1 Prologue

Mental control refers to the ability we have to control our own minds. This may be accomplished by focusing our mental energy on a particular task, drumming up motivation for a particular action, inhibiting distractors or temptations, bringing various mental events into alignment, or perhaps even moving the body or moving to a specific environment in order to change our pattern of thinking. Take the example of playing ping-pong with friends: one uses mental control to focus on the location of the ball in space, to remain engaged at the end of a long rally, to ignore friends who are talking nearby, to coordinate one's conversation around the difficulty of play, and perhaps when moving one's body to a ready position, to better see and respond to a serve.

"Mental control" makes little showing in philosophy. Yet, the primary expression of mental control – attention – has become a more popular topic for philosophers in the past few decades. While scientists wrote extensively on attention in the twentieth century, philosophers focused on consciousness. In the twenty-first century, greater engagement between philosophy and the sciences led to more scientific exploration of consciousness and more philosophical analysis of attention. This Element thus focuses on mental control through the phenomenon of attention, while also covering some related topics (mind wandering and attention deficit hyperactivity disorder [ADHD]) and some general issues in the metaphysics of mind (mental causation and emergence). Because there is so little work on mental control as a broader concept, the Element is limited in providing a thorough analysis and review. It nonetheless provides some initial analysis of this concept, which should indicate a path for future work.

While philosophers have written little on the concept of mental control, they have written extensively on self-control. Mental control is broader than the concept of self-control, which typically refers to the maintenance of a preferred behavior in the face of temptation. Nonetheless, criticisms of the concept of self-control can also be applied to the concept of mental control. In particular, the criticism that self-control implies the existence of an unscientific *homunculus-like agent* also threatens mental control (see, e.g., Sripada 2021). Relatedly, the charge that self-control is *not a natural kind* can also be levied against mental control (see, e.g., Herdova 2017). Yet, this Element will leave open the possibility that a philosophically plausible and scientifically grounded account of mental control remains possible, at least when understood through the concept of attention.

This Element is organized as follows. First, it describes the phenomenon of mental control and nearby forms of control. Second, it covers the relationship

between mental control and attention, including the debate on whether mental control can be automatic. Third, it discusses the phenomenon of both meditation and mind wandering and how they relate to attention and mental control. Fourth, it closely examines a disorder of mental control, ADHD. Finally, it addresses the issue of emergence, one framework for thinking about mental control, discussing both traditional and contemporary accounts of emergence and how they might support mental control. That section puts forward a how-possible account of emergent mental control that makes use of contemporary neuroscience.

A guiding principle of what follows is that mental control is a necessary part of our understanding of the mind, and that further philosophical and scientific work on the topic is thus crucial not only for a complete account of mental control but for a host of other philosophical debates that depend on our understanding of the mind (e.g. legal theory). The Element is thus as necessary as it is incomplete, yet I hope it will serve as a guide for those who would continue this important project.

2 What Is Mental Control?

As James might have said, everyone knows what mental control is. While this may be true in a general sense, a precise understanding is more elusive. The term was not widely used before Wegner's critical work on the subject, starting with the "white bear study" (Wegner et al. 1987).[1] As one psychologist puts it, "There is not any specialized history of conceptions of mental control as there is of memory, sensation, or various other cognitive processes" (Schneider 1993, 33). Wegner thus starts with an intuitive definition: "Normally, we seem to have a measure of control over our thinking" (Wegner 1988, 683). Use of the term remains closely tied to both the intuitive notion and Wegner's own work.[2] In this section I aim to draw out contemporary use of the concept while connecting it with nearby work in philosophy.

According to the intuitive use of the term introduced above, mental control is control *of* the mind *by* the mind: *We* seem to have a measure of control over *our thinking*. How this occurs is not well understood, so it is often illustrated

[1] Wegner and Pennebaker report in 1993 that "the term *mental control* does not appear in searches of the psychological literature prior to 1987" (3). With the benefit of the Internet I did note some earlier mentions, albeit few. One such mention from around the same time period is the finding of a link between mental control and depression, also explored in a later study (Strömgren 1977; Breslow, Kocsis, and Belkin 1980). As it is used in these other studies, mental control "is probably most clearly associated with aspects of memory functioning which concern attending to the task" (Breslow, Kocsis, and Belkin 1980, 542).

[2] Of the top ten articles yielded by a search in Google Scholar for "mental control" in May 2022, eight include Wegner as an author.

Table 1 Mental control is control of the mind by the mind and can be contrasted with nearby forms of control involving the mind and body

		Control by:	
		the mind	external factors
Control of:	the mind	mental control	mind control
	behavior	action control	sensorimotor control

through everyday examples: "Mental control occurs when people suppress a thought, concentrate on a sensation, inhibit an emotion, maintain a mood, stir up a desire, squelch a craving, or otherwise exert influence on their own mental states" (Wegner and Pennebaker 1993, 1). Recall the ping-pong example from Section 1: In order to play a better game, we might suppress distracting thoughts, concentrate on the ball in flight, inhibit a fear of failure, maintain a sense of calm, stir up the desire to win, or squelch a craving for tasty snacks nearby. All of these would count as forms of mental control.

Excluded is control of other functions by the mind, such as behavior.[3] While one might reasonably interpret the term this way, "the current renaissance in the area of mental control has largely been fueled by those who have taken the meaning in the narrow sense – that is, as an attempt to control thoughts and thought processes" (Schneider 1993, 13). Likewise excluded is control of the mind by external factors, such as other agents. Some explicitly contrast mental control with, for instance, "sensorimotor control": While mental control is "induced intentionally," sensorimotor control is "perceptually (externally) induced and is controlled by recent environmental stimuli" (Schack and Frank 2021, 530). Thus, while the words "mental" and "control" are broad, and one might be tempted to judge their combination to include any place where the two overlap, specialist use of the term is more restricted (see Table 1). Given the potential for confusion, "mental control" is sometimes used with modifying language, such as "personal-level mental control" (Papineau 2015), which can help to distinguish it from these other forms of control.

In their introduction to *The Handbook of Mental Control* (1993), Wegner and Pennebaker explicitly tie mental control to attention: "A view of attention as the central faculty of mental control is part of much contemporary work in cognitive

[3] This includes work on the concept of behavioral control, which is a separate topic that has received much more attention in philosophy (see, e.g., Shepherd 2014; Christensen, Sutton, and McIlwain 2016). This distinction is analogous to that between mental action and bodily action, which are sometimes treated as separate topics (see, e.g., Fiebich and Michael 2015).

psychology" (4).[4] Mirroring typical accounts of attention, they contrast mental control with automaticity, in which case the mind is "beyond our control" (4). While they see this divide as one that overlaps with thorny issues, such as the problem of free will, they take it to be tractable through an operational approach. They thus argue that mental control can be operationally separated from "unintended but appropriate mental activity" using a range of experimental techniques (8). These experimental techniques include, among others, (a) using parallel tasks to determine the use of cognitive resources thought to be limited in mental control; (b) observing somatic markers, such as behavioral freezing in the face of mental conflict, typically associated with an attempt at mental control; and (c) collecting self-reports from participants as to whether they are using mental control.

For the most part, scientists now study mental control through what they call "executive control." Broadly speaking, "executive" functions support the coordination and management of other mental functions; they include "inhibition . . . working memory, and cognitive flexibility" (Diamond 2013, 135). Executive *control* is taken to include a subset of the executive functions – those that direct other mental functions in the service of tasks or goals: "It is needed to overcome local considerations, plan and orchestrate complex sequences of behavior, and prioritize goals and subgoals" (Miller and Wallis 2009, 99). In some cases, executive control is treated as control by a "central executive," or a unitary system of control, explicitly making use of a homunculus-like figure (Baddeley 1996, 1998; see subsection 3.2 for discussion of the homunculus). The more popular contemporary perspective is to see executive control as distributed across several systems (see, e.g., Logie 2016). In both cases, "prefrontal cortex" is likely to play a crucial role, because prefrontal cortex is domain general, and thus an ideal candidate for controlling domain-specific neural networks (Stuss and Alexander 2000; Figure 1e–i). As Buehler (2018) puts it, "Central executive control is likely implemented through the prefrontal cortex's modulation of neural activity in domain specific neural networks" (1974).

Buehler (2018) identifies three core functions of executive control: switching, maintenance, and inhibition. Prominent behavioral tests used to explore these functions include the Wisconsin Card Sorting Test (WCST) and Stroop Task (see, e.g., Derrfuss et al. 2005). The WCST involves the participant choosing one out of four possible cards to match a card that is provided. While the participant is not provided with the underlying rule to help them

[4] Schneider, for instance, divides mental control into three types, with attention supporting the most central type: "We control our thoughts by manipulating our sense organs (external control), by voluntary focus of attention (direct control), and by initiating trains of thought (indirect control)" (1993, 28).

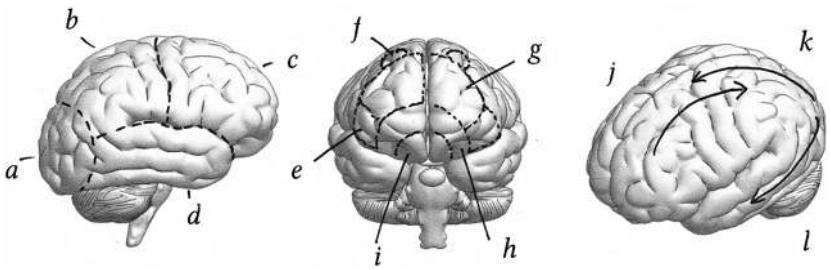

Figure 1 Views of the brain and select brain areas. On the far left is a view of the brain from the right, including (a) occipital or visual cortex, (b) parietal cortex, (c) frontal cortex (including prefrontal areas), and (d) temporal cortex. In the middle is a view of the brain from the front, including (e) ventral lateral prefrontal cortex (right hemisphere is indicated, but each area occurs in both hemispheres), (f) frontal eye fields (FEF), (g) dorsolateral prefrontal cortex (dlPFC), (h) orbitofrontal cortex, and (i) ventral medial prefrontal cortex (vmPFC). On the far right is a view of the brain from the top left, including (j) frontoparietal feedback, (k) dorsal visual pathway (feedforward), and (l) ventral visual pathway (feedforward).

determine the correct card, there are three possible rules: matching the color of the objects on the cards, matching the shape of those objects, and matching the number of objects. The participants can determine the rule through feedback, but the rule may change over a session. The WCST measures how well the participant identifies, maintains, and switches between these rules. The Stroop Task, on the other hand, requires the participant to report on one of the features of a stimulus while ignoring a more salient feature. Typically, a participant will be asked to report on the color of a word while the word itself is a conflicting color term (e.g. the word "blue" in red ink would require the participant to say "red"). The Stroop Task thus measures how well the participant is able to inhibit the salient feature in service of the task at hand. These and other behavioral tests have allowed scientists to determine the scope and limits of executive control, as well as its neurobiological basis.

While there have been some successes in the experimental domain, philosophical work is needed: "Mental control has proven to be an elusive target for scientific analysis, embedded as it is in sticky philosophical problems such as volition and dualism" (Vallacher 1993, 444). Philosophers have, unfortunately, published little on mental control.[5] As is mentioned in Section 1, mental control

[5] The same is true for executive control. As Buehler puts it: "Surprisingly, philosophers have not much engaged with the scientific literature on executive functioning" (2018, 1969). "Executive

includes within it the more popular topic of self-control, which appears in sixty-eight separate entries of *The Stanford Encyclopedia of Philosophy* (*SEP*). The term "mental control" shows up only once, in a single entry https://plato.stanford.edu/archives/sum2020/entries/testimony-episprob/. It is thus natural to start a philosophical discussion of mental control with a brief discussion of self-control.

While self-control is often described in terms of overcoming temptation, a broad definition might note the "competition between smaller sooner (SS) rewards and larger later (LL) rewards . . . [such that] the presence of SS rewards places the agent in a situation of conflict and requires the exercise of self-control" (Kennett and Wolfendale 2019, 34). A classic example is that of the marshmallow task: children are asked which of two options they prefer, one or two marshmallows, and then told that if (and only if) they wait they can have their preferred option. That is, the child can only have the preferred two marshmallows if they are able to refrain from eating the single marshmallow placed in front of them (Shoda, Mischel, and Peake 1990). Famously, the longer a child is able to wait, the more academic success they exhibit in adolescence: "Coherent patterns of statistically significant correlations were found between seconds of delay time in such conditions in preschool and cognitive and academic competence and ability to cope with frustration and stress in adolescence" (Shoda, Mischel, and Peake 1990, 978). These findings were later replicated and extended, leading to the insight that self-control is a "critical early capacity" (Watts, Duncan, and Quan 2018).[6]

Kennett and Wolfendale go beyond the above definition to provide a taxonomy of self-control, noting the different *methods* and *levels* of self-control (2019). In terms of methods, they separate "exerting willpower" from "implementing strategies" as well as "synchronic" from "diachronic" self-control. Take the example of waiting for two marshmallows: one might exert willpower to help one wait or one might implement a strategy, such as looking away from the single marshmallow. Willpower is necessary to overcome a temptation at the time it is given (synchronic) in lieu of a strategy, but strategies can be used if planned in advance of the temptation or for ongoing temptations (diachronic). As willpower is limited, the bulk of self-control depends on strategies enacted diachronically (see, e.g., Bermúdez 2021).

control" shows up in three entries of the *SEP*: "David Hartley," "Neuroethics," and "Philosophical Psychopathology," each time as a single mention.

[6] Worth noting are the structural and social features found to predict success on the marshmallow task, including caregiver reliability, which may or may not predict success through an impact on mental control (Duckworth, Tsukayama, and Kirby 2013; Kidd, Palmeri, and Aslin 2013; Lamm et al. 2018; Michaelson and Munakata 2020). Moreover, the task has not been found to predict other, later outcomes, such as "mid-life capital" (Benjamin et al. 2020).

In terms of levels, Kennett and Wolfendale separate "intentional," "instrumental," and "normative" self-control, with intentional serving as the lowest level. Intentional control requires only the involvement of an intention, instrumental control requires the ability to achieve that intention through intermediate steps, and normative control requires that the entire process is sensitive to considered judgments and values. Kennett and Wolfendale argue that to be an agent at all one would need intentional control, to be an agent over time one would need instrumental control, and to flourish as an agent one would need normative control. They find the importance of self-control to be best displayed in this highest level: "Self-control is a necessary condition of access to a variety of goods that help constitute a life as meaningful, as flourishing, and, importantly, as one's own" (Kennett and Wolfendale 2019, 38).

One might construe self-control through the language of "first-order" and "higher-order" desires (see, e.g., Frankfurt 1988). As Holton and Shute put it: "Accounts agree that self-control consists in a particular kind of control over one's actions – in each case the obvious contrast is with actions that are driven purely by one's (first-order) desires" (2007, 51). Higher-order desires are desires *about* first-order desires. Desiring the marshmallow, for instance, is a first-order desire, while wanting to squelch the desire for the marshmallow is a second-order desire. In self-control, our first-order desires are in conflict with higher-order desires, and we aim to resolve that conflict in favor of the higher-order desires. We might see this as *self-shaping*, since the process ultimately alters the relative strength of our desires. Kane (1999) describes, for example, a case in which a businesswoman decides to stop and help a victim of assault rather than continue on to an important meeting. She has competing desires – to help the victim and to continue on to her meeting – and resolves them in favor of the former.[7] This is a case, Kane argues, in which the businesswoman's effective desire (her "will") was not already determined, but she helped determine it through her action, making it more likely in the future that she would stop to help a victim in the face of a conflicting desire (Kane 1999, 224).

How do these considerations intersect with mental control? While there is significant overlap between self-control and mental control, there is also difference. Take the list provided by Wegner, above: "suppress a thought," "inhibit an emotion," and "squelch a craving" are all strongly associated with self-control. Yet, the other items on the list are less clearly associated. When we "concentrate

[7] While it may not be obvious from the description here, I am treating the desire to continue on to the meeting as a first-order desire, and the desire to stop to help the victim as a higher-order desire (a desire to overcome this first-order desire). This fits Kane's narrative of the businesswoman overcoming temptation in favor of her "moral conscience" (Kane 1999, 225).

Table 2 Points of contrast between self-control and mental control

	Self-control	Mental Control
Source	higher-order desires	the mind
Target	first-order desires	any mental state or function
Method	willpower or strategy	switching, maintenance, or inhibition
Outcome	self-shaping (change to the desire system)	mind shaping (change to any mental disposition)

on a sensation" are we necessarily overcoming temptation? How about when we "maintain a mood" or "stir up a desire"? The language of "temptation" comes from the close ties between philosophical research on self-control and discussions of normativity; mental control is disconnected from those considerations. Further, while self-control concerns the conflict between smaller sooner and larger later rewards, mental control may also include conflict within these levels (between different types of smaller sooner rewards or different types of larger later rewards). Finally, mental control can result in long-term changes, but these need not be to our desires. If I concentrate on the ping-pong ball, I am not necessarily shaping my desires but rather my perceptual and motor capacities. Thus, mental control can be seen as a broader umbrella under which self-control falls, with self-control specifically describing cases of conflict resolution between higher-order and first-order desires, or larger later and smaller sooner rewards, that results in some degree of self-shaping or a change to the desire system (see Table 2).[8]

As mentioned earlier, attention is often cited as the key to mental control. Recall that mental control is studied scientifically through executive control, sometimes understood as control by a central executive. In fact, the "central executive" initially proposed by Baddeley and Hitch (1974) was just attention, as can still be seen in most descriptions of the model (see, e.g., Parkin 1998, 518). While Baddeley has since argued that we should break up the central executive into subcomponents, "including dual task performance, attentional focusing, attention switching, and interfacing with [long-term memory],"

[8] I take these considerations to be separable from temporal ones. Sripada (2021), for example, has argued that self-control occurs over an extended period and is built up from short duration cognitive control states. The account above covers both short-duration mental control (e.g. a single instance of Stroop inhibition) as well as long-duration mental control (e.g. task maintenance and switching in WCST).

attention still plays the primary role (1998, 525). As described by Buehler, "Sometimes this aspect of executive functioning is called 'executive attention'" (2018, 1970). In the next section I will review work that ties mental control to attention, as well as work that challenges this perspective on mental control.

3 Attention and Mental Control

As we saw in the previous section, attention comes up frequently in discussions on mental control. What is attention? Like mental control, its meaning is at once familiar and elusive, with many competing accounts of the phenomenon within philosophy.[9] Consider the example of attending to a dinner guest: we might "pay attention" to the guest one moment, while our attention is "grabbed" by the doorbell the next. That attention can be "paid" and "grabbed" indicates that it has something to do with mental resources, and, in fact, a commonly cited feature of attention is that it is resource limited.[10] Attention is typically described as the process of *prioritizing* these resources.[11] One might say that attention is the prioritization of some mental processes over others, often resulting in the selection of one or more mental processes at the expense of others.

Importantly, attention is not the only selective process in the brain. The selection of particular light wavelengths in the eye is not attention, for example (see, e.g., Mariani 1984). The prioritizing work of attention is best seen as occurring in concert with other forms of prioritization and selection, such as the prioritization of naturally salient stimuli. That is, stimuli that are strong smelling, bright, fast moving, or loud "grab" one's attention due to processes other than attention that prioritize those stimuli (see, e.g., Reynolds and Desimone 2003). Attention theorists are divided on how exactly to separate attention from these other selective processes, but most philosophers describe attention as a "personal" or "subject-level" phenomenon – a phenomenon that essentially involves the person or subject, rather than a proper part of the person or subject – with substantial variance on how to interpret these terms. For example, while Wu (2011) holds that "the subject level process notion should be understood as *selection for action*" (97), Jennings (2012) holds that attention

[9] These range from Mole's (2011) "adverbial" view – that attention is not a process but a way that mental processes might proceed – to Prinz's (2011) "AIR" view that "attention is a process by which perceptual representations become available to working memory" (186).

[10] One potential neural resource is glucose, a general resource for demanding cognitive tasks (Ampel, Muraven, and McNay 2018).

[11] Take, for instance, Buehler's (2019) definition: "Attending to something, in its broadest, most widely accepted sense, involves the directing of processing resources towards that thing. Directing processing resources towards something normally results in faster, more accurate processing of it" (2123).

is "a process of mental selection that is *within the control of the subject*," and need not concern action (535; emphasis mine).[12]

What is the role of attention in mental control? As noted in section 2, many see attention as the central mechanism of mental control: "If we have any mental control, it seems to start with our ability to influence our focus of attention . . . To control our movements, our emotions, our addictions, our desires, our diets, or anything else, we must first control our attention" (Wegner 1988, 683). This reliance on attention has been described extensively in the case of self-control, which is part of mental control.[13] In the marshmallow task, for example, one strategy for overcoming temptation is to simply look away from the marshmallow, a form of "overt" attention (as opposed to "covert" attention, in which the eyes need not move). To move the eyes away from the marshmallow is to prioritize other stimuli over the marshmallow, potentially resulting in the selection of a thought process other than that of the tempting marshmallow. Further, it is this very skill of directing attention away from the distracting marshmallow that is thought to support subsequent academic success (Shoda, Mischel, and Peake 1990, 985). Later in development, self-control can be achieved either by directing attention away from temptations or by directing greater attention toward one's goals (Herdova 2017).

While attention supports self-control by altering the relative prioritization of otherwise salient stimuli, such as temptations, mental control goes beyond overcoming salient stimuli. Recall the ping-pong example from the prologue: mental control helps one to overcome distractions and temptations, but it also helps one to focus, to motivate, and to align mental processes, all of which can occur without the presence of distractions or temptations. Is attention required for all of these functions of mental control, or can mental control occur in some cases without the benefit of attention? Some separate control that depends on attention from "automatic" control. Arango-Muñoz and Bermúdez (2018), for example, distinguish two types of control: "One of them is reflective control – the often slow, effortful top-down control that we exert by recruiting working memory in novel or attention-demanding tasks. The other is intuitive or automatic control – the fast, rather effortless and intuitive control that we exert automatically, without the intervention of working memory" (90). Take, for instance, the direct selection of a response based on a stimulus, such as taking

[12] Within the sciences the issue shows up, for example, in clashes about whether attention modulates primary visual cortex (see, e.g., Mangun 1995 versus Posner and Gilbert 1999).

[13] Alexander (1910) provides an early instance of this claim from within philosophy: "If it should be admitted that there is direct voluntary control of the feelings and emotions, such restraint would be primarily due to attention" (291).