

# Attention and Voluntariness in the Wandering Mind\*

[Penultimate draft. Final version forthcoming in *The Journal of Philosophy*]

*Abstract.* Mind wandering has been a target of a fast-expanding area of research in cognitive science and philosophy. One of the central puzzles that researchers have been grappling with is whether this mental process should be thought of as passive or active in nature. Intuitively, a wandering mind seems passive but mounting empirical evidence suggests otherwise. Irving (2021) defends a prominent account of mind wandering as unguided attention, which aims *inter alia* to resolve the puzzle. However, I present counterexamples that reveal Irving's account to be both too weak and too strong. I then develop the alternative proposal that (stated roughly) mind wandering consists in voluntarily passive attention. After unpacking and defending this idea, I show how it helps to eliminate the conflicting appearance of mind wandering as both passive and active.

## 1. INTRODUCTION

The following experience is surely not idiosyncratic to this writer. Attempting to decipher an assortment of cryptic remarks which, apparently in the eyes of the student who wrote them, pass for an argument, I find my thoughts have drifted towards seemingly random, more rewarding items: From, say, wishing that the weather tomorrow will be more agreeable, to counting down the days till my upcoming vacation, to wondering what the local cuisine there is like, to searching my memory for fresh ideas for the meal I'll be cooking tonight, etc. etc. The experience described is hopefully familiar to the reader first-hand. As a paradigmatic instance of *mind wandering*, it is certainly ubiquitous: An oft-quoted figure estimates that human minds spend between a third and a half of their waking time wandering (Killingsworth & Gilbert, 2010). But as any philosopher would tell you, neither first-person familiarity with, nor the prevalence of some phenomenon offer any promise of understanding its nature. Mind wandering (MW) is a case in point. An adequate account of this pattern of thought has so far eluded researchers, despite the tremendous amount of attention it has attracted. In the cognitive sciences, it is no exaggeration to say that work on

---

\* Comments and suggestions by Arnon Levy, Tom McClelland, Wayne Wu, and an anonymous referee for this journal greatly improved this manuscript. I am very grateful to all of them. Many thanks also to the constructive discussion of material from this paper by audiences at the 'Mental Action and Attention' workshop at the University of Oslo, and the annual conference of the Israeli Philosophical Association. Research for this paper was supported by the Israel Science Foundation (grant no. 381/23).

MW, understood as a unified theoretical construct, has exploded in the past 20-odd years. Meanwhile, philosophical interest in MW has also increased markedly in the past decade or so (although of course, broader interest in the stream of consciousness has much deeper historical roots).

One of the central puzzles that both philosophers and cognitive scientists have been grappling with is whether MW is a *passive* or *active* mental process (Arango-Muñoz & Bermúdez, 2021; Carruthers, 2015, ch. 5; Irving, 2016, 2021; Metzinger, 2013; Murray & Krasich, 2022). Intuitively, MW appears passive: When one's mind wanders, it does not feel like one is actively steering the direction of one's stream of thought, but rather that thoughts are occurring to one randomly or at any rate, without one's active involvement. We often *catch ourselves* mind-wandering, e.g. when we realize that we understand next to nothing of the text we have been reading. However, the apparent passivity of MW conflicts with another widely accepted datum, viz. that there is such a thing as *intentional* MW.

The notion of intentional MW may seem confused, perhaps even a contradiction in terms akin to “aiming to do some activity that is inherently aimless” (Murray & Kraisch, 2020: 436). Yet accumulating evidence suggests it is a real possibility. When prompted, subjects are able to distinguish intentional MW both from unintentional MW and from goal-directed thinking (Seli et al. 2017; Seli, Carriere & Smilek, 2015). Up to 40% of reports by subjects in the lab claim to record episodes of MW with intention. People likewise routinely self-report letting their minds wander on purpose in their daily lives (Kane et al., 2007). Researchers have pointed out the significance of distinguishing the intentional from the unintentional variety of MW, suggesting that each exhibits distinct behavioral and neural patterns, and is associated with different underlying mechanisms (for a review, see Seli et al., 2016).

But now the puzzle comes into view: If MW is sometimes intentional, how can it be essentially passive? The former property seems quintessentially agentive while the latter is associated with patienthood. Irving articulates a prominent theory of MW as ‘unguided attention’ (2021), which

attempts to dissolve this tension while respecting both seemingly conflicting observations. However, there are several problems with Irving's position, which I spell out below (§2). These problems motivate my proposal of an alternative, arguably superior account of MW which, similarly to Irving, can accommodate both the passivity of MW and its occasional intentionality. The account is couched in terms of *voluntary passive attention*. When one's mind wanders, one's attention is allocated in a voluntary passive fashion. The job of unpacking and defending this slogan is carried out in §3. To make the discussion concrete, that section traces briefly the contours of a general and familiar picture of agency as the exercise of a two-way power (§3.1). Having this picture in place sets a useful background against which the idea of MW as voluntary passive attention can be situated (§3.2). In the course of stating and defending the account, the discussion draws out its implications for the hotly debated questions of agency and intentionality within MW. Doing so promises to shed new light on the nature of a widely studied mental phenomenon, while at the same time bearing directly on the broader project of understanding the scope and limits of (mental) agency.

## 2. IRVING'S ACCOUNT: MIND WANDERING AS UNGUIDED ATTENTION

Speaking loosely, MW is dynamically characterized as a meandering train of thought that moves freely from one topic to another without a fixed aim. This observation leads Zachary Irving to propose a novel account of MW in terms of *unguided attention* (Irving 2021).<sup>1</sup> Characterizing the dynamics of MW as essentially meandering or unguided, Irving's account manages to exclude similar processes with a more targeted focus, such as rumination and absorption. But what exactly is the idea of unguided attention at play in his proposed definition?

'Guidance' in particular is a multi-purpose and sometimes obscure notion in philosophy. Irving offers the following explication of the term (quoted here from Irving & Thompson, 2019: 90):

---

<sup>1</sup> Irving's focus on the *dynamics* of MW allows him to circumvent counterexamples to common scientific alternatives, which define MW in *content*-related terms. For details, see Irving (2016), sect. 2.

To say that behavior is guided implies the following: Were one's behavior to go off course or deviate from some standard—as a result, for example, of interfering forces—one would alter that behavior in order to bring it back on course.

More specifically, when the behavior in question is attending, 'guidance' implies that any distraction one experiences would cause one to feel pulled back to the original object of one's attention. Conversely, if one's attention is unguided – that is, if one's mind is wandering – then “there is no information  $i$  such that, if  $\mathcal{A}$ 's attention were not focused on  $i$ , she would notice, feel discomfited by, and thereby be disposed to correct this fact.” (Irving, 2021: 623). Now Irving's account is surely correct in highlighting the connection between MW and attending; the former is in some sense the privation of the latter. However, as we shall see in a moment, the connection between the two is not to be understood in terms of (lack of) guidance. But before presenting counterexamples to Irving's proposal, notice a preliminary point about one of its key motivations. Irving insists (and this writer agrees) that MW should be understood as a passive phenomenon. It is for this reason that Irving finds fault in a Davidsonian causal approach to (mental) action, which he takes to be committed to misclassifying MW as active (Irving 2021: 617-619). For Davidsonians understand actions as events caused by the agent's 'pro-attitudes', and there is evidence of such causation taking place during episodes of MW: Having goals can apparently cause subjects' thoughts to meander to items related to those goals and to how they might be achieved. Thus for example, Morsella and colleagues (2010) told some of their subjects to be prepared to answer questions in the near future about American states, and then gave them a different task. They found that about 70% of their subjects' minds wandered to thoughts about American geography. (See also Klinger 2009; Mac Giolla et al. 2017).

However, despite the mental events in question being plausibly caused by intentions or belief-desire pairs, Davidsonians would not in fact regard these as cases of mental action. This is because the causation involved is *deviant*: The agents' goal is not brought about in the way characteristic of intentional action. The problem of causal deviance is a notoriously intractable obstacle for the

causal-reductive ‘standard story’ of action. Using causation ‘in the right way’ as a placeholder (Davidson 2001) flags the missing ingredient that would set apart genuine mental activity from deviantly caused mental events. For our purposes, the upshot is that mere causation by executive attitudes is actually compatible with the passive nature of MW. The Davidsonian line cannot be impugned for yielding the wrong verdict here.<sup>2</sup>

Moving on from the considerations motivating Irving’s account to its extensional (in)adequacy: We shall now see that the guidance-based account Irving proposes is both too strong, ruling out genuine cases of MW; and too weak, ruling in some cases that are not instances of MW.

Start with the charge that the account is too weak, which is illustrated by the following vignette. You are watching a mildly interesting TV show while running on the treadmill at the gym. The show does enough to keep your attention focused on the screen—you are a captive audience, after all—but just barely. Were the onset of any other salient stimulus to occur, your attention would readily shift to *it* and you would not feel discomfited, pulled back to the show, or disposed to ignore the distraction. As it happens, no such onset occurs and you watch the show till it ends. In the scenario just described, your mind does not wander. Throughout the experience, you remain completely focused on just one stimulus, viz. the TV show. But your attention is unguided in Irving’s sense. This is because the counterfactual conditional for guided attention is not satisfied, as there are plenty of distractions such that, were they to appear, would cause your attention to shift towards them without you feeling pulled back to the show in any way.

The source of the problem for Irving’s account clearly has to do with the counterfactual form of the condition for guidance it proposes (a more precise diagnosis will be offered shortly). One immediate reaction to the problem is therefore to suggest discarding the counterfactual formulation, construing guidance instead as conditional on the subject *actually* overcoming potential distractions. But this tack is hopeless, for the simple reason that attention may remain focused, hence guided, even in the absence of any competing stimuli actually occurring. A

---

<sup>2</sup> Many thanks to an anonymous reviewer for discussion here.

counterfactual formulation seems indispensable. However, there may be a better response available to Irving, which shows that the case does not in fact seriously damage his account. For as described, the scenario may seem to turn on assumptions that render it far-fetched and so not particularly worrisome. To see this, consider that the TV show must be salient enough to attract your full attention throughout but *not* salient enough to sustain your attention in the face of distractions. Satisfying both these requirements at once represents a tight rope, it might be claimed, which could only be walked if, as a matter of incredible coincidence, no distracting stimuli happens to present itself during your workout. But such absence of distractions over an extended period of time is so rare as to make the option of biting the bullet here not particularly damaging. Perhaps, that is, Irving could maintain that your mind is indeed wandering in this limiting case without his account losing much credibility as a result.

So the rejoinder goes.<sup>3</sup> A couple of points can be made in reply. First, the scenario described need not be so rare as it is made out to be. Suppose for example that you enter the gym just as it opens at 6am on a Sunday morning, and are the only person there. You decide to start with a relatively short session of (say) 10 minutes. Under such circumstances, it need not take a rare coincidence for there to be nothing to distract you from the TV throughout your (extended, but not particularly long) 10-minute workout. Second, even supposing that such cases *are* rare does not deflate the objection. For the aim in raising it was never to conclusively refute Irving's theory but rather to identify a weakness in it which, especially when combined with the further weakness exposed below, supports exploring potentially superior alternatives. An independently plausible account that does not suffer from this (or comparable) flaws is surely a welcome prospect even if (contra the first reply) the flaw in question manifests itself only infrequently.

Irving's account is susceptible to counterexamples such as the above because it commits what Robert Shope calls 'The conditional fallacy in contemporary philosophy' (Shope, 1978). This is the fallacy of providing a conditional analysis *C* of phenomenon *P* while ignoring the fact that whether

---

<sup>3</sup> I am grateful to an anonymous reviewer for pressing the rejoinder and getting me to consider it.

or not *P* obtains depends in part on whether or not the *antecedent* of *C* obtains, and not just on whether the entire conditional holds. The fallacy crops up in various contexts where conditional analyses have been proposed, most notably in the voluminous literature on conditional analyses of *dispositions*.<sup>4</sup> In the present context, the fallacy is exposed by the fact that your attention remaining guided depends in part on there *actually* being nothing to distract you from the TV show – not just on whether, if something *were* there to distract you, your attention would remain guided.

Remarks by Irving (2021, §IV.3) may suggest a way for him to handle the counterexample. Briefly, in the passage in question, Irving appeals to what he dubs the ‘categorical basis’ of a disposition, i.e. “the set of properties that give rise to that disposition under normal conditions” (634). When conditions are normal, the categorical basis of guided attention is equivalent to Irving’s counterfactual analysis. But under abnormal conditions, attention may be guided even if the counterfactual is not satisfied, so long as its categorical basis is tokened. Perhaps, then, Irving’s account can be saved by pointing out that, given the abnormal conditions of the gym scenario, the failure of the counterfactual does *not* imply that your mind is wandering after all. Guided attention is guaranteed in virtue of the categorical basis of guidance obtaining.

There is no space here to carefully set out and assess the response just sketched. But we may note the formidable obstacles facing any attempt to develop it. Notice first that invoking the categorical basis seems to undercut Irving’s initial appeal to the counterfactual condition for guidance. It turns out that, at least in some cases, it is not the disposition to refocus attention that blocks one’s mind from wandering. This marks a significant departure from Irving’s account as stated. As such, it would require spelling out more carefully the nature and functionality of the mechanism that *does* guarantee the maintenance of guidance – viz, the guidance mechanism constituting the categorical basis of the disposition. (For some initial remarks explicating the nature of this mechanism in the case of intellectual attention, see Irving 2021: 634-6.) Furthermore, assuming that the dispositional condition is not left out entirely of the revised account, it would

---

<sup>4</sup> Counterexamples of the sort discussed in the text to the so-called ‘simple conditional analysis’ of dispositions are provided e.g. Martin (1994) and Lewis (1999).

need to be made clear when the condition does and doesn't hold. That line is marked by the 'under normal conditions' clause in Irving's definition of the categorical basis above, which states when the disposition comes apart from its categorical basis. But this move gives rise to at least two serious problems. First, it is notoriously difficult to provide an adequate specification of normal conditions for *ceteris paribus* clauses. (For the claim that this cannot be done for conditional analyses of dispositions without rendering the conditional vacuously true, see for example Bird, 1998; Mumford, 2001; Fara, 2005; and Hauska, 2008). Second, even if an adequate formulation can be supplied, it is hard to see how it could block the present counterexample, whose circumstances appear perfectly normal; to treat them otherwise smacks of an *ad hoc* move. But then, if circumstances are indeed normal in the gym scenario, the counterfactual condition should be operative.

Moving on, Irving's account of MW faces, as noted briefly above, the further objection that it is too strong. While the gym scenario demonstrates that the proposed account rules in some cases it should not, the following scenario illustrates the converse charge, that it rules *out* some cases it should not. Imagine reading a novel and being thoroughly engrossed in it. At one point, the detailed culinary scenes trigger pleasant memories from an exquisite meal you once had while on vacation. The memory leads into imagining yourself throwing a dinner party and recreating some of the dishes you so enjoyed, which in turn flows into idle fantasies about becoming a celebrity chef.<sup>5</sup> At various points throughout this experience, you feel pulled back to reading the novel yet you are unable to resist indulging in your flitting train of thought. Here, it seems intuitively clear that there are many moments in which your mind wanders; but your attention is guided in the relevant sense.

It might seem like there is a quick fix to this problem. Instead of focusing on the feelings, dispositions, and intentional states of the subject whose mind wanders, perhaps the account should be couched in terms of actual behavior. That is, rather than characterizing the subject whose

---

<sup>5</sup> Note that this scenario is entirely realistic, indeed ubiquitous; MW during reading is rife (for a review, see Smallwood et al. 2007).



attention is guided as someone who “would notice, feel discomfited by, and thereby be disposed to correct” the fact that her thoughts are drifting, we should characterize her as *actually directing* her thoughts back on course. Irving himself explicitly rejects this way of understanding guidance. Citing an example of a waiter who is guided to stack the pile of plates he is carrying, he says that “[g]uidance does not require that the waiter always balances the plates perfectly; this is too much to ask”. As Irving explains, his understanding of guided attention in terms of motivationally efficacious felt distractions “flows from a general feature of agency”, according to which “what matters is how agents experience and respond to errors” (2021: 633).

Now, Irving may be warranted in refusing to treat ‘guidance’ as a success term for the reason that this would be “too much to ask”. Perhaps one’s attention may indeed be guided towards *S* even when it is not constantly focused on *S*. Be that as it may, the above counterexample brings out that, intuitively speaking, feeling distracted and being poised to refocus attention are *not* sufficient to characterize a non-wandering mind. While reading the novel, your mind is wandering away from your reading even though you are fully aware of this fact and doing your best to change course. Whether this failure of Irving’s account is ultimately down to how it proposes to understand guidance, or alternatively to its very appeal to guidance as the *analysans* of MW, is an (interesting) further question.

We have seen that Irving’s account falls short of adequately explaining the nature of MW, and consequently also the sense in which MW can be both passive and intentional at the same time. The following section proposes a more promising way of doing so.

### **3. MIND WANDERING AS VOLUNTARILY PASSIVE ATTENTION**

In slogan form, the idea defended here is that mind wandering can be understood as the allocation of attention in a voluntarily passive fashion. When one’s mind wanders, one lets in a sense one’s attention be directed by external forces. The idea needs to be unpacked and defended, most of which will be done in §3.2. Before then, §3.1 spells out some background conceptual

distinctions that play essential roles in the subsequent account. The conceptual groundwork laid out in §3.1 – specifically, the conception of action as the exercise of a two-way power – is familiar though not uncontroversial, but no attempt will be made to robustly defend it as that would take us too far afield. Nor is a defense of this sort strictly necessary, since as will become clear, the ensuing definition of MW may in principle be accepted on its own terms. What follows, then, is a framework for understanding some key concepts (with references to extant defenses, including by this writer), with an eye to demonstrating how this framework gives rise to a promising account of MW.

### 3.1 *Background conceptual landscape*

Start with the distinction between activity and passivity. Seeing as the definition being developed characterizes MW as passive, the operative distinction from active must be clarified. The approach endorsed here sees activity (or agency<sup>6</sup>) as the exercise of a two-way power to  $V$  and to refrain from  $V$ -ing. Having the power to  $V$  is understood as having the ability and the opportunity to  $V$  (Kenny, 1975). The same idea can be expressed in the ideology of ‘general’ and ‘specific’ abilities (Honoré, 1964) by saying that an agent has the power to  $V$  just in case she has the specific ability to  $V$ , i.e. she is *in a position* to  $V$  (unlike a piano player held captive in the basement, who has the *general* but not the specific ability to play the piano). The approach of understanding action as the exercise of a two-way power has a distinguished pedigree going back to Aristotle, who in the *Nicomachean Ethics* says that “where it is in our power to act it is also in our power not to act” (*NE* 1113b6)<sup>7</sup>. Versions of this idea have also been defended by Aquinas in *Summa Theologiae*, as well as by Reid (1788).<sup>8</sup> The variations between these writers’ more specific ways of deploying the two-way power conception will not preoccupy us here, as the aim is merely to have in place a basic sketch of the view that we may build on later.

---

<sup>6</sup> Whether or not the two-way power approach to explaining action outlined in the text extends also to explain *agency* depends in part on the delicate issue of whether there can be passive manifestations of agency: potential examples include standing still and not greeting someone you recognize.

<sup>7</sup> See also *Eudemian Ethics* 1223a4–7.

<sup>8</sup> More contemporary defenses include Hart (1968), Geach (2000), Steward (2012), and Alvarez (2013), among others.

Action, then, is understood here as the exercise of a two-way power to  $V$ . (Notice the emphasis on *exercising* the power, over and above merely having it. The latter mode will become important when we turn to explicating the notion of voluntariness, next; for now, the exercise of a power should be kept firmly apart from having the power.) The next issue to address is the appropriate substitutions for the schematic  $V$  (generalizing from any specific answers as may be relevant on given occasions, e.g. ‘wash the dishes’ or ‘eat the apple’). The power to do *what*, more exactly, is action the exercise of? In the case of bodily actions, a plausible answer is ‘move one’s body.’ Whatever else one is doing when acting, one such thing is surely moving one’s body (most often, the limbs). Cast in terms of the present approach, the idea would be that in acting, one exercises one’s power to move one’s body. Another common idea is that the power of agency is the power to cause change (Hyman 2015, ch. 2). However, neither suggestion, certainly not the former, is a viable option when it comes to capturing also *mental* action (see **AUTHOR’S PAPERS**), which is more pertinent for present purposes than bodily action. The central power exercised by agents when they perform mental actions is not to move their bodies (even though certain bodily motions, primarily in the brain, are plausibly an *upshot* of agents’ exercising their agentive power). A much more plausible thought is that mental agents primarily exercise the power to *direct their attention*, or more simply to give their attention to someone or something – to choose what to think about, who to look at, listen to, etc.

This is not the place for an elaborate defense of the idea that the central power of mental agency is the power to attend (for that, go to [**AUTHOR’S PAPER**]). But two principal considerations motivating this line of thought can be cited. The first consideration comes from noticing just how pervasive attention actually is in our mental lives. Virtually every mental act one can think of involves attention: Doing a calculation in one’s head, judging, recollecting, reciting poetry, hypothesizing, visualizing a scene, reading, daydreaming, deliberating, ... Each of these act-types (and a host of others that could readily be added to extend the list) seem clearly to entail attention. It would be impossible to perform any of them without paying at least some degree of attention to the object one is acting upon (reciting *R*, judging *that p*, etc. etc.). This raises the suspicion that

there is a constitutive connection between acting mentally and attending. Second, each of the sophisticated types of mental action, examples of which were just cited – reciting poetry, visualizing a scene, and so on – ultimately involve some form of the more basic perceptual or intellectual modes, such as looking, listening, smelling, or concentrating in thought. And it is hard to make sense of the idea that these basic forms of mental action could be performed without paying any attention whatever. Allan White makes a similar point (1964: 7-8):

Because we focus on what is perceptible by using the appropriate sense-faculty and on what is intelligible by making it the object of our thinking, we can specify the general notion of *attention* in terms of these particular perceptual and intellectual activities [...] when we speak of attention being paid or given, drawn or attracted, it is basically some set of these perceptual and intellectual activities to which we refer.

This, again, seems to implicate the exercise of attention whenever mental action is performed.

Moving on, the next background distinction to be drawn is between the voluntary and the involuntary. What is most important to note for our purposes is that this distinction cuts across the one discussed previously between the active and the passive, as argued forcefully by Hyman (2015, ch. 1). The point often goes unnoticed, owing to a strong tendency among philosophers to conflate the active/passive with the voluntary/involuntary. Here is Gilbert Ryle (1949: 73-4):

Very often we oppose things done voluntarily to things suffered under compulsion. Some soldiers are volunteers, other are conscripts; some yachtsmen go out to sea voluntarily, others are carried out to sea by the wind and tide. [...] So sometimes the question ‘Voluntary or involuntary?’ means ‘Did the person do it or was it done to him?’

Similarly, in his book *Elbow Room*, Dennett asks, “Are decisions voluntary? Or are they things that happen to us?” (1984: 78). But the contrast Dennett sets up with his question is specious, since something’s being voluntary does not exclude it happening to us or vice versa, as will now

be shown. In at least one sense of the term, ‘voluntary’ means what is allowed to happen; what one has the power to prevent but does not prevent. On this understanding of voluntariness, something’s being voluntary does *not* imply that one has chosen or even consented to it. Indeed, one may be strongly opposed to what is taking place, playing along only due to being compelled by a threat, and *still* it would count as voluntary in the intended sense. To illustrate, suppose you hand over your wallet to a robber at gunpoint. Your action counts as voluntary: You could have chosen to hold on to your wallet but did not (because the cost of doing so was prohibitive). There is arguably another sense of ‘voluntary’ on which your action is *involuntary*, precisely because it was performed under compulsion. Kenny recognizes the ambiguity of ‘voluntary’. In one sense of the term, he claims, ‘voluntary’ contrasts with “actions that are reflex, or somnambulistic, or the gasping admissions of an exhausted man broken by torture”, while in another sense ‘voluntary’ contrasts with such things as “the reluctant action of a man sealing up a wad of £10 notes and putting them in an envelope to send to a blackmailer.” (Kenny, 1982: 197-8; see also Hart & Honoré, 1985). It is the former sense that should be kept in mind for what follows.<sup>9</sup>

Talk of different senses of ‘voluntary’ should not mislead into thinking that the issue is at bottom about terminological choice. One is free of course to use one’s terms as one likes, and could decide to treat ‘voluntary’ as interchangeable with ‘active’. But if that usage is chosen, a different term would need to be introduced to designate the property of being something one allows to happen despite having the power to stop it from happening. *This* property certainly does not correspond to the concept ACTIVE as it is not exclusively a feature of actions. It applies also to various different events/processes in which one is acted upon as patient rather than acting as agent. Thus consider being carried voluntarily, falling asleep voluntarily, being kissed voluntarily, or undergoing a medical procedure voluntarily. These are all passive manifestations of

---

<sup>9</sup> Hyman (2015: §4.3) rejects the thought that ‘voluntary’ is ambiguous in the way tentatively suggested in the text. On his view, voluntariness is an ethical concept, which means roughly lack of ignorance and compulsion. I have reservations about Hyman’s position, but this dispute need not be settled here. If VOLUNTARINESS *does* turn out to be the concept Hyman describes, we could simply use a different term to pick out what I am referring to here as ‘voluntary’.

voluntariness. (There are also passive manifestations of *involuntariness*, as when one is being carried or kissed while trying to resist, falling asleep despite trying to stay awake, or being forced to undergo the procedure.) Running together the active with the voluntary thus crucially makes the category of voluntary passivity drop out of view.

Tying this observation with the framework sketched above for understanding the distinction between active and passive, we may note that when passive, one does not *exercise* one's agentic two-way power to *V* and to refrain from *V*-ing; exercising the power is the mark of activity. But while the non-exercise of the power is common both to voluntary and to involuntary instances of passivity, there is a crucial difference between the two when it comes to *having* the power. When involuntarily passive, we lack our agentic power: If we are being carried despite not allowing ourselves to be carried, that must be because we are unable to stop this from happening. But when voluntarily passive, we do have the power though we choose not to exercise it, delegating control instead to some force(s) outside ourselves.

### 3.2 *Defining mind wandering*

We are now finally in a position to put the conceptual machinery introduced in §3.1 to work in defining MW. The general idea, to repeat, is that MW may be understood as a voluntarily passive form of attention. And we now have a clearer picture of what exactly this means. Being *passive*, subjects whose minds are wandering do not exercise their agentic mental power to direct their attention. Being *voluntarily* passive, mind-wandering subjects do not lose the power to direct their attention, but rather delegate it to external forces. They allow external forces to control the direction of their attention, letting them take over even though it remains in their power to prevent or stop this and bring the episode of MW to a close.

Some would be inclined to reject the idea that MW is subject to voluntary control. Thomas Metzinger (2013, 2015) for example, argues that mind-wanderers lack what he dubs 'veto control',

viz. the ability to inhibit or truncate some behavior after it has been initiated. This is because they lack ‘meta-awareness’, i.e. conscious awareness of undergoing some mental experience. When our mind is wandering, according to Metzinger, we lack meta-awareness that it is wandering; consequently, we also lack veto control over whether it is wandering. Metzinger’s argument is problematic, however. First, his claim that lack of meta-awareness is essential to MW is not borne out by the evidence. Participants in Smallwood et al.’s (2007) study reported plenty of cases of ‘tuning out’, where one is aware that one’s mind is wandering while this is taking place.<sup>10</sup>

Second, and more fundamentally, it is not clear why meta-awareness should be thought necessary for veto control in the first place. In order to inhibit a process *F*, one must no doubt know that *F* is ongoing. But crucially, one need not know it under *that* description necessarily. Put differently, awareness *de re* of *F* is sufficient for veto control, even in the absence of awareness *de dicto*. An example will help to illustrate. Suppose I am offending you by sticking my tongue out. I have the ability to stop offending you even if I do not know this is what I am doing, simply by pulling my tongue back in. I have veto control over an offensive act of mine (in part) because I am aware of doing something, which is (or constitutes) an offensive act of mine. The lesson carries over to episodes of mind wandering. Suppose that while reading, my mind wanders to thoughts about getting coffee. I have the ability to stop my mind wandering even if I do not know this is going on, simply by returning to my reading. I have veto control over my mind wandering (in part) because I am aware of something going on – i.e. thoughts about getting coffee – which is (or constitutes) my mind wandering.

It should be clear by now how the account of MW as voluntarily passive attention bears on the question of mental agency. The way in which the account fits with the general schema for understanding mental activity and passivity discussed above confirms the intuitive appearance that MW is passive. But endorsing the account by no means *commits* one to the two-way power approach. All the account requires is acknowledging that MW subjects are able to steer the

---

<sup>10</sup> Cf. Irving & Glasser (2020a, §2.2).

direction of their thoughts, even though they do not actually do so. This much is compatible also with, for example, a causal-reductive take on agency along standard Davidsonian lines. A subject whose mind is wandering would be considered mentally passive by Davidsonians, roughly insofar as her thoughts are not brought about by an intention (or a belief-desire pair) of hers – at least not in the right way. So long as the Davidsonian allows that, in addition to being passive, the subject has the ability to turn her thoughts away from their meandering path, she is free to embrace the account of MW proposed.

Further, the account finds adequate space for attention. As noted, the connection is intuitively clear: MW is in some sense the privation of (sustained, focused) attention. Looked at it one way, MW is inverse to active attending, in which the agent chooses the objects of her attention and her mind remains focused on them. But at the same time, MW is also intuitively distinct from ‘bottom-up’ attentional capture, in which one’s attention is grabbed exogenously by e.g. a loud siren in the street or the sound of one’s name being spoken. And the account proposed here nicely captures this relational structure by situating MW in the middle ground demarcated by voluntary passive attention – in between *involuntary passivity*, of which attentional capture is one instance, and *voluntary activity*, of which active attending is one instance.

One might protest that attentional passivity as understood here overgeneralizes beyond MW, finding application also in cases of *skilled attention*. The expert driver attends at various stages to such things as her blind-spots, objects in her mirror, an oncoming driver, etc. But her attention is arguably automatic and in any event, she does not choose to attend to the particular objects she actually attends to: Being skilled at driving, her attention is allocated without her making such (person-level) choices. How, then, does skilled attention differ from MW on the proposed account? The answer is that the skilled agent is in fact actively, not passively directing her attention. We can (and should) accept that she does not choose to attend *specifically* to (say) her left mirror just before overtaking the truck in front of her. Nevertheless, she does choose to attend to *driving-relevant information*. And this choice considerably circumscribes the range of stimuli she does (and



does not) attend to while driving, even if not fully. In contrast, the objects that the mind wanderer's attention are directed at are *not* similarly circumscribed by any choice she makes, leaving her attention to roam freely (cf. Wu 2014: 35-6).

Active attending is one phenomenon in the vicinity of MW that any account worth its salt should exclude (as the present account does). Another cognitive process that similarly falls outside the scope of MW is rumination or obsessive thinking. This type of behaviour may be roughly characterized as involving “persistent, recurring thoughts that revolve around a common theme and unintentionally enter consciousness, thus shifting attention away from one's current task goals” (Linville, 1996: 121). Ruminators find themselves entertaining recurring thoughts with common, often emotionally loaded, content. Thus a ruminating subject who was fired from his job may relive in her mind over and over again the awkward remark he made to his boss at the office party, berating himself for what he did and reinforcing his confidence that it cost him his job. It is this brooding pattern of thought that is taken to distinguish rumination, not the type of mental content or affect it involves (Nolen-Hoeksema, 1991); rumination can and does occur also in elated and angry moods. Nevertheless, the psychological literature understandably focuses largely on clinically depressed and dysphoric ruminators.

Now a major strand within this expansive literature highlights the extent to which depressive ruminators exhibit deficiencies in *cognitive control*. Ruminating subjects are described as suffering from cognitive inflexibility or perseveration, manifested in failure to modify their behaviour in the light of feedback or changing environmental conditions (Hertel, 1998). For example, one early study by Davis & Nolen-Hoeksema (2000) found that ruminators committed more perseverative errors than non-ruminators on the Wisconsin Card Sorting Task. The aim of the WCST is to determine what rule should be used when sorting target cards to match key cards that vary in colour, shape, and number. Participants received feedback about correct and incorrect matches they proposed, and were required to adjust their performance when the rule unexpectedly changes. Ruminators were found more often to persist in performing the task according to the old rule,

despite receiving feedback that their matches were incorrect. This suggests a lack of cognitive flexibility on ruminators' part.

The type of cognitive inflexibility from which ruminating subjects suffer has been associated with failures of *attentional* control – more specifically, failures to inhibit the processing of distracting information (Linville, 1996; Koster et al., 2011; Whitmer & Banich, 2007; Joormann et al., 2011). Joormann (2004) tested valence-specific inhibitory deficits in keeping information from entering working memory. Her subjects were first asked to ignore the emotional valence of an irrelevant stimulus. In a subsequent trial, the ignored emotional information became either task-relevant or task-irrelevant. Under these conditions, longer response time is normally expected because subjects must overcome inhibition. But this was not observed in depressive ruminators, indicating inhibitory dysfunction. In a similar vein, Joormann (2006) found that an increased tendency to depressively ruminate was associated with impaired inhibition of negative words.

All told, the evidence reviewed suggests that ruminators' capacity for attentional control is impaired, robbing them of the power to refrain from attending to the thoughts that haunt them. And this confirms that rumination is excluded from the proposed account of MW as it is an involuntary form of passive attention. If the verdict that ruminators are powerless to resist their ruminations seems harsh, it may help to recall that attributions of power to  $\mathcal{V}$  in the present context set a rather demanding bar, equivalent to having a specific ability to  $\mathcal{V}$ , or an ability plus opportunity to  $\mathcal{V}$ . Hence, a ruminating subject lacking the power to refrain from thinking about  $\mathcal{S}$  may still retain the (general) *ability* to refrain but lack the opportunity to do so, e.g. perhaps because of her temporary condition. This explains why in some cases we may be intuitively reluctant to deny that ruminative subjects lack the power to control their thoughts. In general, I submit that our intuitions about one's lacking or alternatively having the power to refrain from thinking about  $\mathcal{S}$  track our intuitions about  $\mathcal{S}$ -thoughts being ruminative or alternatively wandering: The less confident we are that a certain subject is specifically able to stop thinking about  $\mathcal{S}$ , the more inclined we are to regard her  $\mathcal{S}$ -thinking as ruminative rather than wandering.

In response, it may be pointed out that not only ruminators but also mind wanderers suffer from lack of cognitive control, hence this could not be what tells the two apart. Studies reveal that working memory capacity (WMC), which is broadly associated with executive control capabilities (e.g. Engle & Kane 2004), is correlated with the frequency of MW episodes: Subjects with lower WMC display increased tendency to engage in unintentional MW when confronted with a cognitively demanding task, and this partially explains their worse performance compared with high-WMC individuals (McVay & Kane 2012; Rummel & Boywitt, 2014; Robison & Unsworth, 2018). The thought is that members of the former group find it harder to prevent unwanted episodes of MW from taking place, which in turn disrupts their task-related activities.

Does this show that lack of executive control is characteristic of unwanted MW? One thing to note is that the data do not conclusively support this line of thought. Soemer & Schiefele (2020) for example found that WMC was negatively related also to intentional, not just to unintentional MW. Furthermore, the results cited in the response above admit of an alternative explanation in terms not of WMC but *motivation*, which has been shown to have an impact on the frequency of both intentional and unintentional MW alike (Seli et al. 2019). Given that low-WMC individuals can be less motivated to engage with demanding tasks (Linnenbrink & Pintrich, 2003), it is possible that they tend to avoid putting in the cognitive effort required to complete the task successfully, preferring instead to let their attention wander to more rewarding items. Moreover, older adults tend to report fewer episodes of MW during demanding tasks (Jordão et al., 2019). This does not sit well with the executive control hypothesis, which would seem to make the opposite prediction given that older individuals display decreased WMC (Wong et al., 2023).

Further, even if—waving the above points—we *should* accept that impaired cognitive control is responsible for increased rates of MW in low-WMC subjects, we are not forced to conclude that these subjects are without the power to steer their attention. For we should bear in mind that powers can be more or less difficult to exercise. Hence, the mere fact that one finds it difficult (even very difficult) to  $\nabla$  does not of itself indicate that one lacks the power to  $\nabla$ . Relatedly, it

seems important that the findings in question concern *low-WMC subjects* in *task-demanding* settings. For this may suggest that WMC and task demands are among the factors that contribute to determining how difficult it is to control one's attention. They do not vitiate the very ability to exercise control.

But does the above explanation of how MW is distinct from rumination run afoul of a salient difference between the two phenomena (as an anonymous reviewer suspects)? After all, as noted above, ruminators are widely seen as suffering from “persistent, recurring thoughts that revolve around *a common theme*” (Linville, 1996: 121; emphasis added),<sup>11</sup> while mind wanderers typically meander between various disparate topics. But the contrast between MW and rumination drawn above is not sensitive to this purported difference between their respective dynamics. In response, I agree that often there is some *one* focal event, theme or feeling whose influence is so seductive as to render ruminators unable to disengage. But I contend that this structural feature is not essential. It is entirely possible for ruminators' thoughts to focus obsessively on several (perhaps associatively related) thoughts. A subject whose mind travels in rigid succession between 3 or 4 mental items, unable to break this inflexible train of thought is, I submit, more intuitively regarded as ruminating than as mind wandering.

### 3.3 *Resolving the puzzle*

To successfully explain the nature of mind wandering, one must dissolve the tension between its passivity and its (occasional) intentionality. Defining MW as voluntary passive attention obviously vindicates the former feature; what about the latter? The present account is strictly neutral on this point. It can accommodate intentional MW if that turns out to be possible. But it may also provide grounds for resisting this idea. Start with the latter. And recall that §3.1 took pains to tease apart the active from the voluntary, contra common tendencies. Now, it is equally

---

<sup>11</sup> That this is a mark of rumination seems also to be reflected in the folk conception of the term (See Irving et al. 2020).

common for philosophers to lump the *intentional* together with the latter two (to verify, the reader need only call to mind the so-called ‘standard story’ of action, which sees action as an event caused by an *intention* to act). But much like the property of being active, being intentional should also be kept apart from being voluntary. The two are no doubt closely related and often go together; many processes are both voluntary and intentional. But not invariably. Indeed, MW supplies one counterexample: Often at least we voluntarily let our minds wander without intending for them to wander. All this suggests a preliminary basis for an *error theory* about the possibility of intentional MW. A sceptic may point out that subjects cannot be expected to reliably tell the (subtle) difference between their minds wandering voluntarily and their minds wandering intentionally. What they take to be the intentional character of their meandering thoughts may in fact be its voluntary character. This confusion may explain the (admittedly incredible) high volume of reported intentional episodes, cited above.

So much for arguing from the present account of MW to *denying* the reality of intentional episodes. As noted, the account can also accommodate the reverse option, that intentional MW is perfectly real. The way to do so is to separate the first- from the second-order with respect to the subject’s relation to her stream of thought. On the present account, MW subjects do not choose which items they attend to, and *a fortiori* do not intentionally choose which items they attend to. But this leaves open the possibility of subjects’ making a *meta-choice* – a choice about who (or what) chooses what they attend to, including external forces. In other words, MW subjects may be choosing to let their minds wander. And such choices may be intentional.<sup>12</sup> Importantly, this also points to a way of acknowledging a kernel of truth in the claim that MW is active: Whenever subjects exercise a second-order choice to let their minds wander, they may be seen as *actively* delegating the power to choose the course of their wandering mind.

---

<sup>12</sup> Compare Irving (2021), pp. 636-9.

## **Works cited**

- Aquinas, T. 1960–73. *Summa Theologiae*, ed. T. Gilby. Blackfriars
- Arango-Muñoz, S., & Bermúdez, J. P. 2021. Intentional mind-wandering as intentional omission: The surrealist method. *Synthese* 199(3): 7727–7748.
- Aristotle 1984. *The Complete Works of Aristotle*, ed. J. Barnes. Princeton University Press.
- Bird, A. 1998. Dispositions and Antidotes, *The Philosophical Quarterly* 48: 227–234.
- Carruthers, P. 2015. *The Centered Mind*. Oxford University Press.
- Davidson, D. 2001. Freedom to Act, in *Essays on Actions and Events* (Clarendon Press), Essay 4.
- Davis, R.N., & Nolen-Hoeksema, S. 2000. Cognitive inflexibility among ruminators and nonruminators. *Cognitive Therapy and Research* 24: 699–711.
- Dennett, D. 1984. *Elbow Room*. Oxford University Press.
- Engle, R.W., & Kane, M.J. 2004. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In: Ross (ed.). *The psychology of learning and motivation* (Academic Press): 145-199.
- Fara, M. 2005. Dispositions and Habituals, *Noûs* 39: 43–82.
- Geach, P. 2000. Intention, Freedom and Predictability, in Teichmann (ed.), *Logic, Cause and Action: Essays in Honour of Elizabeth Anscombe* (Cambridge University Press), pp. 73–81.
- Hart, H.L.A. 1968: *Punishment and Responsibility*. Oxford: Clarendon Press.
- Hart, H.L.A. & Honoré, A.M. 1985. *Causation in the Law*, 2<sup>nd</sup> edition. Oxford University Press.
- Hauska, J. 2008. Dispositions and Normal Conditions, *Philosophical Studies* 139: 219–232.
- Hertel, P.T. 1998. Relation between rumination and impaired memory in dysphoric moods. *Journal of Abnormal Psychology* 107: 166–172.
- Honoré, A.M. 1964. Can and Can't, *Mind* 73: 463–479.
- Hyman, J. 2015. *Action, Knowledge, and Will*. Oxford University Press.
- Irving, Z. C. 2016. Mind-wandering is unguided attention: accounting for the ‘purposeful’ wanderer. *Philosophical Studies* 173: 547–571.
- Irving, Z. C. 2021. Drifting and Directed Minds: The Significance of Mind-Wandering for Mental Agency. *Journal of Philosophy* 118: 614-644.
- Irving, Z. C. & Glasser, A., 2020. Mind-wandering: A philosophical guide. *Philosophy Compass* 15: p.e12644.
- Irving, Z. C., Glasser, A., Gopnik, A., Pinter, V., & Sripada, C. 2020. What does “Mind-Wandering” mean to the folk? An empirical investigation. *Cognitive Science*, 44(10), e12908
- Irving Z.C. & Thompson E. 2019. The philosophy of mind wandering, in Kieran F. & Kalina C. (eds.), *Oxford Handbook of Spontaneous Thought and Creativity* (Oxford University Press): 87-96.
- Jordão, M., Ferreira-Santos, F., Pinho, M. S., & St Jacques, P. L. 2019. Meta-analysis of aging effects in mind wandering: Methodological and sociodemographic factors. *Psychology and Aging*, 34: 531–544.
- Joormann, J. 2004. Attentional bias in dysphoria: The role of inhibitory processes. *Cognition & Emotion* 18: 125–147.
- Joormann, J. 2006. The relation of rumination and inhibition: Evidence from a negative priming task. *Cognitive Therapy and Research* 30: 149–160.

- Joormann, J., Levens, S. M., & Gotlib, I. H. 2011. Sticky thoughts: Depression and rumination are associated with difficulties manipulating emotional material in working memory. *Psychological Science* 22: 979–983.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I. & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science* 18: 614–621.
- Kenny, A. 1975. *Will, Freedom, and Power*. Blackwell.
- Kenny, A. 1982. Duress per minas as a defence to crime II. *Law and Philosophy* 1: 197-205.
- Killingsworth, M. A., & Gilbert, D. T. 2010. A wandering mind is an unhappy mind. *Science* 330 (6006): 932-932.
- Klinger, E. 2009. Daydreaming and fantasizing: Thought flow and motivation. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *Handbook of imagination and mental simulation* (pp. 225–240). Psychology Press.
- Koralus, P. 2014. The Erotetic Theory of Attention: Questions, Focus and Distraction. *Mind and Language* 29 (1): 26-50.
- Koster, E. H., De Lissnyder, E., Derakshan, N., & De Raedt, R. 2011. Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review* 31: 138-145.
- Lewis, D. 1999. Finkish Dispositions, in his *Papers in Metaphysics and Epistemology* (Cambridge University Press): 133-151.
- Linnenbrink, E. A., & Pintrich, P. A. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly* 19(2): 119–137
- Linville P. 1996. Attention inhibition: Does it underlie ruminative thought?, in R.S. Wyer (ed.), *Ruminative thoughts: Advances in social cognition* (Erlbaum): 121-33.
- Mac Giolla, E., Granhag, P. A., & Ask, K. 2017. Task-related spontaneous thought: A novel direction in the study of true and false intentions. *Journal of Applied Research in Memory and Cognition*, 6: 93–103.
- Martin, C.B. 1994. Dispositions and Conditionals, *The Philosophical Quarterly* 44: 1–8.
- McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology*: 141(2), 302–320.
- Metzinger, T. 2013. The myth of cognitive agency: subpersonal thinking as a cyclically recurring loss of mental autonomy, *Frontiers in Psychology* 4: 931.
- Metzinger, T. 2015. M-autonomy. *Journal of Consciousness Studies* 22: 270–302.
- Mole, C. 2011. *Attention is Cognitive Unison*. Oxford University Press
- Morsella, E., Ben-Zeev, A., Lanska, M., & Bargh, J. A. 2010. The spontaneous thoughts of the night: How future tasks breed intrusive cognitions. *Social Cognition*, 28 (5): 641–650.
- Mumford, S. 2001. Realism and the Conditional Analysis of Dispositions: Reply to Malzkorn, *The Philosophical Quarterly* 51: 375–378.
- Murray, S. & Krasich, K. 2020. Can the mind wander intentionally? *Mind and Language* 37: 432-443.
- Nolen-Hoeksema, S. 1991. Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology* 100: 569–582.
- Reid, T. 1788. *Essays on the Active Powers of the Human Mind*, ed. B. A. Brody (MIT Press, 1969).

- Robison, M. K., & Unsworth, N. 2018. Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(1), 85–98.
- Rummel, J., & Boywitt, C. D. 2014. Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review*, 1–7.
- Ryle, G. 1949. *The Concept of Mind*. Hutchinson.
- Seli, P., Carriere, J. & Smilek, D. 2015. Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research* 79: 750–758.
- Seli, P., Ralph, B., Konishi, M., Smilek, D. & Schacter, D. 2017. What did you have in mind? Examining the content of intentional and unintentional types of mind wandering. *Consciousness and Cognition* 51: 149–156.
- Seli, P., Risko, E., Smilek, D. & Schacter, D. 2016. Mind-wandering with and without intention. *Trends in Cognitive Sciences* 20: 605–617.
- Seli, P., Schacter, D.L., Risko, E.F. & Smilek, D. Increasing participant motivation reduces rates of intentional and unintentional mind wandering. *Psychological Research* 83, 1057–1069.
- Shope, R. K. 1978. The conditional fallacy in contemporary philosophy. *Journal of Philosophy* 75: 397-413.
- Smallwood, J., Fishman, D. J., & Schooler, J. W. 2007. Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review* 14(2): 230–236.
- Smallwood, J., & Schooler, J. W. 2006. The restless mind. *Psychological Bulletin* 132: 946 –958.
- Soemer, A., Schiefele, U. 2020. Working memory capacity and (in)voluntary mind wandering. *Psychonomic Bulletin & Review* 27: 758–767.
- Steward, H. 2012. *A Metaphysics for Freedom*. (Oxford University Press).
- White, A. 1964. *Attention*. Blackwell.
- Whitmer, A. J., & Banich, M. T. 2007. Inhibition versus switching deficits in different forms of rumination. *Psychological Science* 18: 546–553.
- Wong, Y. S., Willoughby, A. R., & Machado, L. 2023. Reconceptualizing mind wandering from a switching perspective. *Psychological research*, 87(2): 357–372.
- Wu, W. 2014. *Attention*. Routledge.