivational strength—especially in people who have been taught "how not to misinterpret… physical signs of arousal as indications of actual danger in anxiety-provoking, but safe situations" (Back et al. 2001, p. 39)—Paul's SUDS reports are at best feeble evidence in favor of 2.

Adding in Paul's history cuts no ice here, for reasons offered above. What about Sripada's claim, quoted above, that "It is plausible that these extreme scenarios generate intense desires to withdraw that are indeed the agent's strongest." The affective intensity of a desire is one thing and the motivational strength of a desire is another. I do not see how the extremeness of a scenario gets us any further than do the SUDS reports we would get in such scenarios. And bear in mind that by the time patients behave as instructed in extreme scenarios, they have been through exposure therapy that, as in Paul's case, has wrought significant changes. The jury should still be out, then on whether *H2* is more plausible than *H1* and on whether Sripada's *Block* is more plausible than my motivational shift view.

The search should continue for evidence of the existence of processes and other phenomena that would underwrite *Block*. Phenomena of the sort at issue would be involved in the production of intentional actions performed over a several minute (or longer) span that are contrary to what the agent is, throughout that span, most strongly motivated to do during that very span of time but is not doing or even trying to do, although he is able to do it. I have not been persuaded that Sripada has produced such evidence. But I am not denying that such evidence can be produced. The project of finding evidence of the kind at issue is an important one. If successful, it would shed new light on the springs of intentional action.

I close with some guidance on the evidential project at issue. Randolph Clarke asserts that "There are apparently possible cases where we could have very good evidence that neurological conditions causally affect which action is performed and yet no change whatsoever in the balance of motivational strength has occurred" (1994, p. 8).[9] Obviously, *intentional* actions are at issue; and if the cases Clarke has

[9] Here Clarke has in mind biasing mechanisms or processes of a kind I mentioned above. As I said, I have no objection to the claim that such things are possible. But my topic now is evidence.

in mind are to be directly relevant to present concerns, *proximal* motivation is at issue as well. Clarke writes:

> Suppose that Abe judges doing A better than doing B, but that all indicators of motivational strength point to his wanting more to B than he wants to A. Suppose that Abe now undergoes [various] neurological changes… Then he decides to A rather than B, and he A's rather than B's; the neurological changes are causally relevant to his decision and behavior, and all indicators of motivational strength except the resultant action point to his wanting more strongly to B. (He feels just as tempted as ever to B, he daydreams about it as often and as vividly, he reports no change in the relative strengths of the desires in question, etc.) Here, we would have excellent if defeasible evidence that the relative strengths of the agent's desires have not changed. (p. 8)

He adds: "In the case of Abe… the balance of motivational strength ought to be able to explain his continued feelings of temptation, his self-reports, and his daydreaming" (p. 8).[10]

The evidence that Clarke mentions here for the claim that, at the time of action, Abe was more strongly motivated to *B* then than to *A* then seems primarily to be post-action evidence.[11] Clarke's story about Abe does not portray him as daydreaming about *B*-ing while he *A*-s. Seemingly, it is *after* he *A*-s that Abe daydreams about *B*-ing "as often and as vividly" as he used to. He is tempted to *B* while he *A*-s—and afterward, as well, it seems. And it is not part of Clarke's story that Abe reports on his desires while he is *A*-ing. So perhaps, after he *A*-s, Abe reports that there is no change in the relative strengths of certain general desires (desires for *A*-type action and *B*-type action). But how much does the post-action daydreaming, post-action temptation, and post-action reporting tell us about the strengths of Abe's competing proximal motivations *at the time of action*? And does it, together with the temptation Abe feels at the time of action, provide stronger evidence about what Abe was most strongly motivated to do at that time than his intentional *A*-ing does? Perhaps, on the basis of evidence of the kind Clarke imagines, we would consider predicting that the next time Abe is faced with a choice between *A*-ing and *B*-ing, he will *B*. But we should look for evidence about how good Abe is at exercising diachronic and synchronic self-control (especially regarding competing *A*-type and *B*-type actions) before we make any such predictions.

One moral here is that when it comes to theses like Sripada's *Block* and my motivational shift thesis (and *KS*, *T*, and *T**, for that matter), we need to focus on *proximal* motivation and the best evidence we can get about it. In the case of ordinary intended intentional actions (as opposed, for example, to intentional side-

[10] Notice that a strong desire to *B* can help explain a feeling of temptation and daydreaming, even if the agent has a stronger competing desire. The same is true of self-reports of the kind at issue, unless we assume that people are infallible judges of the relative strengths of their desires (an assumption to which I return).

[11] Readers will have noticed "etc." in the parenthetical sentence in which Clarke mentions temptation, daydreaming, and reporting. Earlier, he also mentioned the following indicators of motivational strength: a person's "past behavior in similar circumstances," what a person talks about, the content of a person's thoughts, and being disappointed when one learns one cannot do something (1994, p. 4).

effect actions, if there are any), as Clarke points out (1994, p. 8), the fact that the agent *A*-ed is some evidence that he was most strongly motivated at the time to *A*. Do we ever have better contrary evidence? Although I have not yet been persuaded that we do, I believe that a careful search for it would prove illuminating.

To make Abe's story more applicable to Sripada's *Block* and my motivational shift thesis, one might imagine a version of the story in which Abe is daydreaming about *B*-ing *while* he *A*-s and, *while A*-ing, reports that he is more strongly motivated to *B* then than to *A* then. Perhaps Abe passed up an opportunity to ride with friends to the beach in order to do some yard work at home. As he works, he daydreams about being at the beach with his friends. When a neighbor drops by, Abe reports that he is tempted to quit working and drive to the beach to join his friends. He says, as well, that at that very moment he is more strongly motivated to quit working right then and to leave for the beach than he is to continue working, and he adds that the whole time he was working he was, in fact, more strongly motivated to go to the beach instead. Abe keeps working as he chats with his neighbor and continues working on his yard for another hour or so. He then spends the rest of the day relaxing at home, occasionally daydreaming about being at the beach.

How powerful is the imagined evidence in this revised case for the claim that Abe acted contrary to his strongest proximal motivation? If Abe is infallible about the relative strengths of his competing motivations, his sincere report alone would settle the matter, of course. But how good are we at detecting the relative strengths of our proximal competing motivations? And *how* do we detect this? Do we use felt intensity as our primary indicator?

There is a lot of much-discussed evidence that we do not know ourselves nearly as well as we think we do (Wegner 2002; Wilson 2002). But I will not make much of that evidence here. Instead, I observe that we do not have direct access to the causal power of our desires. We base any judgments we may make about that matter and about the relative causal power of our competing motivations on evidence that is more directly accessible. So is whatever evidence Abe uses as a basis for his report to his neighbor—together with whatever other evidence there may be for the reported proposition—better evidence about relative strength than is his doing yard work rather than going to the beach? And how good is Abe at interpreting the evidence on which he bases his judgment? Until we know what evidence Abe uses, these questions will not be easy to answer. I am not insisting that Abe is wrong. But the imagined evidence in Abe's story (including the felt temptation, daydreaming, reporting, working, and so on) falls well short of warranting confidence that he is right. There is some recent scientific research in which participants report on "desire strength" (Hofmann et al. 2012a, b).[12] Here are two interesting findings: When people made no effort to resist desires to which they gave a strength rating of

[12] In this literature, "desire" has a narrower extension than it does in the philosophical literature on motivational strength that I have been discussing. For example, Hofmann et al. (2012a) (following Kavanagh et al. 2005) define it as an "affectively charged cognitive event in which an object or activity that is associated with pleasure or relief of discomfort is in focal attention" (p. 1319), and Hofmann et al. (2012b) distinguish desires from motivation-encompassing attitudes for "bodily fitness," healthy eating and drinking, "saving money," sexual fidelity, and so on (p. 584).

"irresistible," they acted on only 71 percent of these desires; and when people made an effort to resist desires they rated as irresistible, they successfully resisted 74 percent of them (Hofmann et al. 2012a, p. 1330). This does not inspire confidence in people's judgments about the strengths of their desires. In any case, the search for good evidence that the kind of claim Abe is making sometimes is true would be a worthwhile project.

Acknowledgments For comments on a draft of this article, I am grateful to Randy Clarke and Chandra Sripada. This article was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this article are my own and do not necessarily reflect the views of the John Templeton Foundation.

References

Back, S., Dansky, B., Carroll, K., Foa, E., & Brady, K. (2001). Exposure therapy in the treatment of PTSD among cocaine-dependent individuals: Description of procedures. *Journal of Substance Abuse Treatment, 21*, 35–45.

Clarke, R. (1994). Doing what one wants less: A reappraisal of the law of desire. *Pacific Philosophical Quarterly, 75*, 1–10.

Emmelkamp, P., Krijn, M., Hulsbosch, A., de Vries, S., Schuemie, M., & van der Mast, C. (2002). Virtual reality treatment versus exposure in vivo: A comparative evaluation in acrophobia. *Behaviour Research and Therapy, 40*, 509–516.

Hofmann, W., Baumeister, R., Förster, G., & Vohs, K. (2012a). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*, 1318–1335.

Hofmann, W., Vohs, K., & Baumeister, R. (2012b). What people desire, feel conflicted about, and try to resist in everyday life. *Psychological Science, 23*, 582–588.

Kavanagh, D., Andrade, J., & May, J. (2005). Imaginary relish and exquisite torture: The elaborated intrusion theory of desire. *Psychological Review, 112*, 446–467.

Kennett, J. (2001). *Agency and responsibility*. Oxford: Clarendon Press.

Kennett, J., & Smith, M. (1996). Frog and toad lose control. *Analysis, 56*, 63–73.

Kennett, J., & Smith, M. (1997). Synchronic self-control is always non-actional. *Analysis, 57*, 123–131.

Mele, (1987). *Irrationality*. New York: Oxford University Press.

Mele, (1992). *Springs of action*. New York: Oxford University Press.

Mele, (1995). *Autonomous agents*. New York: Oxford University Press.

Mele, (1997). Strength of motivation and being in control: Learning from Libet. *American Philosophical Quarterly, 34*, 319–332.

Mele, (1998). Synchronic self-control revisited: Frog and toad shape up. *Analysis, 58*, 305–310.

Mele, (2003). *Motivation and agency*. New York: Oxford University Press.

Mele, (2009). *Effective intentions*. New York: Oxford University Press.

Mele, (2012). *Backsliding*. New York: Oxford University Press.

Sripada, C. (2012). How is willpower possible? The puzzle of synchronic self-control and the divided mind. *Noûs,* 1–38. doi:10.1111/j.1468-0068.2012.00870.x.

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

Wilson, T. (2002). *Strangers to ourselves*. Cambridge, MA: Harvard University Press.