

**THE PLACE OF THE TRACE**  
***Negligence and Responsibility***

Samuel Murray, Duke University\*

*(forthcoming in The Review of Philosophy & Psychology)*

**Abstract**

One popular theory of moral responsibility locates responsible agency in exercises of control. These control-based theories often appeal to tracing to explain responsibility in cases where some agent is intuitively responsible for bringing about some outcome despite lacking direct control over that outcome's obtaining. Some question whether control-based theories are committed to utilizing tracing to explain responsibility in certain cases. I argue that reflecting on certain kinds of negligence shows that tracing plays an ineliminable role in any adequate control-based theory of responsibility.

*0. Introduction*

Imagine that the lead news story describes a man who leads police on a high-speed chase down the highway while intoxicated. So intoxicated, in fact, that his BAC levels are three times the legal limit at the time of arrest. Before he is stopped, the man nearly runs over two pedestrians and rams several police cars. The story closes with the criminal charges: aggravated assault, DUI, fleeing the scene, all of which add up to the prospect of serious jail time.

The man's behavior might provoke shock, though the criminal charges do not. Intoxicated drivers, particularly those that injure others or damage property, are often punished in one way or another. On reflection, this could seem odd. After all, the man (at the time of his driving) is not really aware of what he is doing and does not really possess

---

\* Thanks to Manuel Vargas, Juan Pablo Bermúdez, and Santiago Amaya for extensive discussion and comments on different versions of this project. Research for this project was supported by the Stanford University Neurobiology Department, as well as grants from Florida State University through the Philosophy and Science of Self-Control project and the John Templeton Foundation's "Getting Better at Simple Things: understanding and improving vigilant control" project. The views expressed in this chapter are my own and do not necessarily reflect the views of the Templeton Foundation, Florida State University, or Stanford University.

any self-control. So in virtue of what is this individual rightly blamed and accused? What accounts for the man's responsibility for his actions?<sup>1</sup>

One intuitive line of thought is that the man is responsible for his behavior in virtue of being responsible for his earlier decision to drink to excess (cf. Vargas, 2005). Drinking heavily partially explains the subsequent drunk driving. Thus, insofar as the man possesses self-control and awareness at the time of beginning to drink heavily the man is responsible for any later harms or damages that he brings about in his incapacitated state.

This explanatory strategy is called *tracing*, and responsibility theorists have a long history of appealing to the concept of tracing to explain the responsibility of agents like drunk drivers.<sup>2</sup> In general, theorists use tracing to explain why agents that lack control are still responsible for what they do by connecting (i.e., tracing) lack of control back to some earlier possession of control that explains the later lack of control.<sup>3</sup>

Tracing is central to many theories of responsibility, particularly those theories that locate responsible agency in the possession of control. Recently, however, critics of tracing have raised a challenge that calls into question whether tracing principles can do all the theoretical work assigned to them. For example, suppose that the captain of a ship

---

<sup>1</sup> My use of 'responsibility' refers to moral responsibility as opposed to causal, legal, or other forms of responsibility. The sense of moral responsibility at issue here is the accountability sense rather than the attributability and answerability senses of responsibility (cf. Shoemaker, 2015). A culpability-imputing form of blame, then, is central to the sense of moral responsibility discussed here. I focus on moral blameworthiness (as opposed to praiseworthiness or creditworthiness), where moral blameworthiness picks out a range of judgments and attitudes bound up in non-trivial ways to the Strawsonian reactive attitudes of resentment, indignation, and guilt (Strawson, 1962). Lastly, I look only at retrospective moral responsibility, or responsibility for what one has done.

<sup>2</sup> Many theorists rely implicitly or explicitly on some notion of tracing: Vargas (2013); Smith (1983); Ekstrom (2000); van Inwagen (1989); Fischer and Ravizza (1998: 50); Dennett (1984: 13); Kane (1996: 39); Ginet (2000); Fischer and Tognazzini (2009); Timpe (2011); Audi (1991). Notable exceptions to this trend include Harry Frankfurt's (1998) later work on responsibility, Mason (2015), Scanlon (1998, 2008), and Angela Smith (2005, 2008, 2015). This is only a sample of the relevant literature.

<sup>3</sup> The explanatory relation between some tracing anchor and the tracing terminus does not need to be causal.

is piloting her boat on a routine journey. Because the route is so familiar, she inadvertently slips into a daydream and, as a result of this, crashes into some submerged rocks.<sup>4</sup> Here, the captain seems responsible for the crash despite the fact that she did not really possess control over the process that resulted in the crash (after all, it seems odd to say that the captain controls whether she slips into a daydream or not). There is no past moment at which she decides or intends or chooses or wills to either daydream or put herself in a position where she is likely to crash the boat.

In the abstract, the challenge stems from cases that describe some (relatively) non-deliberative agent doing something for which she is responsible despite the fact that no earlier possession of control explains the subsequent non-deliberative state.<sup>5</sup> We can see this in the example of the daydreaming captain. In these cases, there seems to be no suitable moment in the agent's moral history that might anchor a trace and so no plausible explanation of responsibility in terms of tracing. Still others have noted that tracing cannot account for responsibility for intrinsically involuntary states like emotions or desires.<sup>6</sup> Again, the thought is that no prior possession of self-control or awareness explains the agent's subsequent possession of these sorts of attitudes.<sup>7</sup>

This represents a serious challenge to tracing. Advocates of control-based theories of responsibility introduced tracing to explain responsibility in cases where the agent seems responsible for some action or outcome despite failing to possess control at the time

---

<sup>4</sup> This case is adapted from Sher (2009: 24).

<sup>5</sup> See Vargas (2005: 275). Cf. McKenna (2008).

<sup>6</sup> A state is intrinsically involuntary only if the agent that is the subject of that state cannot directly control whether the state obtains.

<sup>7</sup> And, insofar as possession of these sorts of attitudes is sometimes morally objectionable, tracing seems unable to furnish an explanation for the appropriateness of this sort of moral assessment. See Smith (2015: 118-20).

of the action or outcome.<sup>8</sup> There are, however, cases where tracing seems unhelpful despite the fact that the cases seem to be those that tracing was meant to explain. The failure of tracing in these cases gives us reason to think that introducing tracing principles into control-based theories of responsibility is misguided.

There are four possible responses to this challenge. Let's focus on the example of the drunk driver mentioned at the beginning to see how these responses divide up. The first is to embrace skepticism about responsibility in cases that seem to require appeals to tracing.<sup>9</sup> The upshot of this approach is that we must be skeptical of the claim that the drunk driver is responsible for his actions.<sup>10</sup> This is unattractive in part because the driver seems responsible and appeals to tracing seem necessary to explain the responsibility of the drunk driver. That's not a decisive consideration against the skeptic, but skepticism should, in general, be an option of last resort.

The second possible response is to claim that appeals to tracing are a mistake and that the failure of tracing points to an underlying problem with the control-based theory of responsibility that tracing was meant to supplement. Thus, the drunk driver is responsible, but not because he possesses or displays some level of control over his behavior or decisions; rather, the driver is responsible because he exhibits some morally

---

<sup>8</sup> Theorists generally appeal to tracing in the context of blameworthiness, though the concept can also be utilized to explain distinctive kinds of praiseworthiness. In this paper, I confine myself to discussing tracing as it applies to blameworthiness, though I note this extra dimension of tracing as an important (if underappreciated) aspect of the concept. For some discussion of this other dimension of tracing, see Fischer and Ravizza (1992); McDowell (1978, 1979); van Inwagen (1989, 1994).

<sup>9</sup> Three recent examples of this approach are Rosen (2004), Zimmerman (2008), and Levy (2009).

<sup>10</sup> The skeptic can pursue one of two lines here. Either we might be skeptical that our judgments that some individual (like the drunk driver) is responsible are never justified *or* we might be skeptical that some individual is actually responsible. In the former case, the drunk driver might be responsible though we could never have a justified belief that this is the case. In the latter case, the drunk driver might not be responsible at all. None of this, however, directly threatens the system of legal sanctions and punishments mentioned at the outset. One could imagine a society where people adopt widespread epistemic or metaphysical skepticism about *moral* responsibility that does not extend to the *legal* responsibility of wrongdoers.

objectionable attitude that makes appropriate certain forms of moral evaluation.<sup>11</sup> On this approach, we give up the need for appeals to tracing because we give up the need to anchor responsibility in control. While this approach has been developed and refined in a number of interesting and subtle ways, I will not consider it here. The merits and shortcomings of this approach require separate treatment.<sup>12</sup>

The third possible response is to agree that appeals to tracing are mistaken, but to retain the centrality of control in explanations of responsibility. The motivation for this is that appeals to tracing belie an ill-considered or flat-footed theory of action. Given an adequate understanding of the nature of temporally extended action and the distribution of intentionality through actions embedded in some larger plan, we can explain the responsibility of the drunk driver and the daydreaming captain without appealing to tracing.

The fourth possible response is for advocates of control-based theories of responsibility to face the challenges to tracing head on. This response requires, in part, adequate explanations of responsibility that utilize tracing in cases like the daydreaming captain mentioned above. The motivation for this response stems from a commitment to the idea that tracing is essential for control-based theories of responsibility.

Matt King has put forward the most recent argument for the third response. King argues that an adequate control-based theory of responsibility need not incorporate any appeals to tracing to account for responsibility in cases where an agent apparently lacks control.<sup>13</sup> The advantage of King's position is that it allows a control-based theory of

---

<sup>11</sup> See note 2 for references.

<sup>12</sup> To get a taste of the shortcomings, see King (2009: 583-87) and Vargas (2013: 142-46).

<sup>13</sup> King (2014).

responsibility to retain its explanatory scope without appealing to any theoretically problematic tracing principles.

The argument for this position has a simple structure. One can explain an agent's responsibility, he claims, in paradigmatic 'tracing cases' (like the case of the drunk driver) without utilizing any appeals to tracing. These tracing cases can be modeled and evaluated either as instances of reckless or negligent behavior.<sup>14</sup> An agent is reckless with respect to some form of behavior if the agent knowingly engages in the behavior with an awareness of the risks associated with that behavior. An agent is negligent with respect to some form of behavior if the agent engages in the behavior without any suitable awareness of the risks associated with that behavior. For example, suppose someone throws a baseball-sized rock over the edge of a cliff while she is hiking. If the agent is aware of some risk that the rock might hit someone, then the agent is reckless in throwing the rock. If the agent is not aware of any such risk and there is some risk, then the agent is negligent in throwing the rock.<sup>15</sup> The explanation of responsibility in either case does not require tracing, so an adequate control-based theory of responsibility does not require a tracing principle.

King's argument, then, relies on three claims:

- (1) We can model tracing cases exhaustively as instances of recklessness or negligence.
- (2) If we model tracing cases as instances of recklessness, then tracing is superfluous.
- (3) If we model tracing cases as instances of negligence, then tracing is unhelpful.

---

<sup>14</sup> King (2014: 465-66).

<sup>15</sup> This characterization is, admittedly, rough. For instance, negligence generally presupposes that the potential risks are such that the agent should have known about the risks or that there is some reasonable expectation that the agent be aware of the associated risks.

For King's argument to work, it must be the case that *every* instance of responsibility for either reckless or negligent behavior can be explained without appeals to tracing.

In what follows, I target the third premise of this argument. I claim that the argument that explanations of responsibility for negligence do not require tracing wrongly assumes that explanations of responsibility in these cases have a uniform structure. I argue that there are at least two different kinds of negligence, and explanations of responsibility for one kind require tracing.

Responding to King's argument provides a way to generalize the account of tracing offered here in two ways. First, the account of tracing indicates a wider range of cases than drunk drivers where tracing is needed to explain responsibility. Second, the argument shows that tracing is indispensable for a control-based theory of responsibility, in part because tracing is needed to explain certain instances of responsibility for negligence. Generalizing the account in this way provides an occasion to consider why our temporally extended agency leaves us susceptible to negligence and why this makes tracing a fundamental feature of any control-based responsibility theory.

### 1. *The control-based theory of responsibility*

Before proceeding to King's argument, we should get clear on the theoretical commitments of a control-based theory of responsibility. A control-based theory of responsibility locates responsible agency in the possession and exercise of agential capacities implicated in the production of effective self-governed action and the rational governance of attitude formation and cultivation.<sup>16</sup> Normally, control-based theories

---

<sup>16</sup> Cf. Vargas (2013: 203). Here, responsible agency *just is* the possession of capacities constitutive of effective and rational self-governance.

countenance two kinds of capacities as constitutive of this morally significant self-governance.<sup>17</sup> The first are those capacities that enable agents to exhibit self-governance. Self-governance, broadly construed, picks out the activity of mechanisms and processes that contribute to engaging in goal-directed behavior. This includes, but is not limited to, the capacity for behavior regulation, the ability to adopt and update goals, and the coordination of plans over time. The second set of capacities constitutive of control are those that afford agents awareness of their moral environment (foresight, means-end reasoning, etc.). Thus, control picks out agents that exhibit suitable degrees of self-governance and awareness.

Control-based explanations of responsibility pick out some suitable relation (at the moment of wrongdoing) between an agent's responsibility-relevant capacities that are constitutive of control and some bad action or outcome as a necessary condition for ascribing responsibility for that action or outcome.<sup>18</sup> These explanations are attractive because they pick out a relation that ties the agent to her wrongdoing in a way that makes it appropriate to blame the agent for her wrong conduct.<sup>19</sup>

What makes drunk driving cases paradigmatic instances of tracing cases is that in cases of drunk driving there appears to be no immediate relation between the agent's responsibility-relevant capacities and the wrong conduct. While drunk, the agent is locally incapacitated in a way that disrupts connecting the agent (via her capacities) to her bad conduct, where local incapacitation consists in some (but not all) agential capacities being masked or thwarted temporarily. This highlights the need for tracing: without it, we

---

<sup>17</sup> Vargas (2013: 200). Some of my terminology diverges from that of Vargas.

<sup>18</sup> In what follows, I will focus exclusively on behavior that seems to merit distinctively negative moral reactions. This is because King focuses on these kinds of cases and others in the literature do the same. As I noted before, however, there are interesting questions about the way in which advocates of control-based theories can utilize tracing to explain distinctive forms of praiseworthy behavior.

<sup>19</sup> See Sher (2009: 73-75).



would have no way to connect agents of this sort to their bad conduct and thus no obvious means of explaining why such agents are responsible for their bad conduct.

The goal of the non-tracing control-based theory of responsibility that King defends is to offer some account of responsibility in these paradigmatic tracing cases without utilizing appeals to tracing. This retains the basic contours of a control-based theory of responsibility while jettisoning any theoretically costly or implausible machinery like tracing. In the next section, I consider King's discussion of negligence and offer some criticism.

## 2. *Modeling Negligence: Regress and Utility*

Negligence can be defined in terms of the unwitting risk of wrongdoing. An agent is negligent when the agent behaves in a way that greatly increases the risk of failure or wrongdoing without ever realizing that she is at risk of failure or wrongdoing until it is too late. What distinguishes cases of recklessness from negligence is that in the latter case the agent *unwittingly* risks wrongdoing and is never aware of this risk until it is too late. Acting negligently (or being negligent with respect to some action) consists in acting in such a way that risks failing to act successfully (e.g., driving negligently consists in driving in such a way that one unwittingly increases the risk of crashing).

King suggests that negligent agents are sometimes responsible for their negligent wrongdoing, though the explanation of responsibility for negligence cannot invoke tracing.<sup>20</sup> This is because tracing explanations of responsibility for negligence either devolve into an infinite regress or add unnecessary theoretical machinery. I explain each of these points in turn before offering some criticism.

---

<sup>20</sup> King (2014: 474).

Tracing explanations of negligent wrongdoing devolve into infinite regress in the following way. Suppose that Pete crashes while driving drunk. Further, suppose that Pete is so drunk that he lacks control at the time just before the crash. Pete is responsible for crashing while drunk (he was not, let's suppose, coerced or manipulated into drinking too much). When we try to apply a trace to Pete, however, we run into a problem. Pete has been negligent, so, by definition, there is no prior moment where Pete possesses suitable awareness of the harms likely to result from his risky behavior. And, because he lacks awareness, Pete also lacks control at all relevant prior moments. So, when we try to draw the trace to explain Pete's responsibility for negligence, we find no suitable terminus for the trace.

The only prior moments available that might anchor a trace are those where Pete *should* have, but failed to, exhibit more control. In these moments, however, Pete is also negligent. For example, we might try to trace responsibility for crashing back to some earlier choice point during the night. Pete, however, also lacks awareness in these earlier moments, so Pete's decision to drink heavily (or continue drinking) is also negligent. Thus, we have to apply a trace to Pete's negligent decision to drink to explain his responsibility for that negligent decision.<sup>21</sup> The regress, at this point, becomes clear.

Suppose that we *could* formulate an explanation of the responsibility for the negligent decision that both explains the responsibility for the later crash *and* does not lapse into an infinite regress.<sup>22</sup> We might be tempted to think that tracing is helpful in this

---

<sup>21</sup> King (2014: 474).

<sup>22</sup> King thinks it is not possible to explain responsibility for negligence in terms of control (see King, 2009). The point of the futility argument against tracing is that, *per impossible*, were someone to formulate a tracing explanation of responsibility for negligence, there would no longer be any need to invoke tracing.

scenario because we could explain Pete's responsibility for crashing in terms of his negligent decision-making earlier at the bar.

But this will not work. Consider the negligent behavior ( $N_1$ ). Either we can or cannot explain responsibility for  $N_1$  in terms of control. The tracing theorist assumes we cannot (which, recall, is part of the motivation for introducing tracing principles in the first place). So, we must find an earlier negligent moment ( $N_2$ ) to anchor responsibility for  $N_1$ . Either we can explain responsibility for  $N_2$  in terms of control or we cannot. If we *can*, then King claims we can apply the explanation of responsibility for  $N_2$  to  $N_1$  without ever needing to invoke tracing. If we cannot, then we need to find an earlier negligent moment ( $N_3$ ), and at this point the regress again becomes clear. Thus, even if one could solve the regress problem, tracing still has no role to play because tracing becomes unhelpful.<sup>23</sup>

The problem with King's suggestion to transfer the explanatory schema for negligent decision-making to negligent behavior is that there might be qualitatively distinct kinds of negligent failures. Some negligence occurs in virtue of having no awareness of the moral qualities of one's behavior or the likely consequences thereof. Other instances of negligence occur in virtue of having substandard awareness of one's behavior and the moral environment. The difference between lack of awareness and substandard awareness turns on the possession of some epistemic capacity (or capacities) that enable an agent to appreciate certain considerations and act in light of those considerations. When an agent possesses the appropriate epistemic capacities but fails to exercise them appropriately in a case of negligent wrongdoing, the agent possesses substandard awareness. When an agent does not possess these capacities in a case of negligent wrongdoing (either because something thwarts the proper functioning of these

---

<sup>23</sup> King (2014: 475).

capacities or the agent simply does not possess the requisite capacities), then the agent possesses no awareness.

Take the example of Pete. When he is sitting in the bar deliberating about whether to drink (or continue to drink), he possesses the requisite epistemic capacities to appreciate reasons to stop drinking (or, at least, to drink more moderately) but fails to appropriately exercise these capacities. Here, Pete possesses substandard awareness.<sup>24</sup> Contrast this substandard awareness with the total lack of awareness that Pete possesses moments before the crash. At the time of the crash, Pete might possess no awareness of what he is doing or why he is doing it.

The upshot, then, is that we can explain responsibility for negligent action (an instance where an agent should have been aware of some risk of harm, but could not have been aware) by appealing to some negligent decision (where the agent both should have and could have been aware of the risk of harm). This blocks the regress concern that King raises because we can, presumably, explain responsibility for some negligent decisions directly without appealing to prior choice points (cf. Clarke, 2014: 167 and Murray, 2017 for two proposals on direct responsibility for negligence). However, this also makes clear why we cannot simply apply the explanatory schema for the negligent decision directly to the negligent behavior. A crucial part of explaining responsibility for the negligent decision is the presence of certain unexercised epistemic capacities that the agent *could* have exercised, and these capacities ground the explanation of responsibility.

---

<sup>24</sup> There are a number of ways to analyze what substandard awareness consists in. For example, it could be that Pete has, in some sense, all of the information necessary to correctly infer that he should stop drinking and he fails to perform the crucial inference. Or it could be that he is simply unaware of some fact that he needs to draw the appropriate conclusion about what ought to be done in the circumstances. Or Pete might only be aware of potential risks of his current conduct at a level of granularity that is too coarse-grained. In any case, the crucial point here is that when Pete possesses substandard awareness, he has the ability to correctly evaluate the circumstances but fails to exercise this ability.

In the cases of negligent behavior at issue here, the agent is locally incapacitated, so there is no explanation of responsibility in terms of available, unexercised capacities.

At this point, one might wonder why we should bother with substandard awareness at all. Someone might object that agents either have awareness or they don't. In the former case, there is recklessness; in the latter case, negligence. There is no middle category of partial or substandard awareness. And even if there were, it doesn't seem like the *substandard* aspects of the awareness are doing any explanatory work.

Substandard awareness, however, is useful because it picks out two distinct phenomena, both of which relate to negligence in interesting ways. Pete is aware of something (the color of his beer or the music in the bar, perhaps), but he is not aware of any risk of harm. What makes Pete's awareness substandard is that he possesses an ability to become aware that he fails to exercise. However, we can imagine alternative cases of negligence where the negligent agent's awareness is constituted by attitudes with extremely coarse-grained content. These alternative cases suggest the additional utility of substandard awareness.

Suppose that Pete *does* have some thought while sitting at the bar that something bad might occur, but the thought does not have content that is any more specific. In this case, it does not seem like Pete has the right kind of awareness to constitute recklessness (or, he does not have the kind of awareness to make appropriate the description of subsequent behavior as reckless). There is a certain granularity of awareness at which one fails to foresee the occurrence of a certain event. In fact, I would venture a guess that most people fail to pick out many of the foreseeable upshots of their plans with fine-grained descriptions. Pete is one such individual, and the description of his awareness as 'substandard' seems apt. Thus, we can apply the substandard awareness label to instances

where agents have coarse-grained attitudes that fail to connect these agents to the foreseeable upshots of their behavior in recklessness-constituting ways.<sup>25</sup>

### 3. *Awareness and allocation*

Thus far, the account of negligence developed here has depended on notions like substandard awareness, incapacitation, and epistemic capacities. These can be filled out in empirically tractable ways, and doing so lends some plausibility to the idea that some instances of negligence result from substandard awareness. This, in turn, explains why substandard awareness sometimes generates local incapacitation, thereby setting up the need for tracing-based explanations of responsibility.

Recent work on the psychology and neuroscience of decision-making suggests that people implement cognitive strategies when performing tasks over time. These strategies consist in differential allocations of psychological resources over time (Braver, 2012). Some strategies are more proactive, where these involve sustaining goal representations in working memory continuously to bias attention, perception, and action systems toward goal-relevant processing (Miller & Cohen, 2001). Other strategies are more reactive, where these involve transient, bottom-up reactivation of goal representations at the

---

<sup>25</sup> Someone might think that Pete, even when he possesses this coarse-grained awareness, is still reckless. Awareness is awareness, and if it's present, then the agent is reckless. This objection, however, leads to a problem. If we think that substandard awareness (as I've characterized it above) is not negligence-constituting, then we might eliminate negligence as an interesting category of action. Every minimally conscientious person has, at some point, had the thought that something bad will likely come about through future exercises of one's own agency. If we think of awareness, no matter how coarse-grained, as recklessness-constituting, does this awareness distribute pervasively, such that almost all future acts of wrongdoing are reckless? If not, in what way do we block the distribution of this substandard awareness in a principled way? This problem suggests that we take substandard awareness as negligence-constituting (of course, there is a further issue of what the standards of content granularity are, but that's a familiar debate from action theory and substantive disputes about the nature of foresight).

moment of action, relying on spontaneous recall or bottom-up associative cues to trigger goal implementation (Jacoby et al., 1999).

Strategy selection is a form of managing psychological resources. Given limited psychological resources, people need to flexibly deploy different resources to different tasks at different times. Since people cannot pay attention to everything all the time, they must choose where, when, and how intensely to deploy attention. This gives us an empirical way to explain the epistemic capacities implicated in responsible agency and the foregoing account of negligence. The relevant epistemic capacities are whatever is realized by the systems that compute and implement cognitive strategies for resource allocation.

Similarly, we can understand awareness and substandard awareness in terms of allocation. People can focus either proactively or reactively on whatever they're doing. The allocation of underlying attentional and mnemonic resources constitutes an agent's awareness. Substandard awareness results from allocating insufficient resources to a task or implementing a faulty strategy. For example, someone who lets their mind wander while driving busy city streets exhibits substandard awareness, because the city environment is not conducive to reactive task engagement. Hence, the allocation of psychological resources is substandard in virtue of greatly raising the probability that something bad will happen (e.g., an accident).

Strategy selection systems, and the inputs to which these are sensitive, may seem to be subconscious and subpersonal. However, some of the considerations that determine strategy selection rise to the level of consciousness. In this way, strategy selection manifests a unique form of reasons-responsiveness. Broadly speaking, strategy selection is responsive to perceived task factors and individual differences variables (Braver, 2012: 108-10).

Recent computational work has identified components of the cost/benefit analyses that determine strategy and which variables differentially influence these analyses. Overall, the core variables are the importance of the task (Savine & Braver, 2010), the probability of successfully performing a task relative to a particular allocation (Kool, Gershman, & Cushman, 2017), the anticipated workload of the task (Speer et al., 2003), the anticipated level of distraction in the task environment (Burgess & Braver, 2010), and the opportunity costs associated with focusing on one task rather than others (Inzlicht, Shenhav, & Olivola, 2018; Kurzban et al., 2013). These can be grouped into three overarching factors: (1) the level of expected effort needed to perform the task well; (2) the value of performing the task, and; (3) the opportunity cost of focusing on task performance (Shenhav, Botvinick, & Cohen, 2013; Botvinick & Cohen, 2014).

These factors provide a schema for identifying different kinds of negligence. Some negligence derives from motivational deficits. For example, someone might neglect to prepare for their upcoming lectures because they don't care much about teaching. This lack of motivation results in insufficiently allocating resources to teaching preparation. But not all negligence derives from motivational deficits. Consider the ship captain mentioned in the Introduction. She, presumably, cares about the ship and the safety of the passengers (or at least herself!). In this case, negligence might derive from being overly confident in being able to successfully navigate without thinking hard about piloting (this, in turn, might result from travelling a familiar route).<sup>26</sup> Finally, some negligence might derive from underestimating the effect of a current allocation policy on the ability to make future allocations. For example, suppose someone stays up late playing video games

---

<sup>26</sup> Of course, it's entirely possible that the negligence derives from motivational deficits. The point here is that not all negligence reflects lack of motivation.



despite having a job interview in the morning. This could result from underestimating the effect of staying up late on performance in the interview. In each of these cases, poor allocation policies result from different causes that reflect the computational elements of strategic decision-making.

Sometimes, a particular allocation of psychological resources can result in downstream local incapacitation. Pete's drunk driving is an example of this (as is the aforementioned negligent teaching prep, the negligent piloting, and negligently under-sleeping). Though the terminology is ham-fisted, the idea is that Pete pursues drinking at one point in a way that makes him incapable of properly allocating his psychological resources in a manner conducive to pursuing his goals here and now (i.e., driving home safely). However, these bits of local incapacitation should not always excuse wrongdoing, and we need some way to tie the agent to their wrongdoing even when locally incapacitated. Tracing enables the transfer of responsibility in these cases.

Thinking of negligence in terms of psychological resource allocation also expands the range of cases where tracing is useful. Consider, for example, a case of procrastination. Suppose that someone needs to submit a manuscript, finish a course syllabus, write some letters of recommendation, and submit approval to bring a proposal in front of the faculty senate. There are ways to pursue these different goals over time that puts one in a position where it's practically impossible to complete all these different goals. Procrastinating is a way of allocating one's resources such that one becomes (locally) incapable of allocating one's resources in a way that enables successful pursuit of one's goals. Responsibility for, say, failing to write your recommendation letters or finish your manuscript is explained in terms of implementing a strategy that insufficiently allocated psychological resources.

This provides a better picture of when tracing is useful for explaining responsibility for negligence. When people commit to performing a variety of tasks, they typically commit (implicitly) to completing those tasks within a certain timeframe. They also (typically implicitly) agree to be governed by certain norms and standards of information management, keeping track of what has been done and what must still be done with respect to each task (see Murray & Vargas, Forthcoming for discussion of these norms of information management). People can be negligent with respect to their tasks when they insufficiently allocate psychological resources to those tasks or implement strategies that are unlikely (and known to be unlikely) to successfully balance plural goal pursuit. On the occasions when this leads to local incapacitation (and subsequent negligent wrongdoing), the negligent wrongdoing remains culpable in virtue of being explained in terms of the earlier allocation or faulty strategy.

Notice that Pete's case fits this scheme. There's a negligent decision to drink, where the negligence likely derives from improperly attending to the risks of drinking. This results in local incapacitation and Pete's drunk driving. The responsibility for the drunk driving is rooted in the responsibility for the earlier negligent decision to drink. So, supposing that we formulate some non-regressive explanation of Pete's responsibility for his negligent decision to drink, the appeal to tracing will be helpful (and probably essential) to explaining Pete's responsibility for drunk driving. Tracing can do real theoretical work without slipping into a regress.

#### *4. Ineliminable Tracing*

The argument to this point shows that there are cases of negligence where an agent's negligent behavior results from local incapacitation. Responsibility for this negligent

behavior is rooted in prior substandard awareness (insufficient psychological resource allocation) or faulty strategy selection that leaves open the (relatively high) probability of goal neglect. Tracing allows us to keep negligent wrongdoers on the hook, so to speak, despite their local incapacitation.

This shows that the third premise in King’s argument is wrong. Not every instance of culpable negligence can be explained without tracing. The previous two sections show that tracing explanations of responsibility can be formulated in a useful, non-regressive way. In this final section, I generalize the above argument to show that tracing is an indispensable part of any control-based theory of responsibility. This is because the need for tracing derives from our nature as temporally extended, computationally bounded agents.

Human beings adopt a wide range of commitments, projects, and goals that they aim to complete over varying time scales. Part of the reason for this is evolutionary: our earliest ancestors needed to find food, build shelter, and avoid predators. Increased social organization modulates this plural goal pursuit. Today, apt social functioning requires occupying various roles, and each role is associated with different expectations and goals. Occupying more than one goal requires plural goal pursuit.

Human beings are also inherently limited creatures. Some of this derives from anatomy. People cannot walk in two different directions simultaneously. But some of this derives from our cognitive architecture. People, for instance, cannot perform two two-digit math problems simultaneously. Many of our psychological capacities—such as attention, working memory, and cognitive control—are severely limited (Anderson & Milson, 1989; Bundesen, 1990; Cohen, 2017). These limits establish obstacles to

successful plural goal pursuit. Roughly, people are trying to do a lot of things with precious few resources.

To reconcile plural goal pursuit with limited psychological capacity, human beings scatter their agency over time. Scattered agency consists in the ability to variably apportion an action plan over time that enables efficient plural goal pursuit while compensating for limited psychological resources. In other words, scattered agents can start a project, put it aside for awhile while doing something else, all while remembering to come back and finish at an opportune time.<sup>27</sup>

The possibility of negligence arises from scattered agency. Because we scatter our actions through time, it is possible to never finish something that one started. It is possible to plan to get around to something only to never actually get to it. It is possible to make a commitment and never follow through. Scattered agency makes negligence possible because only scattered agents switch between tasks that they have not yet completed or shift focus away from their goals and commitments to focus on other things.

Consider the drunk driver mentioned in the Introduction. Earlier in the night, he probably made a plan to get some drinks and then drive home. But when out at the bar, there's no need to focus on driving home later; instead, one focuses on conversing with friends. Without the goal of driving home firmly in mind, it's easy to order a few more drinks than you originally planned. Perhaps the thought of driving home just doesn't cross the mind. More likely, with the additional drinks comes overconfidence at one's

---

<sup>27</sup> The notion of scattered agency derives from the notion of a temporally scattered event (see Sorensen, 1985). This is distinct from a spatially scattered event (e.g., the broadcasting of Welles' *War of the Worlds* performance), where the parts of the event occur simultaneously over a large region (Lewis, 1983). An action counts as a temporally scattered event when the time of the action has parts that exist at different times and these times do not necessarily overlap with the time of the action. Scattered agents, then, are capable of acting in ways that consist in temporally scattered events.

ability to drive while intoxicated. Thus starts the cycle of negligence, which begins with shifting focus away from driving home.

Perhaps there's a simple solution to negligence, namely always to stay focused on all of your goals. But that's not possible with the limited psychological resources that humans possess. Perhaps people should form more modest goals. But a moment's reflection shows that many of the meaningful and rewarding activities of life are fluid in a way that opens up the possibility of negligence. There are few activities that people perform all in one go, and limiting oneself only to such activities makes for a relatively mundane existence.

To summarize, then, scattered agency opens up the possibility of negligence. But scattered agency itself appears to be the only efficient solution to navigating the tension between the kinds of lives we want to lead (plural goal pursuit) and the kinds of creatures we are (psychologically limited). Negligence is a quirky upshot of the scattered agency we exhibit.

As we saw, some kinds of negligence can, in the right circumstances, generate local incapacitation. Poorly allocating one's psychological resources in acting over time can generate local inability to properly allocate psychological resources to various commitments. Procrastination is a vivid example of a poor allocation strategy that can lead to local incapacitation. This local incapacitation should not, however, function as a general excuse for wrongdoing. Thus, some negligent wrongdoing is culpable *even when* such wrongdoing results from incapacitation.

When we scatter our agency, our actions consist in temporally scattered events. There is temporal discontinuity between performances of the same activity (e.g., writing a manuscript over many years). In the abstract, one can set out to A, switch over to doing

B, and leave A-ing unfinished. In order to be responsible for one's negligence with respect to A-ing, there needs to be some way to transfer responsibility through these temporal gaps. For a scattered agent, the stream of agency is continuously disrupted and diverted in pursuing various goals over time. Tracing enables distributing responsibility through these temporal gaps to preserve responsibility.

At the outset, I noted that control-based theories of responsibility locate responsible agency in the possession of capacities constitutive of self-governance and awareness. We now see that these capacities are also those that constitute scattered agency. The kinds of capacities central to responsibility are just those capacities that we manifest in ordinary exercises of agency. This indicates that tracing is a fundamental feature of control-based theories of responsibility (or, those theories of responsibility that describe creatures like us). The possibility of negligence follows from scattered agency, the kind of agency that compensates for pursuing a plurality of goals under conditions of bounded rationality. Explaining responsibility for such negligence sometimes requires tracing. So, a theory of responsibility that hinges responsibility on particular manifestations of agential capacities requires tracing to explain responsibility for negligence.

Offering a control-based theory of responsibility that does not include a tracing principle seems implausible, and reflecting on the reasons why reveals something interesting about the role that tracing plays in control-based theories of responsibility. Tracing is necessary to explain why scattered agents are responsible for negligence even when the parts of one's negligence are temporally discontinuous. Tracing, then, provides a plausible way to distribute responsibility over the career of a scattered agent.

## REFERENCES

- Anderson, J.R. and Milson, R. 1989. "Human memory: An adaptive perspective," *Psychological Review* 96:4, 703-719.
- Audi, R. 1991. "Responsible Action and Virtuous Character," *Ethics* 101:2, 304-321.
- Botvinick, M. and Cohen, J. 2014. "The computational and neural basis of cognitive control: charted territory and new frontiers," *Cognitive Science* 38:6, 1249-85.
- Braver, T. 2012. "The variable nature of cognitive control: a dual mechanisms framework," *Trends in Cognitive Sciences* 16:2, 106-13.
- Bundesen, C. 1990. "A Theory of Visual Attention," *Psychological Review* 97, 523-47.
- Burgess, G.C. and Braver, T.S. 2010. "Neural mechanisms of interference control in working memory: effects of interference expectancy and fluid intelligence," *PLoS ONE* 5, e12861.
- Clarke, R. 2014. *Omissions: Responsibility, Agency, and Metaphysics* (Oxford: Oxford University Press).
- Cohen, J.D. 2017. "Cognitive Control: Core Constructs and Current Considerations," *Wiley Handbook of Cognitive Control*, Ed. T. Egner (London: Wiley), 3-28.
- Dennett, D.C. 1984. *Elbow Room* (Cambridge, MA: MIT Press).
- Ekstrom, L. 2000. *Free Will: A Philosophical Study* (Boulder, CO: Westview Press).
- 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. 2009. "The Truth About Tracing." *Noûs* (43:3), 531-56.
- Frankfurt, Harry. 1998. *Necessity, Volition, and Love* (Cambridge: Cambridge University Press).
- Ginet, C. 2000. "The Epistemic Requirements for Moral Responsibility," *Philosophical Perspectives* 14, 267-77.
- Inzlicht, M., Shenhav, A., and Olivola, C. 2018. "The Effort Paradox: Effort Is Both Costly and Valued," *Trends in Cognitive Sciences* 22:4, 337-49.
- Jacoby, L.L., Kelley, C.M., and McElree, B.D. 1999. "The role of cognitive control: Early selection versus late correction," In S. Chaiken & Y. Trope (eds.), *Dual-process theories in social psychology* (New York: Guilford Press), 383-400.
- Kane, R. 1996. *The Significance of Free Will* (New York: Oxford University Press).
- King, M. 2009. "The Problem with Negligence," *Social Theory and Practice* 35:4, 577-595.
- King, M. 2014. "Traction without Tracing: A (Partial) Solution for Control-Based Accounts of Moral Responsibility," *European Journal of Philosophy* 22:3, 463-482.
- Kool, W., Gershman, S.J., and Cushman, F.A. 2017. "Cost-benefit arbitration between multiple reinforcement-learning systems," *Psychological Science* 28, 1321-1333.
- Kurzban, R., Duckworth, A., Kable, J.W., and Myers, J. 2013. "An opportunity cost model of subjective effort and task performance," *Behavioral and Brain Sciences* 36, 661-679.

- Levy, N. 2009. "Culpable Ignorance and Moral Responsibility: A Reply to FitzPatrick," *Ethics* 119, 729-41.
- Lewis, D. 1983. "New work for a theory of universals," *Australasian Journal of Philosophy* 61:4, 343-77.
- Mason, E. 2015. "Moral Ignorance and Blameworthiness," *Philosophical Studies* 172, 3037-3057.
- McDowell, J. 1978. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society, Supplementary Volumes* 52, 13-29.
- McDowell, J. 1979. "Virtue and Reason," *The Monist* 62:3, 331-50.
- McKenna, M. 2008. "Putting the lie on the control condition for moral responsibility," *Philosophical Studies* 139, 29-37.
- Miller, E.K. and Cohen, J.D. 2001. "An integrative theory of prefrontal cortex function," *Annual Review of Neuroscience* 24, 167-202.
- Murray, S. 2017. "Responsibility and Vigilance," *Philosophical Studies* 174:2, 507-27.
- Murray, S. and Vargas, M. Forthcoming. "Vigilance and Control," *Philosophical Studies*.
- Rosen, G. 2004. "Skepticism about Moral Responsibility," *Philosophical Perspectives* 18, 295-313.
- Savine, A.C. and Braver, T.S. 2010. "Motivated cognitive control: reward incentives modulate preparatory neural activity during task-switching," *Journal of Neuroscience* 30, 10294.
- Scanlon, T.M. 1998. *What we owe to each other* (Cambridge, MA: The Belknap Press of Harvard University Press).
- Scanlon, T.M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Belknap Press of Harvard University Press).
- Shenhav, A., Botvinick, M.M., and Cohen, J.D. 2013. "The expected value of control: An integrative theory of anterior cingulate cortex function," *Neuron* 79, 217-240.
- Sher, G. 2009. *Who Knew? Responsibility without Awareness* (Oxford: Oxford University Press).
- Shoemaker, D. 2015. *Responsibility from the Margins* (Oxford: Oxford University Press).
- Smith, A. 2005. "Responsibility for attitudes: Activity and passivity in mental life," *Ethics* 115, 236-271.
- Smith, A. 2008. "Control, Responsibility, and Moral Assessment," *Philosophical Studies* 138:3, 367-392.
- Smith, A. 2015. "Attitudes, Tracing, and Control," *Journal of Applied Philosophy* 32:3, 115-132.
- Smith, H. 1983. "Culpable Ignorance," *The Philosophical Review* 92:4, 543-571.
- Sorensen, R.A. 1985. "Self-Deception and Scattered Events," *Mind* 94:373, 64-69.
- Speer, N.K., Jacoby, L., and Braver, T. 2003. "Strategy-dependent changes in memory: effects on behavior and brain activity," *Cognitive, Affective, and Behavioral Neuroscience* 3:3, 155-67.
- Strawson, P.F. 1962. "Freedom and Resentment." In *Proceedings of the British Academy* 48, 1-25.



- Timpe, K. 2011. "Tracing and the Epistemic Condition on Moral Responsibility," *The Modern Schoolman* 88:1-2, 5-28.
- van Inwagen, P. 1989. "When is the Will Free?" *Philosophical Perspectives* 3, 399-422.
- van Inwagen, P. 1994. "When the Will is Not Free," *Philosophical Studies* 75, 95-113.
- Vargas, M. 2005. "The Trouble with Tracing," *Midwest Studies in Philosophy* 29, 269-291.
- Vargas, M. 2013. *Building Better Beings* (Oxford: Oxford University Press).
- Zimmerman, M. 2008. *Living with uncertainty* (Cambridge: Cambridge University Press).