Responsibility for Forgetting To Do

Grünbaum, Thor

Responsibility for forgetting to do

Thor Grünbaum

Section of Philosophy	Department of Psychology	CoInAct Research Group
University of Copenhagen	University of Copenhagen	University of Copenhagen
Karen Blixens Plads 8	Øster Farimagsgade 2A,	Øster Farimagsgade 2A,
DK-2300 Copenhagen S.	DK-1353 Copenhagen K	DK-1353 Copenhagen K
Denmark	Denmark	Denmark

Abstract

Assuming that an agent can be morally responsible for her forgetting to do something, we can use recent psychological research on prospective memory to assess the psychological assumptions made by normative accounts of the moral responsibility for forgetting. Two accounts of moral responsibility (control accounts and valuative accounts) have been prominent in recent debates about the degree to which agents are blameworthy for their unwitting omissions. This paper highlights the psychological assumptions concerning remembering and forgetting that characterise the accounts. The paper then introduces and reviews recent empirical literature on prospective memory. Finally, it uses the literature to assess the various assumptions. One important implication is that a direct capacitarian control account implies implausible assumptions about the psychological capacity for remembering. A second important implication is that an indirect capacitarian control account and a valuative account highlight different but complementary aspects of remembering and forgetting.

1. Introduction

Many philosophers (Arpaly & Schroeder, 2014, 216; Clarke, 2014; Murray & Vargas, 2020; Sher, 2009; Smith, 2005; Vargas, 2020) and non-philosophers (Graf, 2012; Murray et al., 2019) agree that agents can be morally blameworthy for their unwitting omissions. Here is a possible case:

> Despite promising her partner in the morning, Sarah forgot to unlock the back door before leaving for work. Therefore, later that day, the plumber was unable to get in to repair the furnace and left a note saying that he would not be able to come again until next week (case freely adjusted from Ginet, 2000).

Suppose Sarah has not planned her forgetting. She did not suffer from any kind of neurological or psychological damage or breakdown. She was not operating under any unusual stress or duress. She was not coerced in any way. She has a good everyday memory. Nevertheless, on this occasion, she forgot to unlock the back door as she had promised she

would do. This omission resulted in an extra bill from the plumber and complicated logistic rescheduling. Sarah's partner seems warranted in holding her morally accountable for forgetting to unlock the back door before leaving for work. Sarah will not get off the hook by simply pointing to the fact that she forgot.

Cases of forgetting to do have attracted the attention of philosophers in part because different theories of moral responsibility make opposed predictions about responsibility. Roughly, philosophers who explain responsibility mainly in terms of exercising a capacity for control (control accounts) claim that people can be responsible for forgetting to do something because they are in control of what they remember to do, while philosophers who explain responsibility in terms of manifesting due care for others (valuative accounts, sometimes the position is called attributionism) claim that people can only be responsible for forgetfulness that manifests a morally relevant lack of care or concern for others.

According to control accounts, forgetting to do something is a counter-case to the valuative accounts because an agent can be morally responsible for forgetting to do something even if the action was important to the agent (Clarke, 2017b; Murray & Vargas, 2020; Nelkin & Rickless, 2017). Thus, control theorist could hold that Sarah would be blameworthy even if unlocking the door was important to Sarah – that is, even if her forgetting is not an expression of her lack of care for her promise and for her partner. By contrast, according to valuative accounts, forgetting to do something is a counter-case to the control accounts because an agent can be blameworthy for her forgetting to do something even if she is not in control of remembering (Smith, 2005; Talbert, 2017). That is, since Sarah is not aware of having forgotten to do something, it is not up to Sarah whether she recalls or forgets to unlock the door.

The dialectical situation is like a showdown at High Noon: the kind where the two philosophical gunslingers fatally shoot each other at the same time. The central ammunition for this shootout is a set of psychological assumptions about what can be controlled in situations of forgetting and how subjective importance and forgetting are related to each other. Remarkably, very little work has been done on elaborating and arguing for these psychological assumptions. This is all the more remarkable because there is a rich and substantive psychological literature about the psychological mechanisms of remembering to do future actions. The aim of the present paper is to articulate the psychological assumptions endorsed by control accounts and valuative accounts, respectively, and to evaluate the assumptions by comparing them to experimental results and models of prospective memory. Since the psychology of remembering and forgetting to do something is relatively well understood, we can evaluate this set of assumptions by looking at their plausibility from the perspective of contemporary cognitive psychology.

By reviewing recent experimental research on prospective memory, I show that some control capacities attributed to forgetful agents by some accounts of moral responsibility are implausible, whereas others are not. Furthermore, I argue that research on prospective memory suggests that, contrary to the general view, control accounts and valuative accounts of moral responsibility might not be making conflicting predictions about the responsibility of forgetting. The factors that our best psychological theories of prospective memory describe as being central to the probability of remembering one's intention are also the aspects typically identified by these conflicting accounts of moral responsibility. According to psychological theories of prospective memory, certain forms of control and importance are equally central to

remembering to do something. Consequently, forgetting to do cannot be used as a critical test case of which of the two conflicting types of account of moral responsibility is correct.

The paper proceeds as follows. In Section 2, I introduce the two competing accounts of moral responsibility with a special focus on their conflicting explanations of an agent's moral responsibility for forgetting to do. In Section 3, I describe some of the substantial psychological assumptions that characterise the two accounts. These are assumptions about the psychology of remembering and forgetting to do something. In Section 4, I introduce a number of main findings in recent cognitive psychology of prospective memory, and, in Section 5, I use the findings to assess the assumptions described in the previous section. Section 6 outlines some consequences for a theory of moral responsibility for forgetting to do.

2. Control accounts and valuative accounts

When a person violates an obligation, we are sometimes warranted in reacting to the behaviour of the person with resentment or indignation. In that case, the person would be blameworthy in the sense of being an appropriate target of reactive attitudes. This is the sense of moral responsibility that will concern us in this paper (the accountability sense, Shoemaker, 2011; Watson, 1996). I focus exclusively on situations in which an agent forgets *to do something* she is obligated to do. For stylistic reasons, I will use the phrase "moral responsibility for forgetting", but keep in mind that the phrase refers to a situation where the agent has forgotten to perform some act in some particular situation. "Forgetting" does not in this paper refer to situations where a person forgets some fact (like *semantically* forgetting the order of the various French revolutions or *episodically* forgetting when you saw a friend the last time; see Grünbaum & Kyllingsbæk, 2020).

2.1. Control accounts of moral responsibility for forgetting to do

Control account is here understood in a broad sense. It includes any account that conceives of the lack of control as an excusing condition. It allows for a variety of ways of understanding control and thus includes understanding control in terms of choices and volitions (Frankfurt, 1971), moderate reasons responsiveness (Fischer & Ravizza, 1998), or cognitive and volitional capacities (Clarke, 2014; Vargas, 2013). Control in this sense includes both a freedom aspect (the agent is excused if the action or omission is unfree) and an epistemic aspect (the agent is excused if she did not have the requisite beliefs or knowledge), without prejudging whether or not these aspects are independent of each other (Mele, 2010). Finally, a control account of moral responsibility can be either *direct* in the sense that the agent is directly responsible for her action or omission (Clarke, 2017b) or *indirect* in the sense that the agent's responsibility can be traced back to her responsibility for some other action or omission (Fischer & Tognazzini, 2009).[1]

Cases of blameworthy forgetting cannot easily be accommodated by control accounts. Sarah is blameworthy for forgetting to unlock the back door. But how could she be blameworthy in virtue of having the right kind of direct control in the situation? She wants to do the action, and the only reason she does not do it is because she forgets. Being in a state of

---

[1] In this paper, I will first consider a version of the direct control account of forgetting and only later in the paper (in Section 5) will I consider a version of the indirect account of forgetting.

blissful ignorance about the danger of forgetting or the fact of having forgotten, Sarah is temporarily unresponsive to reasons for unlocking the door. After all, she has already accepted to unlock the door and accepted that it is the right thing to do. Short of being prompted by external cues, she is temporarily ignorant of her omission and its bad consequences. Thus, the choice aspects, the reasons-responsiveness aspects, and the epistemic aspects of control of her omission all seem to be abolished in the situation.

Given this challenge, what would a direct control account of forgetting look like? Is there any form of control that Sarah might have for the duration of her failing to remember that could ground a form of direct responsibility for her forgetting? Not many options are left other than to insist that some unexercised control capacity grounds her direct responsibility. According to the *direct capacitarian control account* of forgetting to φ, the agent is responsible for her omission if (1) she is free to φ (that is, she is free from breakdowns, coercion, and manipulation, has the general capacity to φ, and the opportunity to φ) and (2) when she forgets she falls below some standard that applies to her, given her general cognitive and volitional capacities and the opportunity to exercise them in the situation (Clarke, 2014; Murray & Vargas, 2020).

Given that Sarah is normally a conscientious person that remembers to do what she planned to do, what she is expected to do, and what she has promised to do, if Sarah has the general capacity to unlock the door and the opportunity to exercise the capacity on this occasion, and if she has the general capacity to recall her intention to unlock the door and the opportunity to recall on this occasion, Sarah is blameworthy for not unlocking the door.[2] Today, when she forgets to unlock the door, she falls below her own cognitive standards, and she is blameworthy for this slippage. She could have and she should have recalled and she could have and she should have unlocked the door. Sarah's partner is therefore warranted in blaming her for forgetting to unlock the door.

According to this direct capacitarian control account, the agent is directly morally responsible for not doing something because, in the situation of forgetting to act, the agent had the type of control sufficient for moral responsibility. Notice that, according to this account, conscious awareness is not a necessary condition for moral responsibility. That is, being unaware of reasons to unlock the door or her state of having forgotten is not a circumstance that defeats Sarah's specific capacity to remember and to perform the act. Furthermore, it is important to notice that according to this account an agent can be blameworthy for her omission without being blameworthy for her forgetting (Clarke, 2014, 2017a). Forgetting and recalling might be too passive to be appropriate targets of praise or blame. Thus, the direct capacitarian

---

[2] Terminology varies in the literature. According to McGeer and Petit (2015), the conjunction of the agent's general capacity and the opportunity to exercise it is the agent's *specific capacity*. According to Cyr and Swenson (2019), we should understand a *specific ability to φ* as the opportunity to φ plus some epistemic condition (awareness or know-how). For the purpose of this paper (and using the terminology of Cyr & Swenson, 2019), I will assume that a specific ability equals *an opportunity + relevant knowledge*, where an agent has opportunity to φ in the actual world roughly when there is a possible world W in which the agent φs and, at the very least, everything except S's φ-ing, is the same as in the actual world. To really work, this rough definition would need some refinement (see, Cyr & Swenson, 2019; Franklin, 2011; Swenson, 2016). Cyr and Swenson argue that a specific ability to φ does not entail a general ability to φ and that in some cases the general ability is not necessary for blameworthiness. For the sake of the argument, I will assume that the general ability is necessary (but nothing in my argument really hangs on this assumption). The important assumption is that the specific ability is necessary for moral responsibility. If I omit to open the door when someone knocks because I am tied to a chair, then I am not blameworthy for the omission.

denies the principle that a person is responsible for her unwitting omission only if she is responsible for being unwitting, that is, for her ignorance (Clarke, 2017a).

2.2. Valuative accounts of moral responsibility for forgetting to do
Valuative accounts avoid the difficulties of specifying the required kind of control in situations of forgetting by relating moral responsibility exclusively to the agent's attitudes and evaluative judgements. An agent is blameworthy for her action or omission if and only if (1) the action or omission is an indication or expression of attitudes, values, or cares, (2) these attitudes, values, and cares are morally objectionable, and (3) it makes sense to inquire about the agent's reasons for her attitudes, values, and cares (Mason, 2015; Smith, 2005; Sripada, 2015; Talbert, 2016). Consequently, Sarah is morally blameworthy for her forgetting to unlock the door to the extent that her forgetting is an expression of the low importance that keeping the promise has for her and her assignment of low importance is morally objectionable.

Assuming that we can be responsible for forgetting, one obvious problem for this type of account is that agents forget to do things that are important to them. Take Sarah. She might not have any of these morally objectionable attitudes. She cares about her partner, she cares about keeping promises, and she thinks it is important not to annoy the plumber and inflict the household with an extra bill and logistic rescheduling – she simply forgets to unlock the back door. Thus, her forgetting to unlock the door is not an expression of wrongful attitudes, values, or cares. Consequently, some valuative theorists would draw the conclusion that Sarah is not blameworthy for her forgetting to do (even if regret and annoyance might be appropriate responses, Björnsson, 2017; Talbert, 2017). If ordinary cases of forgetting are like Sarah's case, some versions of the valuative account entail a general scepticism about the moral responsibility for this type of unwitting omission.

3. Psychological assumptions
The aim of the paper is to articulate the kind of assumptions about human psychology of cognitive control and remembering that are made by the accounts of the moral responsibility of forgetting to do. The accounts are different from one another (among other things) in virtue of the assumptions they are each making about the human psychology of remembering and forgetting.

3.1. The direct capacitarian control account
The direct capacitarian control account of forgetting assumes that an agent, in a situation like Sarah's, is free to perform the action in question and has the general psychological capacity and the opportunity to exercise it. Since an agent like Sarah is under the obligation to perform the action in question and has the capacity to remember, it is reasonable to expect her to remember. And since we are justified in expecting her to remember, then, given there are no excusing conditions, we are warranted in blaming her for not performing the action (Clarke, 2017a, 244-245). The direct capacitarian account depends on the correctness of the claim that an agent in Sarah's situation has the specific capacity to remember to do something. How plausible is this claim? We need to evaluate two things in order to assess the plausibility.

First, how should we understand the relevant general capacity to remember to do? According to Clarke (2017b), it is the capacity for unaided and effortless recall of information

when and where the agent needs it. As Clarke puts it, the agent has "a capacity to think at appropriate times, without making any effort to recall, to do the things she has only recently promised to do" (Clarke, 2017b, 71). That is, the agent is not required to take any deliberate or strategic steps in order to ensure recalling her planned action. All that is required is that the agent possesses the general ability to think to do the action (which she has in "the absence of madness, hypnotism, or some such condition" (Clarke, 2017b, 71)). The agent has the opportunity to exercise this ability if nothing in her situation blocks her thinking and performance. Murray and Vargas (2020) tell a related story in terms of strategic allocation of cognitive resources in order to manage information processing in situations of multiple goals. Given the agent's standard of vigilance, she has the capacity to allocate the required cognitive resources for remembering and doing.

The exercises of these mental capacities are our "doings, and failures to exercise them, when we have them, are our failures to do things we are capable of doing" (Clarke, 2017a, 242). We can have these mental powers also when we do not manifest them. Just as salt has the power to dissolve in water, also when there is no water around, the agent has the general capacity to recall also when she is not exercising her capacity. However, unlike the power of salt to dissolve in water, remembering to do is a rational capacity. It is not merely mechanically triggered by an environmental condition. Rather, according to Clarke and other direct capacitarians, we can expect the agent to exercise the capacity if she has (moral or prudential) reasons to exercise it (and she is free from "unfreedom").

Second, how should we specify the ordinary situation of a forgetful agent? It is a situation where the agent has a decisive reason to remember to do. This reason might be moral, as in Sarah's case where she promised to unlock the door, or it might be of a more prudential character, as in the case where I decided to phone a friend after dinner (but only I know about it). It is also a situation where the agent has forgotten. This will usually imply that nothing in the agent's perceptual or mental context operates to remind her. In the situation of having forgotten, she is completely unaware of what she has to do and the fact that she has to do something.

Given these considerations, the direct capacitarian control account of forgetting makes two substantial psychological assumptions. In a situation of forgetting where the agent is completely oblivious to her reasons to φ and is not mentally processing any cues associated with φ-ing, the account assumes that (1) the forgetful agent has a general capacity for unaided and effortless recall of her intention to φ and the opportunity to exercise it, and (2), given basic freedom from breakdowns, stress, and coercion, we should expect the agent to exercise her capacity to remember to φ. Consequently, if the agent is morally responsible for forgetting but the agent in the situation of forgetting does not have the specific capacity to recall her intention and we should not expect her to recall, then the account of her moral responsibility of forgetting could not be a direct capacitarian account. That is, assuming that Sarah is blameworthy for forgetting to unlock the back door, if we have reason to think that she did not have the specific capacity to recall and we should not expect her to recall, then we have reason to reject the direct capacitarian account.

3.2. The valuative account

The valuative account of forgetting assumes that an agent is morally responsible for her actions and omissions only if they reflect and are sensitive to her evaluative judgements, cares, and commitments. According to this view, we are responsible for our actions and omissions because they reflect our evaluations, and we are responsible for our evaluations because "we can be asked to defend or justify them insofar as they reflect our judgments about what is important. And insofar as our attitudes reflect *objectionable* judgments, they may be a source of the sort of moral offense that reasonably elicits the responses involved in blame" (Talbert, 2016, 45).

This has consequences for the view's conception of unwitting omissions, like forgetting something or not noticing something. According to a valuative account, an agent is morally responsible for her unwitting omission to the extent that the omission reflects and is sensitive to the agent's evaluative judgements and the agent can reasonably be asked to defend her evaluation. As Smith puts it, "we can be criticized and asked to acknowledge fault for failing to notice [or remember] something if this failure can reasonably be taken to reflect an (objectionable) judgment that the thing in question is not important" (Smith, 2005, 270).

A valuative account about forgetting is thus standardly assuming a substantive relationship between a person's evaluative judgements and what she remembers. Though perhaps not strictly necessary, it is natural for the account to assume that the more important a task is for an agent, the more likely it will be that she remembers to perform it (Smith, 2005, 244). This implies that we can use facts about what a person fails to notice or remember as evidence for what the person cares about and values because there exists a rational connection between a person's evaluations and what she notices and remembers (and fails to notice and remember).

4. Psychology of prospective memory

In this section, I describe a number of psychological factors that influence the probability of remembering one's intentions. I will assume that agents can form future-directed intentions by making decisions about doing something at some point in the future (but intentions can be formed in other ways as well). Further, I will assume a Bratman (1987) inspired account according to which a future-directed intention is functionally characterized by conduct-control, stability, and certain rational roles in reasoning. According to this type of account, once a person has formed her intention, causal processes support the persistence of the intention (Grünbaum & Kyllingsbæk, 2020). An agent's intention can persist without being operative in the sense of playing a role in the agent's reasoning, planning, decision-making, and action control. Let us call a persisting intention which is absent from an agent's mind in this way *a standing intention* (following Mele, 2007). When the intention is operative in the agent's mind, it is *an occurrent intention*. When the intention is standing, it is not currently ready to play any such role but it is disposed to become ready given the right conditions. Retrieval of one's intention means that one's standing intention (represented in long-term memory) becomes occurrent (represented in working-memory). Forgetting one's intention means that it does not become occurrent in working memory when it should. An agent should remember her intention not only when the time and place has come to execute it but also during crucial stages of practical reasoning. It is important to remember one's intention when the intention is only

abstract and schematic and is in need of further planning in order to be carried out. It is also important to remember one's intention if one starts considering doing some other action at the same time as the originally planned action.

4.1. Automatic retrieval

It is generally accepted in research on prospective memory that agents can maintain intentions or task instructions occurrent in working memory only at a cost to ongoing performance (Sculling, McDaniel, & Einstein, 2010; Smith, 2003; Smith & Bayen, 2005). Rather than using up her limited cognitive capacity by keeping multiple intentions present to her mind, it would be better for the agent to use her long-term memory. In an ordinary situation, like the one where Sarah promises to unlock the back door before leaving for work, in the time span between making the promise and preparing to leave for work she is most likely not aware of her intention or only aware occasionally. We should not expect her to maintain her intention in working memory, since this would involve a cost (in terms of decreased accuracy and increased response time) to all of her ongoing morning activities.

Retrieval of standing intentions from long-term memory is generally thought to be an "automatic" process in the sense that it is usually not a conscious, voluntary search activity. Agents generally rely on environmental cues to trigger the retrieval (McDaniel & Einstein, 2000; McDaniel & Einstein, 2007). The cue might be consciously perceived but the triggered process of retrieving the intention does not require any conscious steps. Sometimes the context triggers a cognitively costly monitoring of the environment for cues determining the when and where of action (Guynn, 2003; Smith, 2016), but, often, the intention simply pops into one's mind (Kvavilashvili & Mandler, 2004; Reese & Cherry, 2002).

4.2. Content and importance of the intention and the probability of retrieval

Cognitive psychology has made significant progress in studying the conditions under which standing intentions tend to become occurrent, and the extent to which agents can control the process. Two general properties of intentions regulate their disposition to become occurrent: their content and their subjective importance.

Let us first look at the content. I will follow the psychological literature in focusing on the situation-based conditional intentions (known as *event-based* in the psychological literature on prospective memory). The content of the intention represents what the agent will do in a situation: I shall φ when situation S obtains. The content includes information about the situation (objects and events), the motor action, and propositional information about the place of the intention in a larger goal hierarchy (Grünbaum & Kyllingsbæk, 2020).

The content can vary in the dimension of specificity. It can be very abstract, such as an agent's intention to host next year's summer barbecue. When the agent first forms the intention (say, nine months before the party), it contains only abstract and limited information about objects, events, and motor actions, but it nevertheless contains schematic barbecue related information. The content of an intention can also be very specific, such as an agent's intention to raise her hand the moment the speaker has finished speaking. This type of intention contains very specific information about attentional cues and motor actions. Abstract and specific content are often related to each other by instrumental practical reasoning.

Convincing evidence suggests that the degree of specificity of one's intention is related to the probability of remembering the intention. Various lines of experimental evidence support this claim, generally from experiments using some form of the so-called event-based prospective memory (PM) paradigm. In the event-based PM laboratory paradigm, participants are typically instructed to complete a forced choice ongoing task (e.g., lexical decisions; press "F" for word, and "J" for non-word). At the outset of the ongoing task, some participants are instructed to remember to perform a third alternative response (e.g., press the "F7" key) if they are presented with a PM target event (e.g., a particular word such as *elephant* or a particular syllable, *tra*).

First, in an early study, Ellis and Milne (1996) showed that if a participant is preparing to react to specific cues for action (in contrast to abstract cues), it increases the probability of remembering to perform the PM task. Second, in a number of studies, McDaniel and Einstein have shown that (so-called focal) cues with substantial processing overlap with cues for the ongoing task increase the probability of remembering to perform the PM task (Anderson, Strube, & McDaniel, 2019; Scullin et al., 2010). Third, studies have demonstrated that when participants imagine the action-cue relation when they are encoding the intention into memory, it increases the likelihood of retrieval (Sculling et al., 2017; Spreng, Madore, & Schachter, 2018). When researchers compare the performance of participants required to mentally imagine performing the prospective memory task to performance of participants required to mentally imagine performing the ongoing task, participants in the experimental group are more likely to remember to perform their delayed intention. Fourth, classic experiments on *implementation intentions* have obtained similar results using verbal instructions (Gollwitzer, 1999). In one series of experiments, participants in the implementation intention condition were asked to repeat statements like "When I see any word that belongs to the category fruits, I will remember to press the Q key", whereas participants in the control conditions were asked to repeat "My category is any fruit words, and my target key is Q". Again, participants in the implementation intention group are more likely to remember their delayed intention (Sculling et al., 2017).

Collectively, these lines of evidence indicate that the probability of remembering one's intention and executing it is partly a function of the degree of specificity of its information about objects and events, motor actions, and the rational place of the intention in a larger planning hierarchy. These lines of evidence have led to the formulation of the *multiprocess view* according to which agents are able to make strategic use of bottom-up, automatic cue-driven retrieval of intentions from long-term memory by forming conditional intentions that represent a cue-action association (McDaniel and Einstein, 2007; Shelton, Scullin, & Hacker, 2019).

The second general property of standing intentions that regulates their probability of becoming occurrent is their importance to the agent. We can think of subjective importance as the sum of the expected utility of performing the intention and the estimated subjective costs of planning and performing the action. This means that the subjective importance is a combination of the estimated value, probability of success, and effort involved in planning and executing. In studies of prospective memory, the subjective importance is usually manipulated as one single factor. Only recently have researchers started to investigate the involved components separately (Shenhav, Botvinick, & Cohen, 2013; Shenhav et al., 2017). The subjective importance is a dynamic property. It may change over time due to new information

concerning the possible rewards, the possibility of succeeding, and the intrinsic costs of planning and executing the intention (Kool, Shenhav, & Botvinick, 2017; Kool, Gershman, & Cushman, 2018). The importance is a relative property. What matters is the importance of an agent's intention relative to the importance of each of the other intentions of the agent. The more important an intention is for an agent relative to her other intentions, the more likely it is that the agent will recall her intention – that is, the more likely it is that the standing intention will become occurrent (Grünbaum, Oren, & Kyllingsbæk, 2021).

Let me mention three forms of evidence in support of the role played by importance in retrieval of standing intentions. First, experiments have manipulated intentions in a standard PM paradigm by monetary reward for remembering to do the PM task or punishment for forgetting to do it. Results generally show better performance on the PM task when more monetary value is added to the intention (Walter & Meier, 2014). For instance, Cook and colleagues showed that situations where the cue representations remain the same between various reward conditions, the probability of remembering the PM task is a function of the level of reward (Cook, Rummel, & Dummel, 2015).

Second, Grünbaum, Oren, and Kyllingsbæk (2021) used mathematical modelling to demonstrate that in a task situation where an agent could retrieve and perform one of several relevant intentions from long-term memory, an important factor would be expected reward value. In their task situation, the participants would be able to choose one of four possible task-sets. Their modelling demonstrated that the probability of retrieving a task-set could be described as weighted by the expected reward.

Third, in some situations, the agent's motivations have changed drastically and executing the intention is no longer relevant to the agent. Even if there is a match between a cue represented by the conditional intention and a perceived cue, the intention might not be recalled because the importance of the intention is now low. This is the case when participants have been instructed that the PM task is no longer relevant (Scullin, Einstein, & McDaniel, 2009; Scullin & Bugg, 2013).

Summing up, two factors regulate the probability that a standing intention becomes occurrent: the match of the content of the intention to the context and the relative subjective importance of the intention. To be sure, these might not be the only factors. These two factors do not deterministically ensure that a person will recall the intention. Rather, they influence the probability that she will recall the intention. This means that even if a very specific content finds an accurate match and the intention is very important to the agent – that is, even if there is a high probability that the intention will become occurrent – it is still possible that she will forget it (Grünbaum, Oren, & Kyllingsbæk, 2021).

This account of the factors involved in regulating the retrieval of a standing intention also provides us with an account of forgetting. Forgetting is a failure to become occurrent at an appropriate time and place, which means that some other intention (or no intention at all[3]) has become occurrent instead. Think of it as a selection problem. One's mind has to select which standing intention to make operative. Forgetting the intention to φ happens when some other intention is selected and becomes occurrent, ready for use in reasoning, planning, decision-

---

[3] I will set aside this type of situation. It could occur if the agent falls asleep, falls into a goalless drowsiness, starts involuntarily to mind-wander, or some similar non-intentional state or process.

making, and action control. According to the account I have sketched here, three different elements can be involved in forgetting to do: (1) Another intention has a better match with the context, (2) another intention is more important to the agent, or (3) another intention might be selected due to chance (even if it has a worse match and is less important to the agent).

The experimental data on prospective memory suggest that a person has some voluntary control over the probability of remembering her intention. This kind of control over the probability of remembering works at encoding. By encoding, I mean the process by which occurrent content becomes persistent as a representation in memory. The agent influences the encoding by the way in which she articulates the content of the intention by imaginative rehearsal or instrumental reasoning whereby the conditional relationship between context cues and action becomes articulated. It is important to notice that control at encoding is not something the agent has only once, when the intention is first formed. The agent has this kind of control over the persisting intention every time that the intention becomes occurrent, which is usually many times. Schematic intentions need to become occurrent when the agent engages in further means-end reasoning. Intentions also need to come to mind when the agent is making other plans – to ensure that she does not double plan. Finally, intentions tend to occur to the agent with increasing frequency as the time for execution approaches (Elis et al., 1999). Every time that the intention becomes occurrent, the agent has the opportunity to increase the specificity of the content and thereby increase the probability of future retrieval.

5. Psychological assumptions and the scientific evidence

Let us return to the psychological assumptions made by the accounts of moral responsibility for forgetting, starting with the *direct capacitarian control account*. Recall that according to this account the forgetful agent is morally responsible for her unwitting omission if she should have and could have recalled her intention. This suggests two assumptions.

First, in the situation where the agent is not mentally processing any cues associated with the forgotten task, the forgetful agent has the specific capacity for unaided and effortless recall. As we saw above, this requires merely that the agent possesses the general ability to think to do the thing and has the opportunity to exercise the general capacity (i.e., nothing in the situation blocks the exercise of the capacity). The problem here is the direct capacitarian's conception of opportunity. Clarke (2017a) assumes that this general capacity is a rational power rather than a merely mechanical triggering of some disposition in the right context. However, if we focus exclusively at the moment of forgetting or the moment of successfully retrieving one's intention, evidence suggest that the process is more mechanical and environmentally dependent than the capacitarian assumes. If the cue specificity is low or non-focal, it will significantly increase the risk of committing so-called "omission errors" (McDaniel & Einstein, 2007), where participants miss PM targets and forget to do the PM task. Moreover, if cue specificity is high or cues are focal, it will increase the risk of committing "commission errors" (committing the action even if it is no longer relevant and the importance is low, Bugg & Streeper, 2019). The experimental manipulation of these types of mistake supports the claim that if the prospective task is completely absent from the participant's mind, while the agent is engaging in her ongoing task completion, retrieving her prospective intention seems to depend almost fully on her coming across some information that can trigger her cue-action association and make her intention "pop" into her mind. This indicates that absent the awareness or

processing of cues that can trigger retrieval of the intention, the agent does not have the specific capacity for retrieval. Without the right context and awareness, the agent does not have the opportunity to recall her intention.

The direct capacitarian might reply that the agent has the opportunity to exercise her capacity to recall her intention, it is just that due to lack of vigilance she falls below her own cognitive standard and fails to exercise the capacity (Clarke, 2014; Murray & Vargas, 2020). Had she been adequately vigilant, she would have noticed the cue and she would have remembered to perform the action. She is to blame for not being sufficiently vigilant, and thereby not noticing the cue and not remembering to perform the action.

It is difficult to understand what vigilance could mean here other than active monitoring of the environment for cues for action. This type of vigilance presupposes that the agent is already aware of her intention. The intention is already operative, and she is monitoring the environment for cues indicating when to execute her intention. This means that the intention has already been retrieved and is now maintained in working memory. Since we cannot maintain multiple intentions active in working memory without a significant cost to our ongoing actions, this insistence on vigilance would only push the problem to an earlier time point. Since agents cannot generally maintain their intentions active in working memory from the time of deciding to the time of execution, there will generally be a period in which the intention is standing and the agent is not yet monitoring for performance cues. During this period, whether the agent recalls her intention and starts monitoring for performance cues often depends on her coming across cues for retrieval of the intention. In absence of these retrieval cues, the agent is not in a context that affords her the opportunity to recall her intention.

Second, according to the capacitarian, we should expect the agent to exercise her general capacity to remember if she has a decisive reason for doing so, also in the situation where the agent is completely oblivious to her reasons. The expectation is here both of a descriptive and a normative sort. It is related to both what we can predict that the agent will do and what the agent should rationally do, given various social norms and her prior commitments. Both sorts of expectation are dubious in the present scenario. The experimental data on omission and commission errors suggest that contexts with no cue-perception or where cues occur non-focally influence the probability of retrieval. The best prediction in such contexts might be that agents will forget. Furthermore, if we adopt an internalist perspective on the agent's rational reasons for engaging in the activity of retrieval, we have to remember that the agent is in a situation where no internal or external cues remind her of her intention or action. Rationally, we should not expect her to jump from her focus on her ongoing tasks to suddenly recalling her intention, unless she came across some reminder (Levy, 2017).

This might not be a decisive argument against Clarke or any other direct capacitarian control account. The direct capacitarian can insist on norms in some external sense. The direct capacitarian could point to norms for social practices rather than norms for individual instances of performance. This is Vargas' view (2013) according to which you can expect me to successfully recall my intention because your expectations over time play important social roles in shaping my self-control and moral agency. To the extent that the direct capacitarian insists on normative expectations for the forgetful agent, the direct capacitarian seems to commit to substantive consequentialist assumptions and metaethical assumptions about

external reasons – at least with respect to moral responsibility. The psychological data do not address these normative and metaethical options.[4]

For the sake of the argument, let us assume that it is a virtue of an account of the moral psychology of agency that it avoids making substantive assumptions in normative ethics and metaethics. The psychological data indicate that we should not psychologically and internally expect an agent to retrieve her intention if she is not aware of any relevant cue. Consequently, conditional on the psychological data, the direct capacitarian account seems to be a less probable account of the moral responsibility of forgetting than the valuative accounts.

There might be other ways to salvage a capacitarian control account, even conditional on the psychological data. Some capacitarian control theorists are sceptical about the claim that responsibility-level control can be exercised during the time of failing to remember to φ (Nelkin & Rickless, 2017). The solution might be to relocate the opportunity to exercise the relevant form of control to an earlier point in time. When it comes to the responsibility of forgetting to do, it is important that what is relocated is the *opportunity to exercise* the capacity rather than the *actual exercise* of the capacity. We cannot trace Sarah's blameworthiness back to any acts of explicitly deliberating about the risk of forgetting to unlock the door or explicit performance of some "benighting" act leading to forgetting (Levy, 2014; Rosen, 2004; Zimmerman, 1997). Consequently, if an account relocating the specific control capacity is to succeed in explaining the blameworthiness for forgetting, the responsibility should be relocated or traced back to situations where the agent had an opportunity to exercise her control, rather than back to conscious choosing of actions or omissions with foreknowledge that forgetting would be the result.

According to such an "opportunity tracing account" (Nelkin & Rickless, 2017), when making the promise to her partner, Sarah was aware of the general risks of forgetting and had the opportunity to strengthen her memory by internal rehearsal or by associating the action with a cue or by producing some external prompt that could help her (such as a note on the front door, etc.). If the agent is morally blameworthy for her forgetting to do, then, at times prior to her forgetful state, the agent "must have had an opportunity to act preventatively" (Nelkin & Rickless, 2017, 120) with respect to the risk of forgetting. Consequently, at some prior time, the agent must have been aware "that there is something she can do then that will sufficiently raise the likelihood of avoiding omission at [a later time]" (Nelkin & Rickless, 2017, 120).

This form of indirect capacitarian control account finds some support in empirical studies suggesting that agents have some voluntary control over the probability of successful future retrieval of intentions. Agents have the opportunity to exercise the control at encoding. Encoding takes place not only when the intention is first formed but also at all of the other times when the intention comes to mind in advance of the situation of executing it. The agent exercises this partial control by forming conditional or implementation intentions where the action is associated with cue and by rehearsing the cue-action association by mental imagery.

One might worry that the epistemic conditions of the indirect capacitarian control account are too demanding (Rudy-Hiller, 2019; Vargas, 2005). It is necessary that the agent is aware of her opportunity to exercise her control and is aware of the risk of forgetting if she adopts the

---

[4] I am grateful to a reviewer for helping me understand and express this point.

wrong strategy for encoding her intention. But some evidence suggest that people might actually possess the required awareness or at least the capacity to make the required assessments. According to the dominant multiprocess view of prospective memory, people have a good metacognitive awareness of their own memory abilities and assessment of the situation of retrieval (Einstein & McDaniel, 2005). If the importance is high or the situation of retrieval assumed to be demanding, people will adopt an encoding strategy with high cue specificity or focality, thus enabling a more automatic intention retrieval. The general assumption of this framework is that the less confident a person is with respect to successful retrieval of her intention, the more likely the person is to adopt a bottom-up cue-dependent retrieval strategy.

Some, though not many, studies have investigated the metacognition of prospective memory (Kuhlmann, 2019). Studies comparing the participants' prediction of their own performance in a PM task with their actual performance generally show that participants are underconfident (Cherkaoui & Gilbert, 2016; Gilbert, 2015a; Gilbert, 2015b; Meeks et al., 2007). That is, participants predict that they will perform worse than they actually do. These predictions are often sensitive to contextual factors suggesting potential retrieval difficulties and to motivational factors (importance). In sum, this data indicate that agents generally do have some metacognitive awareness of the opportunities for minimizing the risk of future forgetting. This data is scarce and far from definitive. But the data should suffice to draw into doubt any armchair judgement about the lack of required metacognitive awareness of the risk of forgetting. Summing up the argument thus far, conditional on psychological data and models of prospective memory, *indirect* capacitarian control accounts are more probable than their *direct* versions.

How about the valuative accounts? The valuative account of forgetting claims that the forgetful agent is responsible for her unwitting omission only to the extent that her forgetting is the expression or indication of her morally blameworthy attitudes or evaluative judgements. Often proponents of this type of account assume a relationship between how important an action is for the agent and the probability that she remembers to do it. As documented above, this assumption is supported by the psychological data. Experiments manipulating the importance of PM tasks generally find a positive relationship between importance and probability of retrieval.[5]

Even though, the psychology of forgetting to do does not fit the assumptions of a *direct* capacitarian account very well, forgetting to do is not a case that can critically test a capacitarian control account against the valuative account – as claimed by parties on both sides of the divide. In contrast to a dialectical High Noon situation, the psychological assumptions of the indirect capacitarian control account and the valuative account are equally supported by psychological models and experimental data. In line with the indirect capacitarian account, models of prospective memory typically leave an important place for encoding strategies.

---

[5] Maybe one could worry (as one of my reviewers did) that the notion of subjective importance in the psychological literature does not match the notion of importance in valuative accounts, where importance is often tied to notions such as "quality of will". I would rather turn it around as a strength. The psychological notion of subjective importance understood as (*subjective utility* x *probability - cost*) disambiguates and specifies what could be meant by "evaluative attitude". It relates the notion of an evaluative attitude to factors that can be manipulated concretely in psychological experiments and estimated in models.

Agents often have the opportunity to exercise some control over how they encode their intention and thereby over the probabilities for retrieving the intention when the time comes. They have this partly in virtue of having the required metacognitive awareness of the risks of future forgetting. In line with the valuative account, experimental data suggest that importance influences retrieval. The higher the importance assigned to the intention, the higher is the probability of retrieving it. Unwitting omissions can therefore be a sign of low importance, i.e., lack of interest or care. I will get back to this dialectical issue in the next section.

I am not claiming that the psychological data and models settle deep normative disagreements between the different accounts of moral responsibility for forgetting to do. The data from the psychological experiments on prospective memory cannot tell us what type of general or specific capacity are at the heart of the control required for moral responsibility. Neither can the psychological experiments tell us whether control is required at all. However, they can tell us something about the plausibility of the psychological assumptions about remembering and forgetting made by the different positions. In other words, empirical results cannot tell us what kind of capacities are normatively necessary for moral responsibility but they can tell us something about whether humans are likely to realise the capacities in question.

6. Moral responsibility for forgetting

Experimental data and cognitive models suggest that two factors are important in the remembering and forgetting to do something: 1) Match of the content of the intention with the context, 2) importance assigned to the intention relative to the importance of other intentions. These two factors are not deterministic. Rather, they function to influence the probability of retrieval. So, chance is a third factor that we should take into account when explaining remembering and forgetting.

Mapping onto these three factors (content, importance, and chance) are three different aspects concerning the assignment of moral responsibility for forgetting: 1) control, 2) evaluative attitudes or care, and 3) luck. These different aspects are usually emphasized by competing accounts of moral responsibility. One upshot of the argument of the present paper is that some of these alternative accounts emphasise different factors involved in remembering one's intention. To the extent that these competing accounts make different predictions about the case of forgetting to do, the predictions seem to be equally satisfied. Given that the forgetful agent is blameworthy, the indirect control account makes the normative prediction that the agent had some awareness of the risk of forgetting and the opportunity to strengthen the memory trait (by specifying the content and rehearsing the cue-action association). The valuative account makes the normative prediction that the agent had the required metacognitive awareness and encoding strategies partly as a function of her assessment of importance. These encoding strategies influence the probability of retrieval. In addition, some models of retrieval of intentions also assign the long-term representations a weight related to the relative importance of the intention to the agent. Thus, in addition to encoding strategies, this weight also influences the probability of retrieval (Grünbaum, Oren, & Kyllingsbæk, 2021).

The fact that both control capacities at encoding and subjective importance are involved would explain some of the inherent ambiguity involved in most examples used in the literature, like the case with Sarah. When Sarah forgot to unlock the back door, she might have forgotten because she neglected to encode the intention with a sufficient degree of specificity, because

relative to some other intention the intention to unlock the door was not sufficiently important, or because she was unlucky (or because of some combination of the three). Just from the mere description of the case, we cannot tell which of these explanatory factors were at play.

To know whether Sarah should be blamed for forgetting to unlock the door, it would be natural to engage in conversation with Sarah to enquire whether she had taken sufficient care to articulate the content of the intention at encoding and to assess the degree of importance assigned to the intention relative to other intentions. On the one hand, if Sarah's partner has reasons for thinking that she had the opportunity but failed to exercise her ability to encode the intention with a sufficient degree of specificity, the partner seems to be warranted in blaming Sarah for forgetting to unlock the door. This would be the case if Sarah had never moved beyond the merely verbal and abstract intention to unlock a back door ("yah, sure, I will take care of it"). On the other hand, it might be that Sarah had planned vividly when and how to unlock the door. Nevertheless, due to a persistent pattern of forgetting things she promises to do around the house, the partner seems warranted in believing that Sarah finds household tasks unimportant and that she has assigned the unlocking of the door a low importance. If that were the case, the partner would be right in demanding a justification for this attitude and blaming Sarah for not caring enough about their shared life.

To be sure, the psychological data and models cannot tell us whether control or importance alone are sufficient for blameworthiness. What they can tell us is that either control at encoding or a low level of importance tends to be present in cases like Sarah's where we would intuitively judge that an agent is morally responsible for forgetting. If this is correct, the indirect capacitarian control account and the valuative account might not be in any real conflict when it comes to the psychology of forgetting.

One might worry that this conciliatory attitude misses a fundamental difference between capacitarian control accounts and valuative accounts. Whereas the capacitarian control account is concerned with the agent's moral responsibility for her actions, omissions, and their consequences, the valuative account is concerned with the agent's responsibility for her attitudes. The worry is that the objects of moral assessment for the control account and the valuative accounts differ. The object according to the control account is the action, omission, or their consequences. The object according to the valuative account is the evaluative attitude. They make conflicting claims about the ontological domain of the object of moral assessment.

My reply is that both types of account allow for a wide flexibility with respect to their objects. Let me gesture at two possible ways to develop the reply. If the worry concerns the idea that control-based assignment of blame concerns the agent's actions, omissions, and consequences and valuative-based assignment concerns the agent's mental attitudes, then there are at least two ways of aligning the ontological domains of the targets of blame.

First, one could extend the valuative account to actions and omissions. One could point out that also the valuative account is able to account for a person's resentment or indignation about another person's action or omission by describing the action or omission as an expression of the agent's objectionable attitude. In this way, the agent could be said to be indirectly responsible for her omission. Second, one could shrink the control account to the mental domain. One could point out that also the indirect capacitarian account in the final analysis locates the ultimate target of blame in the agent's mental capacities. In this way, the agent could be said to be directly responsible for the exercise of her mental control ability (Graham, 2017;

Khoury, 2018). Thus, independently of whether the ultimate object of moral responsibility lies within the mental realm or extends to bodily movements or events in the world, the psychological data support the claim that both control factors and importance factors are central to the moral assessment of forgetting.

What about the third factor of chance? How does chance as a factor in the psychology of remembering and forgetting relate to discussions of moral luck? Imagine that Sarah had planned with sufficient detail and cared sufficiently about their shared household. The content was sufficiently specific and the assigned importance higher than for competing intentions, Sarah was just unlucky and another intention ended up being selected by chance. Whether we think that the agent in such a situation is blameworthy for forgetting depends on one's view on moral luck (Nagel, 1979). It strikes me as reasonable that given a high specificity of the content and a high importance assigned to the intention, Sarah should be excused for forgetting to do the task. The point here is not to provide an argument to the effect that when one forgets because of bad luck one is excused. The point here is simply that chance is an aspect of the psychology of remembering and forgetting to do, and that this aspect, just like control of content and importance, maps onto discussions of moral responsibility for forgetting.

7. Concluding remarks

I used the case of forgetting to do to split the direct capacitarian control account, the indirect capacitarian control account, and the valuative account of moral responsibility into their conflicting psychological assumptions. I focused on the assumptions concerning the psychology of remembering and forgetting to do something. The factors (match of content and relative importance of intention) that our best psychological theories of prospective memory describe as being central to the probability of retrieval of standing intentions from long-term memory are also the aspects (control and evaluations) typically identified by conflicting accounts of moral responsibility. Agents seem to have a certain amount of control over remembering and forgetting, but it is a control they have the opportunity to exercise at encoding. When assessing the agent's involvement in her forgetting to do something, we are interested in her encoding strategies and her assignment of importance. These are important psychological aspects we assess when we assess the appropriateness of targeting the forgetful agent with resentment and indignation.

References

Anderson, F., Strube, M., & McDaniel, M. (2019). Toward a Better Understanding of Costs in Prospective Memory: A Meta-Analytic Review. *Psychological Bulletin*, *145*(11), 1053-1081.

Arpaly, N., & Schroeder, T. (2014). *In praise of desire*. Oxford: Oxford University Press.

Björnsson, G. (2017). Explaining (Away) the Epistemic Condition on Moral Responsibility. In J. W. Wieland & P. Robichaud (Eds.), *Responsibility: The epistemic condition* (pp. 146-162). Oxford: Oxford University Press.

Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.

Bugg, J.M. & Streeper, E. (2019). Fate of suspended and completed prospective memory intentions. In Rummel, J., & McDaniel, M. A. (Eds.), *Prospective memory*, pp. 44-59. Abingdon: Routledge.

Cherkaoui, M., & Gilbert, S. J. (2017). Strategic use of reminders in an 'intention offloading'task: Do individuals with autism spectrum conditions compensate for memory difficulties?. *Neuropsychologia*, *97*, 140-151.

Clarke, R. (2014). *Omissions: Agency, Metaphysics, and Responsibility*. New York: Oxford University Press.

Clarke, R. (2017a). Ignorance, Revision, and Commonsense. In J. W. Wieland & P. Robichaud (Eds.), *Responsibility: The epistemic condition* (pp. 233-251). Oxford: Oxford University Press.

Clarke, R. (2017b). Blameworthiness and Unwitting Omissions. In D. K. Nelkin and S. C. Rickless (eds.), *The Ethics and Law of Omissions* (pp. 63-83). Oxford: Oxford University Press.

Cook, G. I., Rummel, J., & Dummel, S. (2015). Toward an understanding of motivational influences on prospective memory using value-added intentions. *Frontiers in human neuroscience*, *9*, 278.

Cyr, T. W., & Swenson, P. (2019). Moral Responsibility without General Ability. *The Philosophical Quarterly*, *69*(274), 22-40.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, *14*(6), 286-290.

Ellis, J., & Milne, A. (1996). Retrieval cue specificity and the realization of delayed intentions. *The Quarterly Journal of Experimental Psychology Section A*, *49*(4), 862-887.

Ellis, J. A., Kvavilashvili, L., & Milne, A. (1999). Experimental tests of prospective remembering: The influence of intention frequency and delay on performance. *British Journal of Psychology*, 90, 9-23.

Fischer, J. M. and Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Fischer, J. M. and Tognazzini, N. A. (2009). The Truth about Tracing. *Noûs* 43: 531–56.

Frankfurt, H. (1971). Freedom of the Will and the Concept of a Person. The Journal of Philosophy, 68(1), 5-20.

Franklin, C. E. (2011). The problem of enhanced control. *Australasian Journal of Philosophy*, *89*(4), 687-706.

Gilbert, S.J. (2015a). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, 68, 971-992.

Gilbert, S.J. (2015b). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, 33, 245-260.

Ginet, C. (2000). The epistemic requirements for moral responsibility. *Philosophical Perspectives*, *14*, 267-277.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54(7), 493–503.

Graf, P. (2012). Prospective memory: Faulty brain, flaky person. *Canadian Psychology/Psychologie Canadienne*, *53*(1), 7- 13.

Graham, P. (2017). The Epistemic Condition on Moral Blameworthiness, A Theoretical Epiphenomenon. In Robichaud, P. and Wieland, J. W. (eds.) *Responsibility: The Epistemic Condition*, Oxford: Oxford University Press.

Grünbaum, T., & Kyllingsbæk, S. (2020). Is Remembering to do a Special Kind of Memory? *Review of Philosophy and Psychology*, *Volume 11, Issue 2*, 385-404.

Grünbaum, T., Oren, F., & Kyllingsbæk, S. (2021). A new cognitive model of long-term memory for intentions. *Cognition*, *215*, 104817.

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology*, *38*(4), 245-256.

Khoury, A. C. (2018). The objects of moral responsibility. *Philosophical Studies*, *175*(6), 1357-1381.

Kool, W., Gershman, S. J., & Cushman, F. A. (2018). Planning complexity registers as a cost in metacontrol. *Journal of Cognitive Neuroscience*, *30*(10), 1391-1404.

Kool, W., Shenhav, A., & Botvinick, M.M. (2017). Cognitive control as cost-benefit decision making. In T. Egner (Ed.), *The Wiley Handbook of Cognitive Control* (pp. 167-189). Chichester, UK: John Wiley & Sons.

Kuhlmann, B. G. (2019). Metacognition of prospective memory. In Rummel, J., & McDaniel, M. A. (Eds.), *Prospective memory* (pp. 60-77). Abingdon: Routledge.

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive psychology*, *48*(1), 47-94.

Levy, N. (2014). *Consciousness and moral responsibility*. Oxford: Oxford University Press.

Levy, N. (2017). Methodological Conservatism and the Epistemic Condition. In J. W. Wieland & P. Robichaud (Eds.), *Responsibility: The epistemic condition* (pp. 252-265). Oxford: Oxford University Press.

Mason, E. (2015). Moral ignorance and blameworthiness. *Philosophical Studies*, 172, 3037–3057.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *14*(7), S127-S144.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Sage Publications.

McGeer, V., and Pettit, P. (2015). The Hard Problem of Responsibility. In D. Shoemaker (ed.), *Oxford Studies in Agency and Responsibility*, Vol. 3, 160–188. Oxford: Oxford University Press.

Meeks, J. T., Hicks, J. L., & Marsh, R. L. (2007). Metacognitive awareness of event-based prospective memory. *Consciousness and Cognition*, 16, 997–1004.

Mele, A. R. (2007). Persisting intentions. *Noûs*, *41*(4), 735-757.

Mele, A. (2010). Moral responsibility for actions: Epistemic and freedom conditions. *Philosophical Explorations*, *13*(2), 101-111.

Murray, S., Murray, E. D., Stewart, G., Sinnott-Armstrong, W., & De Brigard, F. (2019). Responsibility for forgetting. *Philosophical Studies*, *176*(5), 1177-1201.

Murray, S., & Vargas, M. (2020). Vigilance and control. *Philosophical Studies*, *177*(3), 825-843.

Nagel, T. (1979). Moral luck. In his *Mortal Questions*. New York: Cambridge University Press

Nelkin, D. and Rickless, S. C. (2017). Moral Responsibility for Unwitting Omissions: A New Tracing Account. In D. K. Nelkin and S. C. Rickless (eds.), The Ethics and Law of Omissions (pp. 106-129. Oxford: Oxford University Press.

Reese, C. M., & Cherry, K. E. (2002). The effects of age, ability, and memory monitoring on prospective memory task performance. *Aging, Neuropsychology, and Cognition*, *9*(2), 98-113.

Rosen, G. (2004). Skepticism about moral responsibility. *Philosophical perspectives*, *18*, 295-313.

Rudy-Hiller, F. (2019). Give people a break: slips and moral responsibility. *The Philosophical Quarterly*, *69*(277), 721-740.

Scullin, M. K., & Bugg, J. M. (2013). Failing to forget: prospective memory commission errors can result from spontaneous retrieval and impaired executive control. *Journal of experimental psychology. Learning, memory, and cognition*, *39*(3), 965-971.

Scullin, M. K., Einstein, G. O., & McDaniel, M. A. (2009). Evidence for spontaneous retrieval of suspended but not finished prospective memories. *Memory & Cognition*, *37*(4), 425-433.

Scullin, M. K., Kurinec, C. A., & Nguyen, K. (2017). The effects of implementation intention strategies on prospective memory cue encoding. *Journal of Cognitive Psychology*, *29*(8), 929-938.

Scullin, M. K., McDaniel, M. A., & Einstein, G. O. (2010). Control of cost in prospective memory: Evidence for spontaneous retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 190-203.

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: monitoring difficulty or different retrieval processes?. *Journal of experimental psychology. Learning, memory, and cognition*, *36*(3), 736-749.

Shelton, J. T., Scullin, M. K. & Hacker, J. Y. (2019). The Multiprocess framework. In Rummel, J., & McDaniel, M. A. (Eds.), *Prospective memory* (pp. 10-26). Abingdon: Routledge.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, *40*, 99-124.

Sher, G. (2009). *Who Knew? Responsibility without Awareness*. New York: Oxford University Press.

Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics*, *121*(3), 602-632.

Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, *115*(2), 236-271.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 347.

Smith, R.E. (2016). Prospective memory: A framework for research on metaintentions. In J. Dunlosky and S. Tauber (Eds.) The Oxford handbook on Metamemory. New York: Oxford University Press.

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A formal modeling approach. *Experimental Psychology*, *52*(4), 243-256.

Spreng, R. N., Madore, K. P., & Schacter, D. L. (2018). Better imagined: Neural correlates of the episodic simulation boost to prospective memory performance. *Neuropsychologia*, *113*, 22-28.

Sripada, C. (2015). Moral responsibility, reasons, and the self. In D. Shoemaker (Ed.), *Oxford studies in agency and responsibility* (Vol. 3, pp. 242–264). Oxford: Oxford University Press.

Swenson, P. (2016). Ability, foreknowledge, and explanatory dependence. *Australasian Journal of Philosophy*, *94*(4), 658-671.

Talbert, M. (2016). *Moral responsibility: An introduction*. John Wiley & Sons.

Talbert, M. (2017). Omission and Attribution Error. In D. K. Nelkin and S. C. Rickless (eds.), *The Ethics and Law of Omissions* (pp. 17-35). Oxford: Oxford University Press.

Vargas, M. (2005). The trouble with tracing. *Midwest studies in philosophy*, *29*, 269-291.

Vargas, M. (2013). *Building Better Beings: A Theory of Moral Responsibility*, Oxford: Oxford University Press.

Vargas, M. (2020). Negligence and social self-governance. In A. Mele (ed.), Surrounding Self-Control (pp. 400-420). Oxford: Oxford University Press.

Walter, S., & Meier, B. (2014). How important is importance for prospective memory? A review. *Frontiers in psychology*, *5*, 657.

Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, *24*(2), 227-248.

Zimmerman, M. J. (1997). Moral responsibility and ignorance. *Ethics*, *107*(3), 410-426.