Two Problems of Moral Luck for Brain-Computer Interfaces

DANIEL J. MILLER

ABSTRACT   *Brain-computer interfaces (BCIs) are devices primarily intended to allow agents to move prosthetic limbs, wheelchairs, and other mechanisms by forming intentions or performing certain mental actions. In this paper I illustrate how the use of BCIs leads to two unique and unrecognized problems of moral luck. In short, it seems that agents who depend upon BCIs for bodily movement or the use of other mechanisms (henceforth "BCI-agents") may end up deserving of blame and legal punishment more so than standard counterparts simply due to factors beyond their control. My aim is to explore whether we can avoid the implication that BCI-agents are subject to these unique sources of moral luck. In doing so I offer a number of possible solutions and then defend one of these solutions as the best. As it turns out, the solution I defend addresses both problems of moral luck at once and has broader implications for theorizing about moral luck as well as the epistemic condition on moral responsibility.*

1. Introduction

Human agency is mediated in diverse ways. We speak, write, sign, or type to communicate. We move with legs, wheelchairs, or prosthetics. We grasp with hands, feet, or mouths. And yet, the use of these can be either absent or lost due to congenital conditions, injury, or degenerative diseases. One example is amyotrophic lateral sclerosis (ALS), a degenerative neurological disease that affects the nerve cells required for voluntary muscle movement.[1] While conditions like ALS threaten to limit agency, brain-computer interfaces offer the prospect of restoring agency in significant ways. Brain-computer interfaces (BCIs) are devices that allow agents to move prosthetic limbs, wheelchairs, or other mechanisms by forming intentions or performing mental actions.[2]

Though BCI technology is in relatively early stages, there are signs of promise.[3] One study focuses on a 51-year-old man with advanced ALS:

> he was using an eyegaze device to interact with family and friends and to run his NIH-funded research laboratory. However, as his eye movements weakened, the device became unreliable, and he and his family despaired of his being able to continue the independent communication essential to his quality of life and to his professional productivity.[4]

With the use of a BCI at home and work, his agency has been significantly restored: "he writes manuscripts and grant proposals, sends e-mail, and reads scientific articles...the BCI controls standard software programs...He also uses the BCI for environmental control and entertainment, such as controlling the room lighting and the television."[5] When requested to explain how the BCI has made a difference in his life, he used the BCI to reply via email: "I couldn't run my lab without BCI. I do molecular neuroscience research and my grant pays three people."[6]

As BCIs become more advanced, people will use them for broader ranges of function that they otherwise wouldn't have. As this gets closer to becoming a reality, however, potential problems arise with respect to moral responsibility. In this paper I illustrate how the future use of BCIs lead to two unique and unrecognized problems of moral luck. In short, it seems that agents who depend upon

BCIs for bodily movement and other functions (henceforth "BCI-agents") may end up deserving of blame and legal punishment more so than standard counterparts simply due to factors beyond their control.[7] My aim is to explore whether we can avoid the implication that BCI-agents are subject to these unique sources of moral luck. I offer a number of possible solutions and then defend one of these solutions as the best. As it turns out, the solution I defend addresses both problems of moral luck at once. While the following discussion bears directly on questions concerning the moral and legal responsibility of BCI-agents, the considerations raised have broader implications for theorizing about moral luck as well as the epistemic condition on moral responsibility.

2. BCIs, Moral Luck, and Blameworthiness

BCIs are designed to recognize brain events that are correlated with certain intentions or mental actions and subsequently cause the corresponding movement or event in the relevant prosthetic limb or assisting device (e.g., a wheelchair, word processing application, etc.). For example, a BCI might recognize some brain activity in an agent as a proximal intention to raise their arm and consequently cause the agent's prosthetic arm to rise.

The proficient use of a BCI requires substantial instruction and training.[8] This is related to the fact that designing BCIs to recognize the right sorts of brain activity (and to ignore the wrong sorts) is a matter of precision. The difference, for example, between brain activity correlated with the mental event of vividly *imagining* raising one's arm may be subtly (though importantly) different than brain activity correlated with the mental event of actively *intending* to do so. Thus, BCIs could possibly misrecognize the former mental event as the latter. As Rainey et al. point out, this means that the control agents exercise via BCIs is often relatively limited, and that the operation of a BCI is often paired with limited foreseeability regarding its outcomes.[9] This naturally raises concerns about whether (and the degree to which) BCI-agents are responsible or blameworthy for their BCI-mediated behaviour. To illustrate this concern, consider the following cases:

> INERT IMAGINATION: Standard Saul is seated at a dull faculty meeting. As his mind wanders, he nonvoluntarily begins to vividly imagine tossing the glass of water he is holding into his colleague's face. Of course, he does not seriously consider *doing* this, and the meeting continues uninterrupted.

> ACTIVE IMAGINATION: BCI-Paul is just like Standard Saul, except that he depends upon BCI to control his prosthetic arms. When Paul nonvoluntarily begins to vividly imagine tossing his glass of water into his colleague's face, his BCI mistakenly recognizes this as a "control command" to move his prosthetic arm, and as a result his prosthetic arm tosses the water into his colleague's face.

Standard Saul is not blameworthy for anything in this case. And there are good reasons to think that neither is BCI-Paul. The first reason is articulated in the view that one cannot be (directly) blameworthy for nonvoluntary items.[10] And if one isn't blameworthy for such things, then one cannot be blameworthy for the consequences of such things.[11] Since BCI-Paul's bodily behaviour is simply the result of nonvoluntary mental events, it's implausible that he is responsible for anything in this case, including the fact that his prosthetic arm tossed water into his colleague's face.

The second reason to think that BCI-Paul is not blameworthy is articulated in the *Comparative Control Principle*:

> **CCP**: Two agents cannot differ with respect to their blameworthiness if the only differences between them are due to factors beyond their control.[12, 13]

Standard Saul isn't blameworthy for anything. And the only difference between Saul and BCI-Paul is due to factors beyond Paul's control (i.e., the device's misrecognition). If CCP is correct, then Paul is not blameworthy for tossing water in his colleague's face.

CCP is controversial, however. Some argue that agents can be subject to *moral luck*, where an agent's moral responsibility or blameworthiness is due at least in part to factors outside of his control.[14] Much of the literature on moral luck has focused on *outcome luck*. To illustrate, consider two agents who each drive recklessly down a dark street at night: the first driver swerves onto an empty sidewalk, but the second driver swerves onto a sidewalk where pedestrians happen to be present, hitting and killing them. Both drivers perform qualitatively identical actions in their driving. The only difference between them concerns the *outcomes* of their actions, and yet this difference depends only upon factors that are not under their control (i.e., whether pedestrians were present). While the only difference between the two drivers lies in factors beyond their control, a number of theorists have found it intuitive that the second driver is more blameworthy (i.e., deserving of more severe blame) than the first.[15]

If this intuition is correct, then CCP is false. However, others may be inclined to maintain CCP and deny that this sort of moral luck exists, arguing that it is fundamentally unfair to judge one agent as more blameworthy for a difference in outcomes that is completely due to factors beyond their control. Fortunately, the concern about agents like BCI-Paul can be resolved while sidestepping this controversy over moral luck, since we can appeal to the first reason offered concerning nonvoluntariness: Paul isn't blameworthy for nonvoluntarily imagining something, and if he isn't blameworthy for this, then he isn't blameworthy for its unfortunate consequence.

While the above concern is easily resolved, further reflection upon how BCIs might malfunction leads to a more formidable problem of moral luck. As Miller writes,

> [T]he concerns with moral luck raised by BCI-mediated behaviour extend beyond [Rainey et. al's] observation that *involuntary* mental events might result in overt behaviour. A distinct problem of moral luck arises from the possibility that BCIs may misrecognize a *voluntarily* formed distal intention (e.g., a plan to commit some illicit act in the *future*) as a control command to perform some overt behaviour *now*.[16]

This problem can be illustrated with the following pair of cases:

> ABANDONED PLAN: Typical Terry stands with his business rival, Gary, on top of a skyscraper (their usual meeting place for arranging high-stakes business deals). Terry judges that it might become advantageous to get rid of Gary. Inspired by the view, Terry voluntarily forms a plan to shove Gary off of the skyscraper on some future occasion (if, for example, their business dealings go south). Moments later, however, Terry has a crisis of conscience and consequently abandons his plan.

> PREMATURE EXECUTION: BCI-Barry is just like Typical Terry except that he depends upon a BCI to control his prosthetic arms. Through instruction, Barry has become aware that BCIs can misrecognize distal intentions to Φ later as proximal intentions to Φ now. When Barry voluntarily forms a plan to shove Gary off of the skyscraper on some future occasion, his BCI misrecognizes this as a proximal intention, resulting in Barry's prosthetic arms shoving Gary off of the skyscraper to his death.[17]

This case-pair gives rise to a more serious problem when paired with two additional considerations. The first consideration is that there seems to be no relevant difference between Typical Terry and BCI-Barry with respect to their control. To see this we can stipulate that, if it weren't for

his BCI's misrecognition, Barry also would have succumbed to pangs of conscience shortly after forming his plan and abandoned it. But then the only difference between Barry's plan and Terry's plan (such that Barry's resulted in someone's death and Terry's did not) is due to factors beyond Barry's control.[18]

The second consideration is that, even so, there are reasons to think that BCI-Barry is more blameworthy than Typical Terry.[19] After all, bringing about Gary's death is significantly worse than momentarily intending to do so. Now, we can stipulate that BCI-Barry did not in fact realize the risk he was taking when he formed this distal intention.[20] Nevertheless, since Gary's death seems to be the *foreseeable* outcome of *voluntarily* forming his plan to kill Gary, we cannot excuse Barry by appealing to the idea that agents cannot be directly blameworthy for nonvoluntary events or their outcomes.[21]

If the foregoing considerations are correct, they leave us with the troubling result that Barry is more blameworthy than Terry, and yet that the difference between the two agents is just a matter of luck. So, unless and until BCI devices are developed to be failsafe, their use puts BCI-agents at a risk of increased blameworthiness due to factors beyond their control. Thus, BCI-agents may unluckily end up being more blameworthy than standard counterparts across a range of cases.

3. A Distinct Problem for BCIs?

One might aim to ameliorate the concern at hand by drawing attention to the fact that we are *all* vulnerable to outcome luck in various ways. For example, whether one is blameworthy only for negligently leaving the gas on in one's stove or also blameworthy for a gas explosion may hinge upon factors outside of one's control (e.g., whether someone happens to light a candle nearby). And whether one is blameworthy merely for running a red light or also for hitting a pedestrian after doing so can depend upon factors beyond one's control. If outcome luck exists *at all* then this is so. Furthermore, one might point out that we *all* use mechanisms that are prone to malfunction. If so, the problem of luck I'm drawing attention may not be specific to BCIs and their users, but rather a problem about mechanism malfunction more generally.

This response neglects the fact that, given the fallibility of BCI devices, BCI-agents are subject to an *additional* source of outcome luck by which they are *uniquely* affected. First, even though many of the mechanisms we use are prone to malfunction, that doesn't mean that, *whenever* they malfunction and cause some harmful outcome, we're blameworthy for the outcome. After all, if someone does something entirely permissible and a bad outcome unluckily occurs, it's hard to see how we can justifiably blame them for the outcome unless they're blameworthy for performing the action that led to the outcome (e.g., if the action itself was negligent or reckless).[22] In contrast, I'm supposing that BCI-agents like Barry (and other BCI-agents discussed below) perform mental actions (they have voluntarily formed distal intentions to do something wrong) *for which they're blameworthy*. Now, we may all form blameworthy distal intentions. But those of us who don't rely upon BCIs to carry them out don't run the same risk of having distal intentions translated into overt consequences *then and there*.

Second, consider any fallible mechanism that agents who are not disabled rely upon—for example, a steering wheel that might malfunction, veering the car left or right. Now replace that agent with a BCI-agent who uses a BCI to control his prosthetic arms. That BCI-agent now has *two* sources of possible outcome luck: malfunction of the steering wheel mechanism *plus* malfunction of his BCI. Even if the steering wheel doesn't malfunction, the BCI might. And this can apply to any number of cases in which BCI-agents use a fallible mechanism (i.e., a BCI) to use another fallible mechanism (e.g., a steering wheel). In such cases the potential for unlucky outcomes is *double-layered*.

Lastly, agents who depend upon BCIs for functionality may almost always have the potential for luck even when they're *not* using another mechanism, whereas this isn't the case for other agents.

Given that it would seem particularly unfair if BCI-agents were subject to an additional layer of luck, it's worth exploring whether we can avoid it.[23]

Before considering possible solutions to this problem, I raise a second, perhaps more troubling problem of luck that arises for BCI-agents with respect to punishment. In the following section I develop this problem, and in Section 5 I suggest a solution that resolves both problems of luck.

4. BCIs, Moral Luck, and Deserved Punishment

A second problem of moral luck arises from recognizing a parallel between BCI-Barry's case and cases of involuntary manslaughter. Consider the following:

> MALFUNCTION: Jerry works hard but struggles to make ends meet. He drives his car to and from work, which is five miles from his home. Unfortunately, the steering mechanism on his car occasionally malfunctions and momentarily veers the car slightly to the right. Jerry wishes that he could have it repaired but cannot afford to. To make matters worse, there is no available public transportation Jerry can use to get to work instead, and there are no other jobs available to him that are nearer to his home. One evening while driving home from work, his car's steering mechanism malfunctions, the car veers slightly to the right, and the car hits and kills a pedestrian walking on the side of the road.

Jerry did not intentionally kill the pedestrian. Moreover, the car's veering to the right isn't even an action on Jerry's part. By all accounts, the pedestrian's death was an accident. But Jerry is plausibly deserving of legal punishment for involuntary manslaughter. The steering mechanism's malfunction and the death that resulted from it were foreseeable outcomes of Jerry's voluntarily driving the car.

BCI-Barry's case is analogous. Given that the connection between Barry's distal intention and the movement of his prosthetic arm is causally deviant, the movement of Barry's prosthetic arms hardly constitutes an action. Nevertheless, Barry's BCI device is known to occasionally malfunction by misrecognizing brain activity (or mental events correlated with such activity) and resultingly causing unintended movement of Barry's prosthetic limbs. The BCI is therefore analogous to Jerry's car, which is known to occasionally malfunction by veering to the right. Furthermore, Gary's death was a foreseeable outcome of Barry's voluntarily forming his distal intention, as the pedestrian's death was a foreseeable outcome of Jerry's voluntarily driving his car.

One might think there's a difference. We can suppose that Jerry was aware of the risk he was taking by deciding to continue to use his car. In contrast, I have stipulated that BCI-Barry was not aware of the relevant risk he was taking at the time that he voluntarily formed his distal intention to kill Gary. Regardless, it doesn't seem that this succeeds in driving a wedge between the two cases. For, we can simply suppose that, when Jerry drove his car home that evening, he wasn't *attending* to the risk that doing so posed. But, of course, we wouldn't think that this excuses him. Rather, we would readily point out that Jerry *should* have been aware of the risk. Given this, we need a reason to think that we can't say the same for BCI-Barry, given Barry's general awareness of the risks that BCIs pose.

The strength of these similarities lends considerable support for the claim that, if Jerry deserves to be punished for the pedestrian's death, then BCI-Barry deserves to be punished for Gary's death. And since it seems that Jerry deserves legal punishment for the pedestrian's death, it would seem to follow that Barry deserves legal punishment for Gary's death.[24]

But now return to the comparison between BCI-Barry and Typical Terry, and recall that BCI-Barry differs from Typical Terry only with respect to factors outside of his control. With this in mind, consider an analogue of the Comparative Control Principle which focuses, not on blameworthiness, but on deserved punishment:

> **CCP^{DP}**: Two agents cannot differ with respect to their desert of punishment if the only differences between them are due to factors beyond their control.

It seems unfair to judge that BCI-Barry is deserving of rather severe punishment when Terry is not. This especially so since the primary difference between the two agents is just that Barry depends upon his BCI for bodily movement, and that Barry's dependence upon his BCI is, we may suppose, beyond his control. As with the Comparative Control Principle, so it may be with CCP^{DP}: perhaps luck with respect to deserved punishment is unavoidable. Nevertheless, it's worth exploring whether there is a principled way to avoid an additional and unique species of this type of luck that would stem from reliance on BCIs.

5. Defending a Solution to the Two Problems

There are at least three general ways to avoid the result that BCI-Barry is deserving of punishment for Gary's death and therefore subject to a unique source of moral luck that standard counterparts are not. In what follows I consider each, and then defend a solution that dissolves both the problem of luck concerning punishment and the problem of luck concerning blameworthiness.[25] Consider the following possible solutions:

> *Solution 1* denies that the difference between BCI-Barry and Typical Terry is just a matter of luck, and accordingly attempts to establish a relevant difference in control between the two agents. If this attempt succeeds, then Barry's deserving punishment is not truly an instance of outcome luck, and the problem does not arise. The most obvious candidate for a difference in control is this: since Barry has a general awareness of the possibility that his BCI may malfunction, he has the ability to avoid the sort of outcome in question simply by refraining from forming *any* intentions (distal or otherwise) that might be mistakenly translated into action right away. If he has this ability, then Gary's resulting death is *not* due to factors that are entirely beyond Barry's control.

> *Solution 2* grants that the difference between BCI-Barry and Typical Terry is just a matter of luck and that they are equally deserving (*or* undeserving) of some legal punishment. On this solution, *both* Barry and Terry are deserving of punishment. This is implausible, however. Even if Terry is deserving of *blame* for voluntarily forming the intention to kill Gary, there's no plausible candidate for which Terry deserves *punishment*. Although it is reasonable to maintain that agents are sometimes deserving of punishment for *attempted* crimes, such attempts are typically identified with *actions* that agents take toward the realization of their plans. Thus, an assassin who shoots at but misses his target is plausibly deserving of punishment for attempted murder in virtue of shooting at his target. In contrast, Typical Terry never takes any active steps towards realizing this plan, since he abandons his plan shortly after forming it. Solution 2 can therefore be dismissed.

> *Solution* 3, like the previous solution, grants that the difference between BCI-Barry and Typical Terry is just a matter of luck and that they are equally deserving (or undeserving) of legal punishment. Solution 3 differs from Solution 2, however, in that it maintains that *neither* Barry nor Terry is deserving of any legal punishment. There are two ways to defend this contention:

> > *Solution 3.1* maintains that BCI-Barry is not even blameworthy for Gary's death. Since deserved punishment for Φ plausibly requires deserved blame for Φ, it would follow that Barry is not

6

deserving of punishment for Gary's death.[26] However, given the details of the case, it's not immediately obvious how Barry might be excused. Barry's awareness of the possibility of BCI malfunction (even if not occurrent) seems to make Gary's death a foreseeable upshot of his voluntary mental action of forming his plan. However, this solution merits a closer look.

Solution 3.2 maintains that BCI-Barry *is* blameworthy for Gary's death but not deserving of punishment for it. However, this solution is undermotivated. It is, of course, true that being blameworthy for Φ is not *sufficient* for being deserving of legal punishment for Φ. We would not, for example, think that blameworthiness for purely mental actions (e.g., deciding to tell an insensitive joke), relatively trivial actions (e.g., telling an insensitive joke), or relatively trivial consequences thereof (someone's hurt feelings) suffices for deserving legal punishment for those things. But this solution grants that Barry is in fact blameworthy for something quite serious, viz., bringing about someone's death. It's implausible that something so severe (supposing one is blameworthy for it) could fail meet the appropriate standard for deserving legal punishment.

Solution 1 and Solution 3.1 remain. I what follows I defend Solution 3.1 over Solution 1. We can begin by taking a closer look at how Solution 1 aims to establish a relevant difference in control between Typical Terry and BCI-Barry. The defender of Solution 1 can maintain that, through BCI instruction, Barry has acquired a general ability to foresee malfunction-caused outcomes of distal intentions (in this case, plans to perform some action at a later time). Either before or during the formation of distal intentions, Barry is able to attend to his awareness of possible BCI-malfunction and thus foresee the sort of outcomes that might result from forming those intentions. A proponent of Solution 1 might plausibly argue that these facts are sufficient for Gary's death being a foreseeable outcome of Barry's voluntarily formed plan. And, since Gary's death was foreseeable for Barry, the proponent of Solution 1 might continue, Barry could have avoided that consequence by refraining from forming his plan on the basis of such foresight. If this is so, then Barry is blameworthy for Gary's death.

Whether Solution 1 succeeds, however, depends upon the distinction between an outcome's being foreseeable and an outcome's being *reasonably* foreseeable.[27] Gary's death is plausibly foreseeable for BCI-Barry. But whether it is reasonable to expect Barry to foresee this is quite another question. And, if it's not reasonable to have expected Barry to foresee this, then it's implausible that he is blameworthy for it.[28] More generally, it would be unreasonable to expect BCI-agents like Barry to foresee such consequences because fulfilling this expectation would seem to require that BCI-agents constantly monitor each of their intentions or plans—no matter how momentary and fleeting—with an eye to the mere possibility that such intentions or plans might be misrecognized by their BCIs. This seems to require either an extraordinary capacity for attention and consideration that ordinary humans simply don't have, or else that BCI-agents deliberate, form intentions, and live in perpetual slow motion in order to ensure that they always cover all of their bases, as it were. Even if BCI-agents have the general ability to foresee such outcomes, it doesn't follow from this that it's reasonable to expect the agent to exercise such abilities in the way that would be required to avoid them.

This is bolstered by the experience of people with conditions such as ALS training to use BCIs. One study provides qualitative feedback from BCI users in focus groups who described their experience using a BCI spelling device:

> Four of the eight social groups also expressed concern over the psychological issue of cognitive fatigue involved in using a BCI. These participants doubted their psychological capacity to manage fatigue caused by using the P300 speller, and believed that their endurance would limit the usability

of this technology. '. . .I don't know, I just wonder what my stamina would be. . . I personally wonder how long I'd be able to do that. . . I have doubts that I'd be able to do it for, like, over an hour or so at a time.'[29]

Five of the eight social units remarked that it was psychologically challenging to maintain focus on the spelling task during inevitable, yet unexpected, events. 'And you're sitting there trying to focus on the screen, and all of the sudden the ice cube maker, you know, drops ice cubes. Your train of thought gets staggered. Same thing if a dog was barking outside, or somebody. . .the phone rings, just any normal noise that you would hear in a house can be distracting while you're trying to focus on that.'[30]

Importantly, BCI use is cognitively taxing, and while further technological development may reduce this cognitive load, it may not eliminate it entirely.

As Rainey et. al rightly point out, there is reason to treat differently cases in which able-bodied agents *temporarily* employ the use of BCIs for *recreation* (e.g., using BCIs to control drones), on the one hand, and cases where persons with disabilities are *dependent* on the *constant* use of BCIs in an effort to approximate an average standard of bodily agency.[31] In the former type of case, it may be reasonable to expect able-bodied agents to exercise an extraordinary level of caution for the temporary period that they are using a BCI for recreation. But the expectation that persons with disabilities perpetually sustain this level of caution and attention throughout their daily lives sets a requirement that would be unfairly psychologically burdensome. Thus, while Solution 1 identifies a notable difference between BCI-agents and their standard counterparts, there is reason to think that this difference is (at least often) not a responsibility-relevant difference. This same point also provides reason to doubt that BCI-agents like Barry are blameworthy for the type of consequences of which Gary's death is an instance.

At this point a defender of Solution 1 may grant that it's unreasonable to expect BCI-agents to monitor *all* of their intentions. And yet, it may nevertheless be reasonable to expect BCI-agents to monitor a suitably restricted subset of their intentions, i.e., violent or murderous ones. If so, then consequences of this sort of intention would be reasonably foreseeable for BCI-agents and thus avoidable. On this basis, the defender of Solution 1 may conclude that there is in fact a relevant difference in control between BCI-Barry and Typical Terry: Gary's death on that occasion was a reasonably foreseeable result of Barry's plan, but not Terry's.

This defense on behalf of Solution 1 succeeds only if the scope of intentions BCI-agents must monitor can plausibly be reduced to the point at which it would be reasonable to expect them to do so (and thus foresee the relevant outcomes). But it's far from clear that this is possible. Consider the following two cases that involve BCI-agents:

TAX FRAUD: Carrie is filing her annual tax return online. She has refrained from entering information about some under-the-table earnings for that year. Though she has intermittently experienced feelings of guilt about this omission, she has so far been able to keep these guilty feelings at bay. At the final step in the filing process, Carrie needs only to select 'enter' on her keyboard to submit her tax returns. If she does so, she will have committed tax fraud. At that moment, Carrie is distracted by a text message notification from her phone. She reflexively checks her phone for a message she has been waiting for and begins typing a reply. As she does so, it occurs to her that she hasn't yet submitted her tax return, and voluntarily forms a plan to submit them after she finishes replying to the text message. However, Carrie's BCI misrecognizes this as a proximal intention to submit her tax return now, resulting in her prosthetic hand reaching back to her computer keyboard and clicking 'enter.' Her tax return is submitted, and it is fraudulent. If it weren't for her BCI's malfunction, Carrie's guilt-feelings would have returned, and she would have abandoned her plan to submit the fraudulent tax return.

SPEEDER: Sherry is out for a drive in her new car. Cruising on the highway just at the speed limit, Sherry wants to see how quickly her new car can accelerate to a higher speed. While considering this, Sherry feels a bit light-headed and realizes that she hasn't eaten all day. She figures that accelerating to a higher speed on an empty stomach might be nauseating, and voluntarily forms a plan to test her car's acceleration after stopping for lunch. However, Sherry's BCI misrecognizes this plan as a proximal intention to accelerate now, resulting in her prosthetic leg pressing down hard on the gas pedal, quickly accelerating well past the speed limit. Were it not for her BCI's malfunction, Carrie would have rethought her reckless plan over lunch and refrained from speeding.

These cases help illustrate that the scope of intentions or plans to be monitored by BCI-agents is significantly broader than murderous or violent intentions. And, of course, cases of this kind could be multiplied indefinitely. Importantly, the broader the set of such cases is, the less plausible it is that it's reasonable to expect such agents to monitor these intentions and thereby foresee their possible consequences.[32] The result is that, even if such consequences are foreseeable, they are plausibly not *reasonably* foreseeable.[33] If this is so, then Solution 1 fails.[34, 35]

Solution 3.1 remains. We needn't, however, accept Solution 3.1 simply as the result of the forgoing process of elimination. Instead, notice that Solution 3.1 is essentially a denial of what Solution 1 affirms. Solution 1 maintains that BCI-Barry is blameworthy (and deserving of punishment) for Gary's death, while Solution 3.1 maintains that BCI-Barry is not blameworthy (and thus not deserving of punishment) for Gary's death. My argument *against* Solution 1 is, in effect, an argument *for* the truth of Solution 3.1.

More explicitly, my argument for Solution 3.1 is a conjunction of two claims: (i) an agent is blameworthy for an outcome of some action only if the outcome was reasonably foreseeable for the agent at the time of the action, and (ii) the outcome in question was not reasonably foreseeable for BCI-Barry at the time of his earlier mental action. These claims jointly entail that BCI-Barry is not blameworthy for the outcome in question, and thus not deserving of punishment (the same goes for other relevantly similar BCI-agents like Carrie and Sherry).

In my argument against Solution 1 I focused primarily on defending (ii), while taking (i) as a shared assumption that even the defender of Solution 1 attempts to employ as part of their argument for it (the appeal the defender of Solution 1 makes to (i) misleadingly trades on the linguistic ambiguity between "foreseeable" and "reasonably foreseeable"). Furthermore, the fact that (i) is widely accepted among theorists tells in its favor.[36]

A positive argument for (i) can be made on the grounds that it's required to preserve the link between blameworthiness and fairness. This argument begins from the premise that it's fair to blame an agent for something only if it's reasonable to expect the agent to have avoided it.[37] We can apply this premise to outcomes: it's fair to blame an agent for the outcome of their action only if it's reasonable to expect the agent to avoid that outcome. But it's reasonable to expect an agent to avoid an outcome only if it's reasonable to expect them to be aware that the outcome might result from their action. And this is so only if the outcome is reasonably foreseeable for the agent at the time of the action. It follows from these considerations that it's fair to blame someone for an outcome only if the outcome was reasonably foreseeable for the agent at the time of their action. This reasoning establishes claim (i): an agent is blameworthy for an outcome of some action only if the outcome was reasonably foreseeable for the agent at the time of the action. This, conjoined with claim (ii) (that the outcome in question was not reasonably foreseeable for BCI-Barry at the time of his earlier mental action), entails that BCI-Barry is not blameworthy for the outcome (i.e., Gary's death). The same goes, *mutatis mutandis*, for other BCI-agents. If the foregoing reasoning is sound, then Solution 3.1 is correct.

Notice that Solution 3.1 at once dissolves the problem of blameworthiness and the problem of deserved punishment for BCI-agents. Blameworthiness for some item is plausibly a necessary condition for deserving punishment for it. As Douglas Husak writes, "Penal liability subjects persons to state punishment, and any respectable justification of punishment requires persons to be morally blameworthy".[38] Since blameworthiness is plausibly a precondition of deserved punishment, a successful argument that BCI-Barry is not blameworthy for Gary's death also implies that neither is he deserving of punishment for it. This conclusion generalizes for other BCI-agents whose bodily movements (and their upshots) are not reasonably foreseeable results of voluntary mental actions.[39]

6. Apology, Agent-Regret, and Relationship-Alteration

I've offered an argument that the BCI-agents discussed here aren't deserving of blame or punishment for the outcomes of their mental actions. In doing so, I have denied that they are subject to an additional source of outcome luck by which they are uniquely affected. However, one may feel that, if we simply stop here, there's something left unresolved. Although BCI-Barry isn't blameworthy for Gary's death, something would seem lacking if he simply shrugged it off. And even if Barry sincerely thinks of Gary's death as regrettable, something still seems missing. So, although it wouldn't be appropriate to blame Barry for Gary's death, or for Barry to experience feelings of guilt for it, there may be some related attitude or activity that would be appropriate: something in between mere regret, on one hand, and blame, on the other. Here I briefly explore three such possibilities.

One possibility is what theorists have called *agent-regret*. As mentioned, it would seem insufficient for BCI-Barry to merely feel regret with respect to Gary's death. After all, *anyone* might reasonably feel regret about his death, even if they had nothing at all to do with it. Bernard Williams' distinction between "regret in general" and agent-regret is helpful here.[40] Regret in general can be appropriately felt by anyone who is aware of an unfortunate state of affairs or event: one may sincerely wish that things had been otherwise. Agent-regret, in contrast, is only appropriately felt by an agent who is related to the outcome via his or her own agency. As Williams explains, "The sentiment of agent-regret is by no means restricted to *voluntary* agency. It can extend far beyond what one intentionally did to almost anything for which one was causally responsible in virtue of something one intentionally did."[41] The notion of agent-regret seems to fit well with the BCI-agents discussed here, whose voluntary plans accidentally result in outcomes via involuntary events. These agents are partly *causally* responsible for these outcomes, but not *morally* responsible. Agent-regret might fill the attitude gap where blame or guilt wouldn't be appropriate, since the appropriateness of agent-regret falls somewhere between being unrelated to some outcome and being blameworthy for it. Williams illustrates:

> The lorry driver who, through no fault of his, runs over a child, will feel differently from any spectator... Doubtless, and rightly, people will try, in comforting him, to move the driver from this state of feeling, move him indeed from where he is to something more like the place of a spectator; but it is important that this is seen as something that should need to be done, and indeed some doubt would be felt about a driver who too blandly or too readily moved to that position...there is something special about his relation to this happening, something which cannot merely be eliminated by the consideration that it was not his fault.[42]

Marcia Baron describes agent-regret as involving a distinctive kind of pain. Of Williams' lorry driver, she writes, "The pain that the driver feels about what happened is not just pain about *that*. He was centrally involved in what happened; indeed, he was the agent. That *he was the agent*, even though

what he did was unintentional and even unforeseeable, is ineliminably a part of what he feels."[43] Baron's observation can be adapted for BCI-agents: although the movement of his arms doesn't constitute an *action* on Barry's part, his mental action is an exercise of agency through which the subsequent outcome came about. Thus, he might appropriately feel a distinctive sort of pain upon reflecting that his own agency was (albeit in a causally deviant way) a conduit of Gary's death.

A second possibility concerns the notion of *relationship-alteration* as developed by T. M. Scanlon. While both agent-regret focuses on the agent as the subject, this notion shifts our focus to responses from others. According to Scanlon's account, an agent's blameworthiness for some fault should be understood in terms of how that fault impairs his relationships with others. Consequently, blaming responses toward the agent involve relationship-alterations that are made appropriate by the fault for which the agent is blameworthy (e.g., withdrawal of trust, withdrawal of good will, etc.).[44] Relevant here is Scanlon's view that agents who unluckily bring about bad outcomes are no more blameworthy than their more fortunate counterparts, but that the bad outcomes can still increase the *significance* of their blameworthiness for those affected by the fault.[45] Thus, perhaps different kinds of relationship-alterations would be appropriate toward the BCI-agents as compared with their standard counterparts. Gary's family might reasonably feel or behave differently toward someone who (like Typical Terry) momentarily planned to kill Gary but refrained from doing so than they do toward someone who (like BCI-Barry) actually brought about Gary's death (even if unintentionally).[46] They might reasonably find it difficult to reconcile with Barry, given that he was causally responsible for their family member's death. They might even refuse to be in his presence, given his role in the tragedy.

Lastly, consider the act of *apology*, which concerns how the agent might address others. At first glance, it may seem that it's appropriate to apologize for something only if one is blameworthy for it. If Barry isn't blameworthy for Gary's death, then how could an apology for it be appropriate? Some theorists, however, defend the view that apologies can be appropriate even if no one is morally responsible for anything. On Derk Pereboom's view, an apology can communicate an acknowledgement of wrongdoing, a wish that one hadn't committed it, and a commitment to avoid such wrongdoing in the future.[47] Bruce Waller draws attention to the fact that an apology can acknowledge that one's wrongdoing stemmed from a flaw in one's own character.[48] These features of apology are consistent with the view that one is blameless for who one is and for what one does.

Might apologies of this sort be appropriate for the BCI-agents discussed here? Each of them performs a mental action that reflects a flaw in their regard for others (e.g., at least a momentary willingness to carry out a morally wrong plan). Indeed, it might be entirely appropriate for them to apologize for these mental actions (one can imagine a remorseful BCI-Barry admitting to Gary's family that he did murder Gary *in his heart*, as it were). But notice that an apology for *this* would also be appropriate for his counterpart, Typical Terry. The question we're interested in here is whether it might also be appropriate for Barry to apologize for Gary's death. And there are some reasons to think not. The movement of Barry's arms isn't an action; *a fortiori* it's not a morally wrong action. Furthermore, it's dubious that an accidental outcome (Gary's death) can itself be expressive of a flaw in one's character. While this doesn't settle the matter, it's less than clear that it would be appropriate for BCI-agents to apologize for these unlucky outcomes. And perhaps one of the sources of ambivalence here is that there are different kinds of apology. A certain kind of apology may come in the form of an *expression* of agent-regret: an expression that one deeply regrets that one's own agency was part of the causal nexus that resulted in the tragic outcome. However, if these agents are (as I've argued) not at fault of the outcomes in question, the kind of apology that involves an acceptance of fault or an expression of guilt for the outcome would be less clearly appropriate.

While these three possibilities are not exhaustive, reflecting upon them can help point the way forward in accounting for some of the conflicting intuitions we might have concerning whether to accept certain types of outcome luck.

7. Conclusion

I have presented two unique problems of moral luck for agents who depend upon BCIs for the movement of their prosthetic limbs or the use of other mechanisms. These problems arise due to the peculiar possibility that BCIs allow for the mistaken recognition of distal intentions as proximal intentions. The solution I offer—that BCI-agents simply are not blameworthy in such cases—has the advantage of solving both problems at once.

The considerations raised here extend beyond the focus of this paper, having broader implications for theorizing about moral luck. One lesson is that accepting outcome luck doesn't require that we accept it in *all* its forms. Whether we should accept some species of outcome luck may partly depend upon whether it would affect certain types of agents uniquely, subjecting them to an additional layer of luck from which other agents are free. This discussion also elucidates an important requirement on reasonable foreseeability, a notion that plays a considerable role in the literature on the epistemic condition on moral responsibility.[49] We may mistakenly assume that an outcome is reasonably foreseeable for an agent if that outcome can be foreseen upon the agent exercising some general capacity that the agent possesses. The considerations raised here, however, suggest that an outcome is reasonably foreseeable only if it is also reasonable to expect the agent to exercise this capacity on a particular occasion.

Daniel J. Miller, Department of Philosophy, West Virginia University, 345 Price St., Morgantown, WV, 26505, USA. daniel.miller2@mail.wvu.edu

Acknowledgments

NOTES

[1] National Institutes of Health, 'Amyotrophic Lateral Sclerosis (ALS) Fact Sheet' (https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Amyotrophic-Lateral-Sclerosis-ALS-Fact-Sheet).

[2] That BCIs can help restore a sense of agency is supported by qualitative feedback from focus groups of prospective BCI users. See: Stefanie Blain-Moraes, Riley Shaff, Kirsten L Gruis, Jane E. Huggins, & Patricia A. Wren, 'Barriers to and Mediators of Brain-Computer Interface User Acceptance: Focus Group Findings', *Ergonomics* 55, 5 (2012) at p. 523.

[3] BCI research sometimes fails to prioritize perspectives of those who stand to gain from their use. Accordingly, some have insisted that BCI design must be "user-centred." See, for example: Laura Specker Sullivan, Eran Klein, Tim Brown, Matthew Sample, Michelle Pham, Paul Tubig, Raney Folland, Anjali Truitt, & Sara Goering, 'Keeping Disability in Mind: A Case Study in Implantable Brain-Computer Interface Research,' *Science and Engineering Ethics* 24 (2018), pp. 479-504.

[4] Eric W. Sellers, Theresa M. Vaughan, & Jonathan R. Wolpaw, 'A Brain-Computer Interface for Long-Term Independent Home Use', *Amyotrophic Lateral Sclerosis* 11 (2010): 449-455.

[5] Sellers et al. op. cit., at p. 452.

[6] Sellers et al. op. cit., at p. 453.

[7] The BCIs I imagine in this paper are somewhat more advanced than current BCI technology. This is intentional, since my aim is to address future potentialities for the widespread application of BCIs as they relate to moral responsibility.

[8] Stephen Rainey, Hannah Maslen, & Julian Savulescu, 'When Thinking is Doing: Responsibility for BCI-Mediated Action Requires Special Attention in Terms of Controllability and Foreseeability of Outcomes', *AJOB Neuroscience* 11, 1 (2020): 46-58. Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, et. al., 'Brain-Computer Interfaces for Communication and Control', *Clinical Neurophysiology* 113, 6, (2002): 767-791.

[9] Rainey et al. op. cit.

[10] See, for example: Neil Levy, 'The Good, the Bad, and the Blameworthy', *Journal of Ethics and Social Philosophy*, 1, 2 (2005): 1-16. One might nevertheless be indirectly blameworthy for some nonvoluntary item if it results from some earlier exercise of voluntary control. See, for example: Gideon Rosen, 'Skepticism about Moral Responsibility', *Philosophical Perspectives*, 18, 1 (2004): 295–313. We can assume that this is not the case for BCI-Paul.

[11] This is somewhat simplified. It's possible that one can be responsible for a consequence Z of something Y even if one is not responsible for Y, since Z might *also* be a consequence of something else, X, for which one *is* responsible. I assume that this is not the case here.

[12] This principle (or very similar ones) is formulated in terms of moral responsibility and moral assessment, respectively by the following authors: John Greco, 'A Second Paradox Concerning Responsibility and Luck', *Metaphilosophy* 26 (1995): 81-96; Dana K. Nelkin, 'Moral Luck', *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2019/entries/moral-luck/>.

[13] I take this name for the principle from Robert Hartman, *In Defense of Moral Luck: Why Luck Often Affects Praiseworthiness and Blameworthiness* (New York: Routledge, 2017), p. 7.

[14] See, for example: T. Nagel, 'Moral Luck', in T. Nagel (ed.) *Mortal Questions* (Cambridge: Cambridge University Press, 1979); Hartman op. cit.

[15] Nagel op. cit. For theorists who explicitly defend outcome luck, see Michael Moore, *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics* (Oxford: Oxford University Press, 2009); Hartman op. cit.

[16] Daniel J. Miller, 'BCI-Mediated Behaviour, Moral Luck, and Punishment', *AJOB Neuroscience* 11, 1 (2020): 72-74, at p. 72.

[17] Miller op. cit., at p. 72.

[18] Miller op. cit.

[19] One might deny that BCI-Barry is *more* blameworthy, and instead maintain that Barry is simply blameworthy for *more* items (i.e., the voluntarily formed plan *and* Gary's death). This view rests on the distinction between the *scope* of an agent's responsibility and the *degree* to which the agent is responsible. See, for example: Michael Zimmerman, 'Taking luck seriously', *The Journal of Philosophy*, 99, 11 (2002): 553-576, at p. 560. This is, in a way, to deny the existence of outcome luck altogether. I proceed on the assumption that the scope of an agent's responsibility can affect their degree of responsibility. My focus is whether, assuming there is outcome luck, we can avoid a particular *species* of it that might arise from dependence upon BCIs.

[20] While BCI-Barry has a general awareness that BCIs can misrecognize distal intentions as control commands in the relevant way, we can stipulate that Barry's awareness of this fact was not *occurrent* at the time.

[21] One might object that, if Gary's death (on that occasion) was a *foreseeable* outcome of BCI-Barry's voluntarily formed plan (but not Terry's), then there is a relevant difference between these two agents with respect to their control, and thus this case-pair isn't in tension with CCP. Since it was foreseeable for Barry that his distal intentions might have this sort of outcome, he was able to refrain from voluntarily forming intentions that, if mistakenly translated into overt behaviour, would have such outcomes. If this is correct, then the first consideration I raise is mistaken. I address this objection in detail in Section 5. However, as a first pass, we can plausibly suppose that, even if the outcome was foreseeable for Barry, whether he became *occurrently aware* of the relevant fact was also due to luck. After all, it's generally not under our control whether something (even if foreseeable) occurs to us or not.

[22] Those who choose to use a BCI for functionality may voluntarily open themselves up to the potential for BCI malfunction. Consequently, one might wonder whether BCI-agents are responsible for outcomes that result due to BCI malfunctions. However, we all permissibly open ourselves up to luck by using mechanisms that have a chance of malfunction. I might, for example, permissibly use a toaster even though it could malfunction and start a fire. In general, it's permissible to depend upon fallible mechanisms for certain functions (e.g., driving to travel to work, using computers to file taxes, etc.). We're blameworthy for unlucky *outcomes* of using mechanisms only if our *use* is negligent or reckless, and if those outcomes are (as I will argue later) reasonably foreseeable. We should say the same for agents who choose to depend upon BCIs for functionality.

[23] Although this problem doesn't depend upon CCP (as I acknowledge, perhaps there is *some* outcome luck), something akin to the intuition that gives rise to CCP may underly this more particular concern. It's open to theorists to grant that there may be *something* unfair about this, and yet simply to embrace the claim that BCI-agents are subject to an additional layer of moral luck. I neither intend nor expect to convince such theorists otherwise. Perhaps the best

account of moral responsibility must make peace with some measure of unfairness in the world. But it is nevertheless a worthwhile task to explore if there is a principled way to avoid allowing further layers of unfairness into one's moral ontology.

[24] In what follows I sometimes discuss this problem in terms of punishment more generally. If no punishment is deserved (as I will proceed to argue concerning BCI-agents), then, *a fortiori*, legal punishment is undeserved.

[25] Miller op. cit. lays out similar solutions and remains agnostic on which is best. Here I develop these solutions somewhat differently, explore them in more depth, and proceed to defend a position concerning which solution is best.

[26] I understand blameworthiness as a matter of the desert or fairness of blame.

[27] Notably, it has been argued that (given plausible assumptions) even reasonable foreseeability is not sufficient to satisfy the epistemic requirement on blameworthiness for outcomes. See: Daniel J. Miller, 'Reasonable Foreseeability and Blameless Ignorance', *Philosophical Studies*, 174, 6 (2017): 1561-1581.

[28] See: John Fischer and Neal Tognazzini, 'The Truth about Tracing', *Noûs*, 43, 3 (2009): 531-556. The relevant legal analogue is whether a reasonable person would have foreseen the consequence (Model Penal Code, sects. 2.02(2)(c) and 2.02 (2)(d).). However, these two standards are arguably not equivalent. In some cases it might be that, while a *reasonable* person would have foreseen Φ, it's nevertheless not reasonable to expect the person in question to expect Φ (if, for example, that person is particularly oblivious). See: Miller 2017 op. cit., at p. 1565. Regardless, it's plausible that neither the condition specified here on blameworthiness nor the legal standard is met.

[29] Blain-Moraes et al. op. cit., at p. 520.

[30] Blain-Moraes et al. op. cit., at p. 520.

[31] Rainey et al. op. cit., at p. 54-55.

[32] A proponent of 3.2 might suggest that considerations about reasonable expectation don't fully *excuse* but instead merely *mitigate* blameworthiness for the outcome. And perhaps mitigated blameworthiness isn't sufficient for legal punishment. Nevertheless, on this supposition, the agent is still blameworthy for a serious legal offense (in a way that blameworthiness for hurting someone's feelings isn't). The natural inference to draw from this, it would seem, is that the agent *is* deserving of legal punishment (*contra* 3.2), and only that the punishment itself should also be mitigated. But the claim that they deserve legal punishment at all still seems worrisome. In any case, a proponent of 3.2 would need to offer an argument either for the claim that (a) it is *not* unreasonable to expect BCI agents to foresee the outcomes in question, or else an argument that (b) an agent can be blameworthy for an outcome that isn't reasonable to expect them to foresee.

[33] If I am right about this, then the BCI-agents I have discussed differ from Jerry in *Malfunction*. The scope of behaviours for Jerry to monitor is, in contrast, much smaller and therefore quite manageable: he can reasonably be expected to be aware that driving his car poses clear risks to others.

[34] One might suggest that BCI-agents shouldn't engage in certain activities (e.g., driving) if it's not possible to avoid dangerous malfunctions. However, this suggestion doesn't neatly apply to other categories of which *Premature Execution* and *Tax Fraud* are instances. For example, it doesn't seem there is a promising method by which to group together cases in which BCI-agents might inadvertently shove people in dangerous ways (e.g., walking down busy streets, walking alongside rivers, taking cruises on ocean liners). To take another example, *Tax Fraud* is just one instance in which BCIs might result in harmful outcomes via the accidental click of a button. But there are all kinds of scenarios in which this might occur (e.g., making foolish purchases online, pressing buttons in factories, etc.), and the category of "being near buttons" is hardly a neatly definable category. In sum, I suspect there's no non-*ad hoc* way to capture in well-defined categories all scenarios in which BCIs might malfunction, and then to incorporate these into general obligations for BCI-agents (e.g., "Refrain from being in X-type situations...").

[35] There is room to adopt a hybrid view. One might maintain that Solution 1 is plausible with respect to a very restricted set of distal intentions (e.g., murderous or violent ones) while adopting Solution 3.1 with respect to the wider set of more mundane cases such as *Tax Fraud* and *Speeder*. I am grateful to Gabriel De Marco for suggesting this.

[36] Gideon Rosen, 'Kleinbart The Oblivious and Other Tales of Ignorance and Responsibility', *Journal of Philosophy* 105, 10 (2008) at p. 604; Fischer and Tognazzini op cit., at pp. 537-538; Jan Willem Wieland and Philip Robichaud, 'Blame Transfer', in P. Robichaud and J.W. Wieland (eds.) *Responsibility: The Epistemic Condition* (Oxford: Oxford University Press, 2017), at pp. 286-287.

[37] William J. FitzPatrick, 'Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge' *Ethics* 118, 4, (2008): 589-613; Neil Levy, 'Culpable Ignorance and Moral Responsibility: A Reply to FitzPatrick', *Ethics* 119, 4 (2009): 729-741.

[38] Douglas Husak, 'Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting', *Criminal Law and Philosophy*, 5 (2011): 199-218, at p. 200.

[39] It is a distinct question whether, all things considered, these BCI-agents *ought* to be punished for the relevant outcomes. That would depend upon additional factors (e.g., the social functions of punishment) that are beyond the scope of this paper.

[40] B.A.O. Williams, 'Moral Luck', *Proceedings of the Aristotelian Society*, Supplementary Volumes, 50 (1976): 115-135.

[41] Williams op. cit. at pp. 123-124, emphasis original.

[42] Williams op. cit. at pp. 123-124.

[43] Marcia Baron, 'Remorse and Agent-Regret', *Midwest Studies in Philosophy*, 8 (1988) at p. 261, emphasis original.

[44] T.M. Scanlon, *Moral Dimensions: Permissibility, Meaning, and Blame* (Cambridge, MA: Harvard University Press, 2008).

[45] T.M. Scanlon, 'Forms and Conditions of Responsibility' in Randolph Clarke, Michael McKenna, and Angela M. Smith (eds.) *The Nature of Moral Responsibility: New Essays* (New York: Oxford University Press, 2015), at p. 105.

[46] For a thoughtful discussion of how Scanlon's account might approach the question of outcome luck, see Matthew Talbert, 'The Attributionist Approach to Moral Luck', *Midwest Studies in Philosophy* 43, 1 (2019), at pp. 28-31.

[47] Derk Pereboom, *Living Without Free Will* (New York: Cambridge University Press, 2001) at p. 201.

[48] Bruce N. Waller, 'Sincere Apology Without Moral Responsibility', *Social Theory and Practice*, 33, 3 (2007), at p. 445.

[49] Gideon Rosen, 'Skepticism about Moral Responsibility', *Philosophical Perspectives*, 18, 1 (2004): 295–313; Manuel Vargas, 'The Trouble with Tracing', *Midwest Studies in Philosophy*, 29, 1 (2005): 269–291; Fischer and Tognazzini op. cit.