

## Self-control: a mental action in no need of special motivational powers

Accepted for Brent, M. (ed.), *Mental Action and the Conscious Mind*, Routledge (forthcoming)

Sebastian Watzl

### 1. Introduction

The image of the agent who controls her passions like a charioteer steering her hot-blooded horses, has long sparked the philosophical imagination.<sup>1</sup> Such self-control illustrates, it has been argued, something deep and interesting about the mind. Specifically, it has been suggested that for its explanation we must posit special motivational powers: *will-power* as an irreducible mental faculty (Holton 2009), the active self as a dedicated and depletable pool of *psychic energy* or – in today's more respectable terminology – mental resources (Baumeister, Bratslavsky and Muraven 2018), or a deep division between *reason and passion* – a deliberative and an emotional motivational system (Sripada 2014).

This essay argues that no such special motivational powers are necessary. Yet, at the same time, the tradition is right that self-control powerfully illustrates the importance of a feature of the mind. What it illustrates, I argue, is the importance of the mental activity of attention in the control of *all* action. It is by appeal to this mental activity that we can dispense with special motivational powers.

The significance of attention for self-control, of course, is compatible with several models, including ones I would like to reject.

On the one hand, one might link attention to a willpower or a resource driven view of self-control, starting perhaps from William James (1890, p. 562) who thought that "effort of attention is ... the essential phenomenon of will". The idea might be that attention either is itself a mental resource, a willpower faculty, or the mechanism controlling access to them.

On the other hand, and in stark opposition to the first model, attention has also been linked to the denial that the will is involved in self-control at all, as what dissolves the charioteer altogether (Ganeri 2018). Recently, attention has been suggested as central to a surprisingly

---

<sup>1</sup> Plato, *Phaedrus*, 246a–254e. See Ganeri (2018) on the imagery in Upaniṣads written around the same time. As Ganeri points out, the Buddhists empathetically denied the adequacy of the image.

non-agential view where all strategies for self-control are “distinctively cognitive” (Kennett 2003, p. 139, cf. Kennett and Smith 1996, 1997).

The role of attention I will argue for is opposed to both of these conceptions. The capacity for attention is important for self-control exactly because it is agential, and it is important for self-control even though it has no connection to anything resembling special motivational powers. The interesting feature of the mind self-control illustrates is this: attention acts as a flexible interface between the agent's motivational systems and her actions. Through attention an agent can actively couple or decouple an intention, preference or desire to and from action – by intentionally changing the current priority of her mental states. I call the resulting view *the re-prioritization account* of synchronic self-control.

In one sense, the project of this paper is deflationary: self-control uses no special motivational capacities. Agents use attention for action control also when things go smoothly and when there is no need for controlling any wayward temptation. Once we think correctly about the role of attention in the control of all forms of agency (see Wu 2016, Watzl 2017), there is no explanatory role for willpower, mental resources, or a divided mind. Self-control is not special: if we think of Humeanism as the view that there is fundamentally only one kind of motivational system and that all action is based in that system, then this essay contributes to a defense of Humeanism.

In another sense, the project – though deflationary – is constructive. Any model of agency in terms of only beliefs and desires, motivational and representational states, or preferences and credences, is incomplete. Models of agency need attention as an independent factor. The way attention organizes the mind cannot be subsumed under what the agent wants and how she takes the world to be.<sup>2</sup> A different conception of Humeanism as the view that every mental state is either motivational, representational, or a combination of them, is false.

The view I argue for in this essay aligns with one defended by Inzlicht and Berkman and colleagues in the psychological literature: according to them “the decisions that we label self-control are merely a fuzzy subset of all value-based decisions, which involve selecting a course of action among several alternatives.” (Berkman et al. (2017, p. 423)). I also agree

---

<sup>2</sup> Sinhababu (2017) argues *for* Humeanism and against, for example, willpower (Ch. 8) and a substantial role of the self (Ch. 10), by appealing centrally to the role of attention in the link between desire and action (he suggests also that certain patterns of attention are constitutive of desire; cf. Ch. 1 & Ch. 5). Yet, if attention cannot be successfully reduced to a combination of representational (beliefs) and motivational states (desires) (cf. Watzl 2017), and if attention plays an irreducible role in the (rational) explanation of action (which I argue for below, but see also Wu 2014, 2016), then Humeanism is false in an important sense.

with them that the field "would be better served by abandoning the resource concept altogether" (Inzlicht and Berkman (2015, p. 520)), and that "self-control outcomes emerge organically from the operation of a single, integrative system with input from multiple regions rather than antagonistic competition between two processes" (Berkman et al. (2017, p. 424)). I point readers with an interest in the empirical details to this literature. My goal here is in a corresponding philosophical defence.

My primary aim is descriptive: to show how self-control can work without willpower, mental resources, or mental division. In the conclusion, I will briefly touch also on normative questions: whether self-control can be rational or irrational, and whether exercising self-control is always good for the agent. If my descriptive model is right, then, probably there isn't much in general to say about the normativity of self-control. The descriptively deflationary view may help to show that also in the normative sense we should neither take an overly elevated nor a debased view of self-control.<sup>3</sup>

Here is how I will proceed. In section 2, I will introduce what self-control is, and focus on one variant that philosophers have found especially puzzling. In section 3, I explain one such philosophical puzzle, aiming to show that this variant, intentional synchronic self-control, is impossible without positing special motivational powers. A satisfactory theory of self-control should respond to this puzzle. In section 4, I will first sketch my re-prioritization view of self-control and show why a non-intentional account of synchronic self-control, which shares some its ideas, is unsatisfactory. In section 5, I present the view in a bit more detail and show how it can be used to defend the possibility of intentional synchronic self-control on the basis of four premises. Sections 6, 7, 8, and 9 then defend each of these premises. Section 10 answers how the re-prioritization view can explain the sense of effort and difficulty accompanying self-control attempts, and why self-control may improve through training. In the concluding section 11, I summarize and point briefly to potential normative implications.

## 2. What is self-control?

Agents exercise self-control to counteract a threat of losing control over themselves. Here is a paradigmatic example:

*Cookie Temptation.* After an exhausting day at work, on her walk home, Christina walks by a pastry store. Delicious cookies are on display. Christina stops and looks into

---

<sup>3</sup> Cf. Brownstein (2018).

the window. The temptation rises: she wants one. Eat it. Right here. Right now. But Christina has resolved to become more fit and healthy, and made the specific plan to go on a run first and then eat a salad. She firmly believes that this is what she should do. Part of her resolution was specifically to overcome temptations like the one she is now facing.

When Christina is standing in front of the pastry shop, she experiences a threat of losing control over herself. She feels that her momentary urge to eat that cookie is getting the better of her. But this is not what she thinks she should do. In a situation like this she may exercise self-control: she may pull herself together, and go on the run she had planned.

Here is another paradigmatic case:

*Angry Punch.* Amira is driving home from a dinner with friends. Suddenly, a police car comes up behind him, and pulls him over. When they ask him for his license, he hears in their voice that they treat him as some kind of suspect. Then they interrogate him about whether he had taken drugs, ask him to step out the car, and begin searching every corner inside. Amira feels the anger rising. He feels insulted by their demeanour. He feels like punching one of the police men in the face. But he knows that this would have disastrous consequences. He would probably end up in prison or worse.<sup>4,5</sup>

For Amira, like for Christina, there is a threat of losing control over himself. He believes that no good is going to come of that punch. He needs self-control to do as he thinks he should do: let the insulting procedure pass, get back into the car once it is over, and be home in less than ten minutes.

Starting from such paradigmatic examples, we can characterize self-control in terms of two features (cf. Duckworth, Gendler, and Gross 2016):

The first feature is a certain type of situation, a *loss of control threat*, as I will call it. Loss of control threats are characterized by an asymmetric and subjective conflict. The agent is faced with at least two options. One of the options is, from her own perspective, *better*. The other one, though, has a more *powerful grip* on her. In the cookie temptation case, the better option is to go on a run and eat a salad. The option with the more powerful grip is to eat a cookie

---

<sup>4</sup> Thanks to Johannes Rössler, who pushed me to say more about how my view treats cases like the Angry Punch.

<sup>5</sup> It has not escaped my notice that both of my examples have problematic features. I will return to the issue of normativity and what is good for the agent in the conclusion.

right there on the spot. The conflict in the situation is subjective in three ways. First, the agent is aware of both options. Second, it is from the agent's own perspective that the two options are incompatible. Third, the better option is better by the agent's own lights. Loss of control threats should then be distinguished from situations where the agent is not aware of an important option, ones where the agent does not realize that there is a conflict between what has a powerful grip on her and what she takes as the better option (a form of mental fragmentation), cases where the agent is stuck between two equally desired options (Buridan's ass situations), and cases where one option is better for the agent even though the agent herself does not see it that way.

The second feature of self-control is conflict resolution: the agent counteracts the threatening loss of control or – in other words – she resolves the subjective conflict in favour of the better option. Christina would not have exercised self-control if she went home for her run because the cookie shop is closed, or because her cruel partner does not allow her to eat the cookie. For a case of self-control, the resolution of the conflict must be causally traced to the agent. This, though, is not sufficient: Christina would also not have exercised self-control if she did not eat the cookie because she suddenly feels a stomach ache, or if she happens to remember her overdue taxes whose urgency crowds out even the cookie. Self-control must be agential, in that the conflict is resolved by something the agent herself does non-accidentally.<sup>6</sup>

I will speak of a *self-control event* when an agent, in the non-accidental sense just described, averts a loss of control threat. A *self-control strategy* is a way of averting loss of control threats. And an agent has the *capacity for self-control* insofar and to the degree to which she is competent at averting such loss of control threats.<sup>7</sup>

Given this characterization of self-control, we can distinguish types of self-control by varying self-control cases along various dimensions.

First, there are different ways of understanding when an option is *better* from the agent's own perspective: it is the one that is more congruent with the agent's more long-term goals; more congruent with the agent's overall motivational set; the one that the agent more closely

---

<sup>6</sup> The non-agential view of self-control (cf. Kennett op. cit.), in my view, need not deny this datum. In Section 4 I argue that this view fails, but not necessarily because it cannot explain that self-control is non-accidental.

<sup>7</sup> The capacity for self-control is at issue when one argues that infants or children have less self-control than adults, or that self-control develops throughout childhood, puberty, and early adulthood. Tangeney, Baumeister, and Boone (2014) have developed a "self-control scale" aimed at measuring capacity self-control. For some problems with such a scale see, e.g., Brownstein (2018). The deflationary view developed in this paper may be used to further cast doubt on the use of such psychological measures.

identifies with; the one that the agent would choose if she were fully rational; the one that she has formed a resolution or intention to pursue; or, finally, we may think of an option as subjectively better because it is the one the agent believes or judges to be the option she should pursue. I will here focus on options that are subjectively better in this last sense.

Second, there are also different ways of understanding when an option has a *more powerful grip* on the agent: it may, for example, be the one that promises the more immediate reward, or the one that is phenomenally more powerful. My discussion will focus on options whose powerful grip consists in the fact that the agent in the subjective conflict situation prefers the option over the one she judges to be the one she should pursue (we may also say that she desires the option more).

Third, there are different types of self-control strategies. On the one hand, we have diachronic strategies, which resolve the subjective conflict before it arises. The agent may, for example, eat a banana before leaving the office knowing that this will help her not succumb to eating cookies on her way home. On the other hand, we have synchronic strategies, which resolve the conflict once it has arisen. A second distinction is between situational strategies, which are pursued by selecting or modifying situations so as to minimize or resolve potential conflict – e.g. by walking home on a route that avoids the pastry store (cf. Duckworth, Gendler, and Gross 2016), and intra-psychic strategies, which are pursued by changing one's own mind. I will focus on synchronic and intra-psychic strategies.

In what follows, I will thus understand self-control events as cases where an agent non-accidentally changes her own mind in such a way as to bring her actions in line with her judgment about what she should do when faced – at that time – with an opposing preference. The relevant self-control strategies are ways of changing her mind in such a way. The relevant capacity for self-control is an agent's competence in achieving such a change of mind. The threat of losing control here thus consists in a threat of *akrasia*, i.e. acting against one's best judgment. By engaging in self-control, the agent faces that threat and does what she thinks she should do. In a slogan: self-control here is the enkratic aversion of a threat of *akrasia*.

It's easy to see why this type of self-control might suggest that agents have special motivational powers. If the horses are the agent's preferences or desires, which pull her in various directions, then the enkratic aversion of *akrasia* seems to show that the agent herself – as *opposed to* her preferences – can get the horses to do what *she thinks is right*. But then she must have some special power – her self, her will, her reason – to align her preferences with her beliefs. There must be more than horses.

My own view of self-control is that, as far as motivation is concerned, there is no more than horses. No variety of self-control is resolvable only through special mental capacities. I choose to focus on the enkratic aversion of akrasia because this is the version that seems to raise a philosophical puzzle.

### 3. The Preference Determines Action Principle and the Paradox of Self-control

The puzzle is that this type of self-control, while intuitively quite common, can appear to be impossible given a relatively plausible assumption of how the mind works. This so-called *paradox of self-control* (cf. Mele 1992, Kennett and Smith 1997, Sripada 2014) has inspired special motivational powers as well as the non-agential view of self-control mentioned in the introduction.

Consider the following plausible principle:

*Preference Determines Action.* If, at time  $t$ , the agent prefers option  $A$  over option  $B$ , and believes at  $t$  that both  $A$  and  $B$  are genuine options for her, then the agent will take option  $A$  (and not  $B$ ) – given that she chooses one of them intentionally.

The paradox can now be presented in the form of the following dilemma (Sripada 2014). I will illustrate it with the Cookie Temptation case:

Horn 1: Suppose that when in front of the pastry store, Christina indeed prefers eating the cookie to going home and preparing for her run. It follows from the fact that we have a subjective conflict situation that she recognizes eating the cookie as one of her genuine options. But then it follows from the preference determines action principle that if she chooses one of eating the cookie or preparing for her run intentionally, she will choose eating the cookie. But this then seems to rule out that Christina will engage in another activity, self-control, that will make her *not* eat it.

Horn 2: But then suppose, by contrast, that Christina's tempting urge does not reflect a genuine preference at that time for the cookie over the run. Or suppose that Christina doesn't actually take herself to be free to eat the cookie, i.e. that she doesn't recognize it as a genuine option for her at that time. Then, self-control would seem to be superfluous: if she doesn't actually want the cookie (or prefer it over the run), or thinks that she cannot actually take that option, then she doesn't need to exercise self-control to prevent her from eating it.

The paradox is that the exercise of synchronic self-control can seem either impossible or superfluous. A philosophically satisfactory account of self-control needs to show how akratic aversions of threats of akrasia are possible and efficient at the same time. If the agent indeed, as we have stipulated, prefers a course of action that she believes is open to her, then how could she intentionally do something else that is in direct conflict with what she prefers? In the chariot image: if preference determines action, then the only thing that can change the direction of the chariot are the horses. But then if the horses pull left, it just isn't possible that the chariot ends up going right – unless, of course, it happens accidentally (a gush of wind blows against the horses) or there is a charioteer with special steering powers.

I said that the preference determines action principle is plausible. What exactly is its status?

On one construal, the principle is analytic, following from the definition of "preference". According to some revealed preference theories (cf. Samuelson 1938), choice behaviour determines preferences: from the fact that an agent takes option B over option A, in a situation where both options are subjectively available to her, we can infer that she prefers B over A. If this were right, then we know that Christina, when she "succeeds in self-control" and pursues Run (B) over Cookie (A), must have preferred B over A, and hence our original description in terms of a preference of A over B must have been wrong.<sup>8</sup>

Given appropriate consistency constraints, choice behaviour can without doubt be used to define a preference ordering for the agent (as the so-called revelation theorems show). But the fact that we can *define* such an ordering doesn't show that this ordering is explanatory or psychologically real: preferences should be construed as mental states that *explain* behaviour not as summaries of such behaviour (cf. e.g. Hausman 2000; Bermudez 2009; Dietrich and List 2016).<sup>9</sup> The revealed preference paradigm, while influential, is arguably inspired by a form of behaviourism we have little reason to accept (cf. Hausman 2000; Dietrich and List 2016). Choice behaviour can be *evidence* for preferences without determining them.

I therefore do not take the preference determines action principle as following from the definition of preference. Rather, I take it as a substantial and explanatorily powerful hypothesis about intentional action. It seems to be deeply embedded in our folk-psychology, a "truism" (Kennett and Smith 1996, 1997), and it accords well with modelling in, for example,

---

<sup>8</sup> Thanks to Katharine Browne and Jurgis Karpus for pressing this point on me.

<sup>9</sup> Okasha (2016) argues that behaviourism about preferences might be correct for a normative (and not descriptive) theory of rational choice. This, though, is not what we are concerned with at this point.



the economic sciences. A theorist who rejects the principle will need to show why it fails where it fails and how it needs to be amended. Such rival explanatory schemes will need to be judged by whether they are better than the one promised by the principle itself.

I will show that there is such a better explanatory scheme.

#### 4. Attention and the Non-Agential View

In my view, synchronic self-control strategies are ways of re-focusing attention, or re-prioritizing mental states, as I will call it. The capacity for self-control, accordingly, consists in a skilled competence at such re-prioritizing in situations where the agent is faced with a contrary preference. What the preference determines action principle leaves out is that attention acts as a flexible interface between preferences and actions. An agent's distribution of attention mediates between the agent's motivational systems and her sensory situation. Because their influence is so mediated, the agent can intentionally decouple her preferences from the relevant action. This is what she does when she intentionally engages in synchronic self-control.

Before I present my own re-prioritization view of self-control, let me briefly look at one that looks similar but promises to keep the preference determine action principle untouched. The view is Jeannette Kennett's (2003) (see also Kennett and Smith 1997). Here is how she puts it:

When an agent realizes that her actual desires do not match her judgements of desirability, and that she is therefore in danger of losing control of what she does, there are three ways in which she may focus her attention so as to bring it about that she does as she believes she should. First, she may restore the focus of her attention. Second, she may narrow or redirect the focus of her attention. Third, she may expand the focus of her attention.

(Kennett 2003, p. 136)

Kennett here seems to agree that self-control is achieved through strategies for re-focusing attention.

But *how* do re-focusing strategies work? What is distinctive of Kennett's view is that she thinks that the relevant attention shifts are never intentional actions (they can't be, for Kennett, because of "truisms" like the preference determines action principle). Re-focusing of attention, for Kennett, then is always "distinctively cognitive" or "a matter of her entertaining or excluding certain thoughts at the appropriate time" (op. cit., p. 139). The relevant changes

of attention thus are not explained on the basis of the agent's motivational states. Therefore, there is no motivational conflict with the motivation associated with her preference.

One challenge for the non-intentional view is to explain how the relevant un-motivated changes are interestingly different from an accidental avoidance of the loss of control threat. You don't eat the cookie because of a sudden stomach ache, and a sudden thought about your looming taxes prevents you from punching a police man. The fact that these new "thoughts" occupy your attention at just the right time will, as a matter of fact, avoid the loss of control threat. But these are not instances of self-control.

This specific challenge, arguably, can be answered. For example, one might say that the relevant attention shift, while not an intentional action, must be the result of a reliable competence, a subject-level dispositional capacity. In the stomach ache and urgent tax examples the relevant attention shifts are not the result of such a reliable competence. Maybe we could even call the deployment of such capacities an agent's "activity" (cf. Schellenberg 2018 on perceptual consciousness as an activity).

This still leaves the feeling that something has been left out: the agent's choice of the better option seems intentional and motivated in a way perceiving a red dot (one of Schellenberg's examples) is not: maybe the agent deploys reliable capacities also in the latter case, but perceiving still seems to be just happening to the agent: she has no voluntary control over it, and cannot perceive intentionally.

We can sharpen this problem: the restoring of the focus of attention, or the narrowing, re-directing or expanding of it, all – unlike perception, and unlike having thoughts – *are* exercises of agency: attention shifts, as I argue in detail in Watzl (2017), are always based in the agent's motivational system. While it may be true that "people exercise control over their own *thought* processes simply by having the thoughts that they are disposed to have" (my emphasis), as Kennett and Smith (1997, p. 124) put it, people control their *attention* like they control their (other) actions: try listening to the subtle flute in the big orchestra and keep your attention on it. This is an action you can control like you can control the movements of your fingers. Given that attention *can* be shifted intentionally, an unintentional attention shift seems accidental. The description in terms of attention arguably makes Kennett's view seem unlike the case of the thought about the looming taxes. But it is illegitimate, since it smuggles in an agential element that the view officially disavows.

The re-prioritization account, I will now argue, can explain self-control through attention while fully acknowledging its agential character.

### 5. The Re-Prioritization Account of Self-Control

What is important for the role of attention in self-control is not that the relevant shifts of attention are unmotivated. What is important is how attention affects the agent's mind. Attention organizes mental states in an action-relevant way: how much of an agent's attention a mental state occupies makes a difference to how that state influences the agent's actions (I take it that this is also a truism of folk-psychology). By intentionally changing how much of her attention a mental state occupies, the agent can therefore intentionally change her course of action. Based on this simple idea, I argue that an agent can through an intentional change to her distribution of attention intentionally prevent herself from acting on a preference, and thus intentionally engage in self-control. This self-control strategy, I argue, requires no special motivational powers, and explains everything about self-control that needs explaining.

The attentional account of self-control draws on both the shifting priorities model recently proposed in the empirical literature by Michael Inzlicht and others<sup>10</sup>, as well on philosophical accounts of attention of attention and its role for action provided by Wayne Wu (2014, 2016) and myself (Watzl 2017). I will provide a philosophically satisfactory and elaborated account that shows how “[a]ttention plays a crucial role in ... self-control by gating which options enter the choice set at any one moment and foregrounding their salient attributes” (Berkman et al., 2017, p. 423).

I will first illustrate the basic idea. As a starting point ask how a preference actually brings about an action. The way preferences bring about intentional actions, I argue, is through the agent's distribution of attention: attractive, affectively loaded or action-relevant features of the relevant option will be highly salient to her, drawing her attention to them. Further, her preference directs her attention to relevant targets for action: from all possible targets for action Christina must actually find the cookie before she could eat it. Intentional action, in most cases, requires an alignment of the distribution of attention with one's preferences.

But since preference influences action through the agent's distribution of attention, I argue further, the agent can intentionally interfere at this stage. Importantly, in doing so the agent

---

<sup>10</sup> Cf. Inzlicht, Schmeichel and Macrae 2014; Inzlicht and Berkman 2015; Berkman et al. 2017. For a related discussion of the role of attention for various forms of control in the literature on artificial intelligence, see also Bello and Bridewell 2017.

does not act from any motivation that does not derive from her preferences. It is compatible with a preference of A over B that the agent prefers that the attractive, action-relevant properties of A and the objects targeted by A-ing occupy less of her attention than they actually do (in a self-control situation, the agent may have that latter preference alongside the first because she believes that B and not A is the better option).

But if the agent prefers that certain properties of A or objects targeted by A-ing occupy less of her attention than they currently do, then she can, based on that preference, intentionally shift her attention away from those properties or targets. This can break the causal link between the preference and the action (decoupling the action from the preferences) and so the agent will end up not acting on her preference of A. This in turn will often lead to a preference change so that she now starts preferring B over A. And so, Christina ends up going on her run, and Amira back in his car – with no special motivational powers needed.

In the following sections, I present this idea in terms of a detailed argument based on four premises.

My presentation of this argument will use some aspects of the idea, defended in Watzl (2017), that attention is a form of mental agency where agents act on their own minds by changing or maintaining the priority ordering of their occurrent, subject-level mental states. According to this *priority structure view*, attention consists in the agent's activity of regulating priority structures, which order the parts of the subject's current state of mind by their current priority to the subject: when an agent is perceptually attending to a perceptually presented item or a perceptually presented feature, she is prioritizing some parts of her overall perceptual state over other parts of that state. If attention is, for example, visually focused on an object, then the state of seeing that object is prioritized over other parts of that visual state. When attention is focused on a feature like the colour of an object, then the state of seeing that feature is prioritized. The priority structure view also allows for non-perceptual forms of attention. These are, for example, at issue when Kennett speaks of attention as being "a matter of ... entertaining or excluding certain thoughts" (op. cit., p. 139). Attention can bring occurrent thoughts to the subject's mind and prioritize them over other parts of her current mental state. When attention is non-perceptual what is prioritized thus is a non-perceptual aspect of the subject's on-going mental life: this may, for example, be a thought, a bodily sensation or a feeling, or a mental image. The priority structure view thus unifies all forms of attention by taking as the primary notion the notion of a mental state's relative priority for the subject. The forms differ only in which aspect of the subject's mind has the highest relative priority.

Attention to external objects, on this view, thus gets explained by the relative priority of aspects of the subject mind. The fact that the agent's attention is directed at a cookie in front of her, or at the colour of the cookie, is constituted by the fact that the agent is prioritizing a mental state that is about that object or feature. When the agent is prioritizing seeing the cookie, thinking about it, or feeling an urge to eat it, then she is focusing her attention on the cookie in very different ways. Priority orderings thus are more fine-grained than the items the agent is attending to. The priority structure view also takes attention to be graded. A mental state can occupy more or less of the agent's attention depending on where in her current priority ordering it is located.

What we need from the priority structure view, for present purposes, is something fairly minimalistic and common-sensical: occurrent mental states occupy the agent's attention to various degrees and the degree to which they occupy the agent's attention is a matter of the priority they have for the agent at the relevant time. It is compatible with this view that there is some deeper, further account of what it takes for a mental state to have priority to the agent.

In what follows I will argue for two features of these priority orderings: first, for a role that priority plays in coupling an agent's preferences or intentions to her actions. Second, that an agent can intentionally affect the priority ordering of her mental states, based on her preferences for such an ordering, and that she can do so in the relevant self-control situation.

The argument showing how intentional, synchronic self-control is possible, takes the following form:

1. Non-deviant causal links between a preference for A over B and an intentional action A are mediated by the agent's distribution of attention, which is the associated priority ordering of the preference for A. (The Mediation Claim)
2. An agent who prefers A over B can at the same time, psychologically and consistently, prefer not to have the associated priority ordering of the preference for A. In this case, she has a diverging attention preference. (The Attention Preference Claim)
3. If the agent acts on her diverging attention preference, she can intentionally break the causal link between her preference for A, and her A-ing. (The Intentionality Claim)

4. If the agent intentionally breaks the causal link between her preference for A, and her A-ing in this way, she intentionally engages in synchronic self-control. (The Self-control Claim)

So, 5. It is possible for an agent, by intentionally re-distributing her attention, to intentionally engage in synchronic self-control.

Since no willpower, no mental resources, and no mental division are mentioned anywhere in this, this argument shows that it is possible to engage in synchronic self-control without them.

In the next sections, I defend the four premises of this argument.

## 6. Preferences and their Associated Priority Structures

Much modelling in decision theory and economics, and much philosophy of action, simply takes an agent's preferences and how they lead to action and choice behaviour as given. The mediation claim is the result of thinking more carefully about the link between preference and action.

Let us start with an intuitive idea and compare preferences to desire, where a characteristic link to the agent's distribution of attention has long been recognized: if one has a strong desire for an action, then one's attention will be insistently drawn to appealing properties of that action, or considerations that seem to count in its favour (Scanlon 1998), and one will be disposed to attend to things one positively associates with what one desires (Sinhbabu 2017). The same is plausible for preferences. When an agent has a preference for an option, then when it comes to making a decision, her attention will be drawn to something that seems positive or appealing about that option. If Christina, in front of the store, genuinely prefers the cookie over the run, then something motivating about that option must be on her mind at that time. Her attention, whether in perception or thought, must be directed toward the cookie, all that is good about eating it, and other appealing properties of that option. A preference that doesn't direct one's attention seems toothless and without motivational power. For a preference to become alive in an agent's decision making it must at the relevant time engage her attention and prioritize what is relevant for enacting that preference.

In a series of recent publications, Dietrich and List (2012, 2013) show how to integrate such intuitive considerations into a formal model of preference and rational choice. On their model, an agent's preference order is given by how she weighs her motivating reasons at the time of decision making. They summarize the model as follows (I won't here go into the formal details):

at any time, the agent is in a particular psychological state, represented by his or her set of motivating reasons in relation to the given alternatives, which, jointly with the agent's weighing relation, determines his or her preference order. This preference order then induces a choice function, which encodes how the agent would choose from any concrete set of alternatives (Dietrich and List 2013, p. 126)

According to Dietrich and List, the agent's preference-based choice of an option thus goes through being motivated by certain properties of that option.

When it comes to what makes a particular property motivating to a specific agent at some specific time, Dietrich and List suggest that at least one way is that the agent focuses attention on those properties when she is forming her preferences or is in a relevant choice situation. Drawing on "the concept of attentional salience as frequently used in psychology and behavioural economics" they suggest that a consideration becomes motivating to the agent if she "focuses on it actively or uses it as a preference-formation heuristic or criterion" (op. cit. p. 109). On the resulting view, then, preferences are formed "by focusing—consciously or otherwise—on certain properties of the alternatives as the motivationally salient properties" (Dietrich and List 2012, p. 622), and preferences are changed "when new properties of the alternatives become motivationally salient or previously salient properties cease to be salient" (ibid.).

If an account like Dietrich and List's is right, then preferences are partially constituted by having one's attention drawn or directed to motivating reasons (subjectively motivating properties) for particular actions or options. Therefore, certain priority structures will be constitutive of those preferences, and all preference-based action will be mediated by an associated priority ordering. Preferences lead to choice-behaviour in part through activating parts of their associated priority structures.<sup>11</sup>

We can further support the claim that the preference-action link is mediated by attention by drawing on work by Wayne Wu (2014, 2016) about the role of intentions in the production of action. What Wu says about intentions, we will see, transfers to preferences.

An agent's intentions, Wu observes, are not events that happen at a particular time and kick off an action. They are "standing states that persist over time" (Wu 2016, p. 106). How then

---

<sup>11</sup> Note that in order to accept this last claim, we don't here need to define preferences in terms of those associated priority structures.

do those standing states lead to specific actions at specific times? According to Wu, they are structural causes of action. They structure how an agent selects a behavioural path in a space of behavioural possibilities. In a specific situation, an agent may be aware of various potential targets for action (various things she perceives or thinks of) and in regard to those potential targets she is aware of various behavioural options: what she can do with the targets. Any agent that does anything (whether intentional or not) must couple the potential targets for action to a behavioural response (Wu calls this the Many-Many problem). The causal role of intentions, Wu argues, is to structure an agent's behavioural space: they bias her choice toward certain types of behavioural responses to certain perceptual situations. Intentions set "the weights that bias which selections are made in action" (Wu 2016, p. 110).

How is this linked to attention? For Wu, attention just is the selecting of a specific perception-behaviour mapping in a specific situation (Wu speaks of attention as selection for action). The causal role of intentions in the production of action, for Wu, is that they bias the agent's attentional system toward selecting certain responses in a certain range of perceptual situations. Intention guides action through the agent's deployment of attention.

The question Wu asks with regard to intention also arises with regard to preferences. An agent's preferences are standing states. An agent has certain preferences over some period of time (on the Dietrich and List model this would be the agent's weighing function). She may have formed those preferences at some particular moment, and she may revise her preferences later. But the preferences themselves are not events that occur in the agent's mind at some particular moment. They are standing states. Given what Christina perceives in her specific situation, or what – more generally – she is aware of in that situation, there are many different things she could do with regard to those things. What role do her preferences play in selecting one behavioural path rather than another? We should be inspired by Wu's account of the role of intentions in the guidance of action: preferences lead to action by biasing the agent's attentional system toward selecting responses in a range of situations.

This Wu-inspired thought again minimally implies that preferences have associated priority structures: a preference is linked to certain ways of attending. Preferences do not kick off actions by themselves. They cause actions at least in part by adjusting what the agent will attend to in a certain range of circumstances. In my own terminology: the influence of preferences on action is mediated by the agent's priority structures, i.e. by how much attention she pays to what she is aware of in various situations. Preferences set the weights for coupling



items in the situation the agent is aware of to behavioural possibilities. They are (at least in part) dispositions to select items the agent is aware of as priorities for action.

We can accept that the influence of preferences on action is mediated through attention without accepting Wu's further claim that attention just is selecting items for action. We can see how by putting together Wu's ideas with those of Dietrich and List: attention mediates between preference and action because a preference disposes an agent to attend to certain items *and their motivationally relevant properties*. It lets the agent see a certain action possibility in a positive light. And seeing it in that positive light is an event that causes the agent to choose the action. Preferences are structural causes of action. Focusing on motivationally relevant properties is an event cause.

We have therefore intuitive considerations for the mediation claim, a formal model that supports it, and the philosophical considerations just mentioned.

The claim that choice behaviour is mediated by the agent's distribution of attention is not a mere abstract possibility, though. There is also empirical work in its support. While non-perceptual forms of attention, according to what I have argued so far, clearly also play a role in linking preferences and behavioural choice, some of the most detailed empirical work has shown the importance also of perceptual attention. Krajbich, Armel and Rangel (2010), for example, show that a very simple choice between two images is largely predicted by how much time subjects spent looking at the respective images. A large effect remains even when how much a subject *likes* a specific image is controlled for. Many other such effects are known. Based on a review of more than 65 studies, Orquin and Loose (2013) have concluded "that attention plays an active role in constructing decisions" (p. 203). Attention leads to a variety of "downstream effects on decision making" (ibid.). The known empirical work on attention and choice behaviour thus like the philosophical consideration supports the view that preference-based choice is largely mediated through the agent's distribution of attention leading up to and at the time of decision making.<sup>12</sup>

---

<sup>12</sup> A particularly powerful illustration of the interaction of attention and choice concerns drug addiction. It has been found that drug-related features are highly salient for the addicted person and consequently tend to draw her attention away from affectively more neutral features. This has been tested through the so-called addiction Stroop test, where subjects have to report on the colour in which a word is written. Here addicted people show longer reaction times and higher rates of mistakes when asked to report the colour of drug-related words compared to the control group. For the addict "salience attribution transforms the sensory features of the

## 7. Action preferences and attention preferences

If the link between preference and action is mediated by associated priority structures, as I have argued, then it is possible to change whether the action results by interfering with the priority structures. In this section, I will argue that the agent can *intentionally* interfere with those priority structures and thereby intentionally decouple the action from her preferences. I will again focus on philosophical considerations for this claim, but I take those broadly in line with the conclusion drawn by Orquin and Loose (2013), in the review just mentioned, that the role of attention in decision making is "constructive" since decisions emerge "not as a simple application of preferences and heuristics to choice stimuli but, through complex interactions among stimuli, attention processes, working memory, and preferences." (p. 203).

An agent's preference for an option disposes her to pay more attention to motivationally relevant properties of that option as well as aspects of what she is aware of that will couple her preference to a situation specific action that enacts that preference. This is the associated priority structure of that preference. The point of this section is to argue that this does not imply that the agent also *prefers* to attend to those motivationally relevant properties and perceptually presented items. I will thus argue for the attention preference claim, i.e. the claim that an agent who prefers A over B can at the same time, psychologically and consistently, prefer not to have the associated priority ordering of the preference for A, and that she can intentionally act on that diverging attention preference.

For this discussion it is important to recognize that attention is a form of agency just like embodied action (cf. Chapter 7 of Watzl 2017). An agent can choose to focus her visual attention on an item just like she can choose to pick up that item with her hand. And she can control the focus of her attention in ways that are highly similar to the way she controls her bodily movements. An agent who, for example, engages in a visual search task (looking for Waldo in the famous cartoon drawings) may explicitly adopt a particular strategy: she may scan the picture with her visual attention from left to right and top to bottom.

Specifically, what we need here is that agents often have preferences regarding what to attend to. When listening to a piece of orchestral music, someone with an interest in the flute, might prefer to focus her auditory attention on the melody of the flute over focusing it on the sound

---

incentive stimulus into an especially salient percept, which 'grabs attention', becomes attractive and 'wanted' and thus guides behavior to the incentive" (Robinson and Berridge 1993, p. 261).

of the violins. And you might prefer to look at one of the photographs in your office over some of the other photographs. Agents also often have preferences regarding which *features* to attend to: in one situation, the agent might prefer to focus on the rhythm of the melody that the flute plays rather than how loudly the notes are being played, while in a different situation, she might prefer to focus on the loudness instead. Or she might prefer to look at the shape of her dining room table (a very pleasant shape) rather than visually focusing on its colour (rather less pleasant). The same holds for attention in thought: when thinking about his last vacation, Amira might prefer focusing his attention on how peaceful it felt rather than on the fact that it was raining most of the time. Or consider attention to bodily sensations: you might have a preference for attending to the aftertaste of coffee in your mouth over attending to the unpleasant numbness in your leg.

In other words, agents have preferences over priority structures. They prefer to have some priority structures over others. These preferences, of course, are often not extremely fine-grained: there are many possible priority structures that are compatible with an agent's preference for focusing most of her attention on the rhythm of the flute. The attention preference need not specify how much attention she pays to the violins and how much should be left for visually taking in her surroundings. The same holds for other preferences: Christina's preference for eating a cookie over going for a run does not specify which cookie she eats, the details of how she is going to get it, or how she will put it in her mouth. I will thus assume that agents have attention preferences: ways of attending (i.e. priority structures) are among the options agents' preference relations range over.

One might ask: how is the idea that we have attention preferences compatible with the claim of the last section that attention mediates between preference and action? Isn't there some problematic circularity or infinite regress? No. Attention also mediates between an attention preference and the act of attention. Suppose you prefer paying attention to painting A over paying attention to painting B. How does that preference actually get you to pay attention to A rather than B? The mediation claim applies here just like it applies in other cases. Your preference disposes you to focus on properties that make paying attention to A appealing, and it structures your attention at one time to select targets for your attention at a later time. Once the motivating properties of paying attention to A are salient to you, you likely to end up paying attention to A. There is no circularity or infinite regress. A preference for one priority structure over another biases an agent's *other, mediating*, priority structures toward positive features of that priority structure. What would lead to an infinite regress is if we were to claim that the only way an agent could arrive at the *mediating* priority structures is again via a preference for those mediating structures. But this is not what is claimed here. The problematic claim would be analogous to claiming that the only way an agent's preference for an overt action could lead to that action is if the agent also had a preference for everything that mediates between her preference and that action. But such a claim ha Bello, P., & Bridewell, W. (2017). There is no agency without attention. *AI Magazine*, 38(4), 27-34.

s no plausibility.

Now suppose that an agent has a preference for some overt action A, like eating a cookie, over another overt action B, like going on a run. That preference has an associated priority structure. Call that  $P_A$ . Does this entail that the agent at that time also prefers to have that priority structure over relevant alternatives, like the priority structure associated with B,  $P_B$ ? No. The fact that  $P_A$  mediates between the agent's preference for A and A does not entail that the agent also prefers  $P_A$ . We have, in fact, already seen this in the last paragraph: a priority structure might mediate between a preference and an action without an agent also preferring to have that priority structure. This shows that it is possible that an agent prefers option A over an incompatible option B, yet she also prefers not to have the priority ordering associated with A, but rather have a priority ordering associated or at least compatible with B. In other words, it is possible that the agent's attention preference diverges from her action preference.

One might ask: even if diverging attention preferences are possible in principle, is such divergence ever psychologically realized? I think it clearly is. Suppose you are at a supermarket, preferring to eat a cookie, which draws your attention to various motivationally salient properties of that action like the location of the cookies in a supermarket shelf. But you need not *prefer* to focus your attention on that location in the shelf: you want to eat the cookie, but I also want to buy some laundry detergent (though you want the cookie more). You might prefer to focus your attention on where the detergent might be, since you haven't found it yet. Your attention is *drawn* to the cookie locations, but you prefer it to be somewhere else. This is a perfectly ordinary situation, even aside from self-control cases.

Do diverging attention preferences necessarily make the agent's preferences inconsistent or incoherent? I don't think so. An agent might consistently and coherently prefer A over B, and C might be necessary for A, and yet the agent might prefer not to be in C. If the agent can be in C un-intentionally (i.e. not on the basis of a preference for being in C), then the agent need not prefer C over not-C in order to get to A. Maybe a coherent agent must prefer to take the means if she prefers the end. If that is true, then it is true only because means are intentional actions the agent must perform in order to get what she wants. But priority structures are not a means an agent takes toward an end she aims to achieve. Priority structures mediate between preferences and action, but not as intentional means to the action. Consider this analogy: in order to pick up a coin, I must move my fingers in a specific way. I can move my fingers in this specific way intentionally, but in order to pick up the coin I need not, and will normally not, move them intentionally in this way. Therefore, there is no

incoherence in assuming that while I really want to pick up the coin, I prefer not to move my fingers in the way that is required.

Still, one might say, there surely is some kind of conflict when one has an attention preference that diverges from one's preferences for overt action. And yes: of course, there is. If the agent acts on her attention preference, she will not realize her overt action preference. And if she acts on her overt action preference, she will not realize her attention preference. There is a subjective conflict. The agent can't get everything she wants. That just *is* the subjective conflict that, among other situations, characterizes self-control situations. The existence of a subjective conflict when the agent has divergent attention preferences is not a problem for the present view. It is part of the datum we wanted to explain.

I thus hope to have convinced you of the attention preference claim. It is psychologically and consistently possible for agents to have attention preferences that diverge from their action preferences.

#### 8. Intentionally breaking the preference-action link

I still need to defend the last two premises of the argument. In this section, I defend the intentionality claim, i.e. that if the agent acts on her diverging attention preference, she can intentionally break the causal link between her preference for A, and her A-ing.

Preferences are structural causes of action, I have argued. They bias the agent's attention system toward prioritizing action relevant objects and their appealing properties. Such appealing properties will be salient to the agent: her attention tends to get drawn to them. But if the agent has a diverging attention preference, she prefers to have a different priority structure instead. If she acts on that attention preference, she shifts her attention in a different direction or suppresses the relevant saliences. Since the causal link between her action preference and her action is mediated by those saliences, interfering on the mediating structure will unlink the preference from the action. Since that interference is based in the agent's preferences, and there are no deviant causal chains, the interference is intentional. Therefore, if the agent acts on her diverging attention preference, she can intentionally break the causal link between her original preference and her action. Which is just the intentionality claim.

Let me illustrate how agents intentionally unlink their preferences from action with one of the most famous experimental studies of self-control: the delay of gratification experiments, pioneered by Walter Mischel in the 1960s. Here children are presented with two options. The

first option is to eat one cookie (or Marshmallow) now. The second option is to wait and get *two* cookies later. The cookie in front of the child, we assume, has a powerful grip on her, while the child views the second option as subjectively better (two are better than one). Once the child is sitting in front of the one cookie and the experimenter has left the room, she prefers the one cookie over waiting for a second cookie later. But she thinks that she should wait for the second cookie. One nice feature of the delay of gratification paradigm is that it gives us a quantifiable measure of a self-control strategy's effectiveness in terms of the children's waiting time.

The delay of gratification results illustrate how preference is mediated by attention, and how this allows for intentional interference by the agent.

One important result is that brute effort of attention (unlike what William James may have thought) is not an effective strategy. Suppose that the experimenter primes subjects to try hard to keep attention focused on the two-cookie goal by, for example, instructing children to mentally rehearse "I will get two cookies later." Keeping the two-cookie goal mentally alive prioritizes mental images of or thoughts about two cookies. Such strategies shorten waiting times relative to a neutral base line. With the mediation claim we can easily see why: thinking of two cookies is thinking of cookies. Activating the goal of two cookies makes the one cookie in front of the child, and especially its yumminess, extremely *salient* to her. The child's attention is drawn to a motivationally salient property of the one cookie. Activating the two-cookie goal thus supports the priority structure that mediates between the preference and the cookie. Therefore, it makes the child more rather than less likely to fall for the temptation. Preferring to think of two cookies later is not, after all, a diverging attention preference.

Effective strategies, by contrast, involve intentionally re-orienting attention in a way that *lowers* the priority ranking of states that mediate between the child's preference and the action of eating cookies. One effective strategy is self-distraction. The child might think about a fun activity or sing songs in her head. These are intentional actions. Children who are primed by being informed of the availability of such strategies, and thus are more likely to use them, have increased waiting times. Since here the children are intentionally prioritizing images and thoughts that have nothing to do with cookies, these strategies, unlike refocusing on the two-cookie commitment, do not make the one cookie more salient: by increasing the salience of, say, mental images of playing on the beach, the relative salience of the yumminess of the cookie is decreased instead. Similarly, waiting times are increased when the child

intentionally focuses attention on shape or colour of the cookie. Now she is deprioritizing the perception of the affective properties (its yumminess) of the cookie by prioritizing its non-affective properties (its colour or shape): she intentionally shifts attention away from motivationally salient properties of cookie eating.

In order to make something intrinsically boring, like the shape of a cookie, easier to attend to, techniques for reconstrual help, like thinking of the cookie as a UFO and of the raisins as the aliens that ride it. It also helps to build associations between seeing the tempting object and something else (you may train yourself so that every time you see a cookie you think about salads). The formation of such habits does not make the relevant shifts of attention automatic or unintentional. A habit-based action like switching on the light when entering your house, or putting on your running shoes when the clock strikes, is still an intentional action.

Implementation intentions (Gollwitzer and Brandstätter 1997) are still intentions. And habitual preferences are still preferences.

Of course, a child, or an adult, must know what to attend to in order to live by a specific normative judgment about what she should do. By acting on a diverging attention preference it is *possible* to intentionally disconnect a preference from an action. But this does not entail that you succeed. Self-control attempts by means of intentional re-prioritization might fail, and they often do. The relevant attention preferences must be linked up with the subject's normative judgment and they must be effective: as the case of focusing on the two-cookie goal shows, what the agent *thinks* is a priority structure that will decouple her preference from action might not actually be a priority structure that has this effect.

#### 9. Attention preferences and the Judgments that drive them

Suppose that the agent intentionally breaks the causal link between her preference for A, and her A-ing by a shift of attention that destroys the associated priority structure of that preference. Does that amount to an intentional act of synchronous self-control? According to the last step in my argument, the self-control claim, it does. I argue that it does by answering two objections. One is that the relevant avoidance of the loss of control threat isn't really synchronous. The other that while it might be an intentional action, it is not an intentional act of self-control.

With regard to synchronicity, the objection would be that the re-prioritization strategy really describes diachronic self-control. Specifically, let us ask whether the account can explain *full-blooded* self-control, which Sripada (2014) raises as a problem for a view of self-control by

Alfred Mele (1992) with which the present account has some similarities (for the objection that Mele's view only allows for diachronic self-control see also Kennett and Smith 1997).

According to Mele, agents engage in synchronic self-control by engaging in a second, *ancillary*, intentional action that is compatible with acting on her wayward preferences, and that will, in due course, change her motivation. The re-prioritization account agrees that self-control is achieved through a secondary action. Here this is an intentional change to the agent's priority structures.

Full-blooded exercises of self-control, according to Sripada, are cases where the agent never even begins to act on her wayward preference (or as he says, her "strongest desire"). Sripada believes that there are such cases. Christina might never even begin to take a step toward the pastry shop, and Amira might not begin to initiate his punch. Sripada objects that Mele's view, by contrast, entails that the agent will always begin to engage in the preferred and wayward act. On Mele's view, the agent begins to engage in the wayward action; while it is on its way she engages in a compatible ancillary action, like telling herself "Remember the run!". This, Mele plausibly believes (and Sripada and Kennett and Smith agree), may change her motivations (on Mele's view this indeed often happens through a shift in attention), which in turn interrupts her tempting action, and thus the agent achieves self-control. Sripada thinks that this account neglects the full-blooded cases.

Does the re-prioritization account, arguably like Mele's, rule out such full-blooded self-control? No. An agent's preferences, as I said, always lead to action through her priority structures. Since re-prioritization acts on these priority structures, the agent may never initiate the wayward action. Christina and Amira can decouple their preferences from action through a shift in priorities, before they initiate their preferred action. The re-prioritization account differs from Mele's in exactly the crucial respect: since priority structure mediate between the preference and the action, the act of changing those structures is not one the agent performs while already engaged in her preferred action; she performs it before she ever engages in it.

Sripada and Kennett and Smith might object that the agent *does* initiate her preferred wayward action. Her cookie preference makes appealing properties of eating cookies salient to Christina, and draws her attention to where the cookies are found. Even if she interrupts herself by, say, shifting her attention to a mental image of her last vacation, she has already begun to act on her preference. The crucial question seems to be: is prioritizing positive features of eating cookies already a part of the action of eating cookies, or is it not? If we think of actions as the event specified in the content of the agent's preferences, then the



associated priority structures are not part of the action, and the present objection fails. If, by contrast, we think of actions as complex temporally extended processes that include the mental precursors of the event specified in the agent's preferences (cf Dretske 1991), then the associated priority structures are part of the action. But on this way of individuating action the agent's preferences (and her beliefs) are *also* part of the action: the action is the process of the preferences and beliefs leading to the bodily behaviour. But then the action has already begun when Christina first forms her preference for that cookie, and full-blooded self-control would be ruled out by the very description of the loss of control threat situation. Sripada and Kennett and Smith therefore cannot appeal to it.

I therefore conclude that the re-prioritizing strategy makes self-control synchronic in the relevant sense.

The next objection is that while the agent may do something intentionally to avert the loss of control threat (based on her attention preferences), what she does intentionally is not an intentional act of self-control. The agent must engage in the act *in order* to do as she believes she should do. The mere fact that she does something that as a side-effect leads the agent to choose the option she judges to be correct is not enough to show that her act was intentional *qua* self-control act.

I agree that an intentional self-control act must be the result of the agent's view of what is the subjectively better option, in our case, her normative judgment. If a child in Mischel's experiment just happens to prefer focusing her attention on the shape rather than the yumminess of a cookie, and as a result waits longer, this would not be an intentional act of self-control. We indeed have to demand that the agent must form her relevant attention preference based on the reasons for the better option she is aware of in her normative judgment. But it is guaranteed that there are such reasons the agent is aware of: in a subjective conflict situation that characterizes self-control cases the agent is aware of the better option as better. So, she sees reasons for choosing the better option. For an intentional self-control act those reasons must inform the formation of her attention preference by means of which she decouples her momentary preference from action. Those reasons clearly are motivationally relevant for what to attend to in that situation. Thus, by the Dietrich and List (2012, 2013) account, she can form that attention preference by focusing on those reasons.

Of course, as we have already seen, the fact that the agent's attention preferences are informed by her reasons in this way does not mean that she knows how to attend in order to secure the better choice. The child might try to focus on a mental image of two cookies in order to resist

her temptation. But that, as we have seen, is not an efficient self-control strategy. It is still an intentional self-control attempt, since the child might re-focus her attention in this way based on the reasons for the better option she is aware of. And the attempt might, in some cases, and for some period of time, succeed. In this case, it would be a synchronic act of self-control. For more efficient self-control, the agent must have, what might be called, *an attention skill*: she must be disposed to bring to bear situation specific knowledge of what to attend to in order to secure her the choice of the better option (cf. Pavese 2016). Having a disposition to bring to bear certain knowledge in the relevant situations thus is part of how we succeed in self-control (I here agree with ideas in Kennett and Smith 1996, 1997). But when we re-focus our attention based on that knowledge we engage in an ordinary intentional action, based on our preferences, which we form in light of that knowledge. Synchronic self-control acts can be intentional attempts *to* secure the better option.

#### 10. How to live without special motivational powers

I hope to have shown that we can explain synchronic self-control without appeal to any special motivational powers. Once we understand how the horses pull the chariot, i.e. how preferences are linked to action, we need no more than horses. I also hope to have shown that the possibility of such self-control follows naturally from the role of attention in the control of all action. The resulting picture, I believe, is psychologically plausible as a description of how agents actually engage in self-control and may succeed at it.

I will end by answering two further questions.

First, according to the re-prioritization model, self-control involves no will-power, and no battle between two motivational systems. What then explains the phenomenology of effort that seems to accompany self-control? If self-control just is an ordinary action, why does it feel so hard?

In my view, the sense of effort is probably the result of a number of features of self-control situations.<sup>13</sup> First, re-prioritization in order to achieve the subjectively better result is an error-prone process with an uncertain outcome. The agent cannot be sure that the way she chooses to re-focus attention will actually lead to the correct result, and it is easy to make mistakes. It has been shown that experienced effort and judgments of error-likelihood are strongly

---

<sup>13</sup> In a related context, Sinhababu (2017) suggests that we should explain the sense of effort involved in self-control in terms of effort of attention. This, though, clearly begs the question (what, after all, would explain the sense of effort of attention?).

correlated (see Inzlicht, Schmeichel and Macrae (2014) for this idea, and Dunn, Inzlicht and Risko 2017 for some detailed results). Second, the agent's attention will likely fluctuate between, on the one hand, priority structures that facilitate choice of the preferred option and ones that accord with the one she judges to be better: given that the agent's action preferences persist, they will keep drawing her attention to relevant items and their appealing features even though the agent prefers to have a different distribution of attention. Such fluctuations of attention between conflicting cues and the associated valuation fluctuation have been shown to be linked to experienced conflict and task difficulty (Krajbich, Armel and Rangel 2010; Kiani, Corthell and Shadlen 2014; Desender, Van Obstal, Van der Bussche 2017; cf. Berkman et al. 2017). Third, there is the hypothesis that subjective effort is linked to opportunity costs (Kurzban et al. 2013): given that the subject is aware of two options, one of which she prefers and the other she judges to be best, effort might signal that there is another alternative distribution of attention that the agent is not having. Generally, given the complexities of the psychological processes known to be involved in the generation of subjective effort, it would be a mistake to take such subjective effort as much evidence for the existence of a will-power faculty or a battle between two motivational systems. The re-prioritization account is consistent with what is known about the experience of effort.

Second, and relatedly, one might ask how – on the present view – we are supposed to explain why self-control attempts seem to show ego-depletion effects of the kind summarized in Baumeister, Bratslaviski and Muraven (2018). Holton (2009), for example, mentions those findings in support of the view that like a muscle willpower may become weak, and can be trained by its repeated use (appealing to Muraven and Baumeister 2000).

Self-control can be trained because attention skills can be trained. It is thus compatible with the re-prioritization account that self-control can be improved through training. What the account denies is that willpower can be trained *like a single, specialized muscle*. And for that the evidence is decidedly mixed (cf. Inzlicht, Schmeichel and Macrae 2014). Generally, the evidence for ego-depletion has come under attack, and is subject to replication failure (Carter et al 2015). One of the most surprising apparent findings seeming to support the existence of a psychic resource was that this resource was allegedly measurable in terms of blood glucose levels (Gailiot and Baumeister 2007). Yet, it isn't clear that the results that self-control decreases glucose levels hold water: even rinsing the mouth with glucose might enhance self-control (Sanders et al. 2012). Based on such findings, it has been argued that we can better explain the effects of glucose on self-control without positing any psychic resources. Kurzban

et al (2013), for example, argue that the detection of glucose signals “success” and is thereby motivating (see also Levy 2016). Generally, I agree with Inzlicht, Schmeichel and Macrae's (2014) assessment that "mental resources" may be, just as David Navon (1984) argued more than 30 years ago, a "theoretical soup stone", a mysterious entity that seems important, but that in fact is explanatorily empty once we have a fuller picture of the mechanisms actually involved in self-control.

## 11. Conclusion

I have argued that we can substantiate the intuitive idea that attention and self-control are connected. Self-control is indeed a matter of shifting the focus of attention in the right way. Yet, self-control is mostly not achieved by an effort of attention, trying hard to keep one's better options in clear view. Rather, it is achieved through a complex set of attentional skills. Attention is employed in self-control just like it is employed in other forms of agency. It acts as a flexible interface between our standing preferences and our actions. Agents have a form of agential flexibility over action because they need not translate their preferences directly into action. Through attention agents gain a form of freedom: that freedom, though, is not a special freedom that shows the exercise of a mental faculty or an aspect of the mind that comes into play only when self-control is called for. Rather, it is a freedom that attention weaves into the structure of agency quite generally.

My view of self-control is thus deflationary: we use attention in order to realize our goals, short or long term, we use it to act on our commitments or in favour of situational demands, we use it to realize what we think is right or we use it to act against our best judgment. While some have argued that there is a special rational demand to be enkratic (cf. Broome 2013) – whether or not that ever is achieved through special powers – I believe that the deflationary view argued for here lends some further support to views on which there is no such demand (cf. Audi 1990, Arpaly 2004, Brunero 2013, Reisner 2013). Christina's judgment that she should run and eat salad might be the result of a skewed deliberative process, channelled by problematic societal pressures regarding ideals of beauty and health. The cookie might in fact be good for Christina, making her a better runner, and a happier person, and her "temptation" in front of the store might be the result of sensitivity to exactly those reasons. Amira's anger, similarly, might be a rational response to the police man's behaviour, making Amira more rather than less sensitive to the normative reasons for action present in the situation (see Srinivasan 2018 on anger; and D'Cruz 2013 on reasons that might be in principle inaccessible through deliberation). Self-control might be irrational in the sense that it makes Amira less

responsive to the reasons present in the situation at hand. If self-control is not achieved by reason winning over passion, or by the self taking control of the horses, then plausibly self-control also need not always be what is most rational. If the normative structure of the situation is anything like the descriptive structure, then practical wisdom is likely "unprincipled" (Arpaly 2004).

## References

- Arpaly, N. (2004). *Unprincipled virtue: An inquiry into moral agency*. Oxford: Oxford University Press.
- Baumeister, R. F., Bratslavsky, E., & Muraven, M. (2018). Ego depletion: Is the active self a limited resource?. In Baumeister, R.F., *Self-Regulation and Self-Control* (pp. 24-52). New York: Routledge.
- Bello, P., & Bridewell, W. (2017). There is no agency without attention. *AI Magazine*, 38(4), 27-34.
- Berkman, E. T., Livingston, J. L., & Kahn, L. E. (2017). Finding the “self” in self-regulation: The identity-value model. *Psychological Inquiry*, 28(2-3), 77-98.
- Berkman, E.T., Hutcherson, C.A., Livingston, J.L., Kahn, L.E., & Inzlicht, M. (2017). Self-control as value-based choice. *Current Directions in Psychological Science*, 26, 422-428.
- Bermúdez, J. L. (2009). *Decision theory and rationality*. Oxford: Oxford University Press.
- Broome, J. (2013). *Rationality through reasoning*. Oxford: John Wiley & Sons.
- Brownstein, M. (2018). Self-control and overcontrol: Conceptual, ethical, and ideological issues in positive psychology. *Review of Philosophy and Psychology*, 9(3), 585-606.
- Brunero, J. (2013). Rational akrasia. *Organon F*, 20(4), 546-566.
- D'Cruz, J. (2013). Volatile reasons. *Australasian Journal of Philosophy*, 91(1): 31-40.
- Desender, K., Van Opstal, F., & Van den Bussche, E. (2017). Subjective experience of difficulty depends on multiple cues. *Scientific reports*, 7, 44222, 1-14
- Dietrich, F., & List, C. (2013). A reason-based theory of rational choice. *Nous*, 47(1), 104-134.

- Dietrich, F., & List, C. (2013). Where do preferences come from?. *International Journal of Game Theory*, 42(3), 613-637.
- Dretske, F. I. (1991). *Explaining behavior: Reasons in a world of causes*. Cambridge, MA: MIT press.
- Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, 11(1), 35-55.
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L.E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 92(2), 325-336.
- Ganeri, J. (2017). *Attention, not self*. Oxford: Oxford University Press.
- Gollwitzer, P. M. and Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, 73(1), 186–99.
- Hausman, D. M. (2000). Revealed preference, belief, and game theory. *Economics & Philosophy*, 16(1), 99-115.
- Holton, R. (2009). *Willing, wanting, waiting*. Oxford: Oxford University Press.
- Inzlicht, M., & Berkman, E. (2015). Six questions for the resource model of control (and some answers). *Social and Personality Psychology Compass*, 9(10), 511-524.
- Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, 18(3), 127-133.
- James, W. (1890 [1981]). *The principles of psychology*. Cambridge, MA: Harvard University Press.
- Kennett, J. (2001). *Agency and responsibility: A common-sense moral psychology*. Oxford: Clarendon Press.
- Kennett, J., & Smith, M. (1996). Frog and toad lose control. *Analysis*, 56(2), 63-73.
- Kennett, J., & Smith, M. (1997). Synchronic Self-control is Always Non-actional. *Analysis*, 57(2), 123-131.
- Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, 84(6), 1329-1342.

- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292-1298.
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661-679.
- Levy, N. (2016). The sweetness of surrender: Glucose enhances self-control by signaling environmental richness. *Philosophical Psychology*, *29*(6), 813-825.
- Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. Oxford: Oxford University Press on Demand.
- Mischel, W. (2014). *The marshmallow test: Mastering self-control*. New York: Little, Brown and Company.
- Mischel, W., Ebbesen, E. B., and Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*(2), 204–18.
- Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle?. *Psychological bulletin*, *126*(2), 247-259.
- Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*(3), 774-789.
- Navon, D. (1984). Resources - A theoretical soup stone?. *Psychological Review*, *91*(2), 216-234.
- Okasha, S. (2016). On the interpretation of decision theory. *Economics & Philosophy*, *32*(3), 409-433.
- Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta psychologica*, *144*(1), 190-206.
- Pavese, C. (2016). Skill in epistemology I: Skill and knowledge. *Philosophy Compass*, *11*(11), 642-649.
- Rawn, C. D., & Vohs, K. D. (2011). People use self-control to risk personal harm: An intra-interpersonal dilemma. *Personality and Social Psychology Review*, *15*(3), 267-289.
- Reisner, A. (2013). Is the Enkratic Principle a Requirement of Rationality?. *Organon F*, *20*(4), 437-463.

- Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews*, 18(3), 247-291.
- Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61-71.
- Sanders, M. A., Shirk, S. D., Burgin, C. J., & Martin, L. L. (2012). The gargle effect: Rinsing the mouth with glucose enhances self-control. *Psychological Science*, 23(12), 1470-1472
- Scanlon, T. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.
- Schellenberg, S. (2019). Perceptual consciousness as a mental activity. *Noûs*, 53(1), 114-133.
- Sinhababu, N. (2017). *Humean nature: How desire explains action, thought, and feeling*. Oxford: Oxford University Press.
- Srinivasan, A. (2018). The aptness of anger. *Journal of Political Philosophy*, 26(2), 123-144.
- Sripada, C. S. (2014). How is Willpower Possible? The Puzzle of Synchronic Self-Control and the Divided Mind. *Noûs*, 48(1), 41-74.
- Tangney, J., R. Baumeister, and A. Boone. 2004. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2): 271–324.
- Watzl, S. (2017). *Structuring mind: The nature of attention and how it shapes consciousness*. Oxford: Oxford University Press
- Wu, W. (2011). Confronting Many-Many Problems: Attention and Agentive Control. *Noûs*, 45(1), 50-76.
- Wu, W. (2014). *Attention*, Oxford: Routledge
- Wu, W. (2016). Experts and deviants: The story of agentive control. *Philosophy and Phenomenological Research*, 93(1), 101-126.