

The Sheriff in Our Minds: On the Morality of the Mental

Many people believe that our thoughts can be morally wrong. Many regard rape and murder fantasies as wrong. In a recent essay, George Sher disagrees with this and argues that “the realm of the purely mental is best regarded as a morality-free zone,” wherein “no thoughts or attitudes are either forbidden or required” (484). Sher argues that “each person’s subjectivity is a limitless, lawless wild west in which absolutely everything is permitted” (484). Sher calls this view the Wild West of the Mind.

I argue against Sher’s position. In section I, I summarize Sher’s view. In section II, I outline and criticize Sher’s argument for the Wild West of the Mind. Sher identifies two features of the mental realm that he thinks put our thoughts beyond the scope of morality. The first feature of the mental that Sher appeals to is that rules against actions have discrete boundaries, while rules against thoughts do not. I argue that this problem is equally true of actions and thoughts, meaning that this argument fails to show how thoughts are morally different from actions. The second feature of the mental that Sher points to is that our mental lives are impermeable to and disconnected from other people, meaning that they cannot wrong others. I argue that our thoughts, despite being impermeable and disconnected, can wrong others by inflicting unfelt harms upon them.

In section III, I outline additional objections to Sher’s view. *First*, I argue that Sher’s view should actually be understood as being about the permissibility of any unheard utterance, not just about the permissibility of private thoughts. This clarification, I argue, renders his view implausible. *Second*, I argue that our thoughts can inflict unfelt harms on others, making them sometimes impermissible. *Third*, I argue that Sher’s position on thought-action composites is implausible.

I: Sher's argument:

Sher's thesis is the following:

Wild West of the Mind (WWM): for any purely private thought, T, that is expressed in an agent's, S's, mind, T is not morally wrong.

Sher offers several clarifications of WWM. *First*, he is not denying that one's thoughts "can reflect badly on his character" (484). He agrees that some thoughts suggest that an individual is vicious. *Second*, he agrees that one's thoughts can be epistemically wrong (i.e. epistemically unjustified). With these clarifications, Sher's claim is that "where a person's private mental contents are concerned," condemnation based on viciousness or epistemic wrongness "are the only forms of condemnation that are in place" (484). As Sher later argues, these forms of condemnation are not sufficient to support the view that our private thoughts can be wrong.

Lastly, Sher is not addressing cases where an individual's actions are made better or worse in virtue of her thoughts. For example, imagine a case in which Smith pushes Jones to hurt him and another case where Smith pushes Jones to save him from being hit by a train. Smith's thoughts are clearly morally relevant. It would be implausible to deny that Smith's intentions make a moral difference in how we should evaluate his actions. To avoid this implication, Sher distinguishes between purely private thoughts and thoughts that have both a public and a private component. The former exist only in the mind and have no physical expression in the world, while the latter include both thought and action. In the cases of Smith and Jones, we are dealing with "composite occurrences that have both public and private components" (485). Sher's claim is only that our purely private thoughts cannot be wrong.

II. Problems With Sher's Argument:

Sher defends WWM by arguing that all available arguments against WWM are bad and also by outlining positive reasons for WWM. I will not address Sher's negative argument. Instead, I

object to his positive argument for WWM. In his positive argument, Sher identifies two features of the mental that seem to put our thoughts beyond the scope of morality.

The first feature of the mental that Sher appeals to is that rules against actions have discrete boundaries, while rules against thoughts do not. As Sher says, “when morality or the law forbids me to shoot you, the act that it forbids begins with my decision to pull the trigger and ends with the bullet penetrating your body” (492). However, when it comes to thoughts, “if morality were to forbid me to think of shooting you, the prohibition would inevitably diffuse itself, like dye poured into water, among innumerable other thoughts and feelings” (492). In short, to avoid thinking about shooting someone, one needs to refrain from a host of other thoughts about that person. Sher continues: normal people “can easily resist the transition from the impulse to shoot you to the deed itself,” but most normal people “have far less control over...[their] inferences and associations; so, to guard against thinking about shooting you, I would also have to avoid many other thoughts” (492). If it’s wrong for me to fantasize about shooting someone, I would “have to avoid dwelling on the wrong that [they] have done to me,” because that thought may lead me to think about shooting them (492-493). More generally, “to know which thoughts to avoid, I would have to know which ones might lead to the forbidden thought” and avoid them too (493). So, actions can be morally evaluated, because they have discrete boundaries; but, if we were to morally evaluate our thoughts, we must also evaluate all the thoughts that lead to our purportedly bad thoughts. This seems implausible.

Sher’s criticisms apply equally to actions. To say that it’s wrong for Smith to kill Jones also means that it’s wrong for Smith to do things outside of killing Jones but which will lead to killing Jones. For example, buying a gun, researching how to dispose a body, etc. are all wrong.

So, it seems that actions are just as permeable as thoughts. Since this is not problematic for actions, it should not be problematic for thoughts.

Sher might respond: if Smith buying a gun won't lead to him killing Jones, then buying the gun is not wrong. Only those actions that most likely would lead to the killing are wrong, not the ones leading up to it, because they don't cause the killing.¹

I agree that if buying the gun has no causal connection to Smith killing Jones, then it is not wrong for Smith to buy the gun. But, given Sher's reasoning for the permeability of thoughts, I believe that the problem applies equally to actions. Recall that Sher's reasoning for why the wrongness of thoughts can spread so easily while the wrongness of actions cannot: normal people "can easily resist the transition from the impulse to shoot you to the deed itself," but most normal people "have far less control over...[their] inferences and associations; so, to guard against thinking about shooting you, I would also have to avoid many other thoughts" (492). Sher's argument seems to rely on the claim that the jump from thought X to murderous thought Y is not within our voluntary control, while the jump from action X to murderous action Y is within our control. This explains how blame for thoughts can permeate very far. Wouldn't this same principle be able to apply to some actions? Suppose that Smith knows that if he were to buy a gun, this would vastly increase the likelihood that he would kill Jones. Wouldn't it be wrong for him to buy the gun? Also, it's otherwise permissible for Smith drive down Jones' Street. But, if he does so knowing that this will fill him with uncontrollable rage, leading to him killing Jones, then this is wrong. Using Sher's reasoning, we can show that some actions are equally permeable to thoughts in terms of blame.

¹ I am thankful to an anonymous referee for raising this objection.

One might say that, in general, we have more control over the shift from thought to action than we do over the shift from thought to thought. While that may be true, the important takeaway is that this is not a unique problem for thoughts. For those links between thoughts that are voluntary, we can avoid the problem posed by Sher.²

Lastly, the proponent of my position can bite the bullet say that if we regard thought X as wrong, we should regard all thoughts that knowingly lead to X as being somewhat wrong. Like I said above, this seems like the right answer for otherwise permissible actions that knowingly lead to a bad action.

The second feature of the mental that Sher points to is the fact that our mental lives are impermeable to and disconnected from other people. Sher argues that the mental and the public are different, in that “what is going on in each person’s subjectivity is always independent of, and is often wildly at variance with, what is concurrently going on in his public neighborhood” (494). Given the “impermeability and independence of each person’s subjectivity,” Sher argues that “each subjectivity is almost literally a world of its own” (494). Sher claims that, in light of this gulf between our minds, our thoughts about other people are representations of those people which don’t have moral standing. As he says, “the ‘people’ who populate our mental landscapes are only shadow people, and you can’t have a moral obligation...to a shadow” (494).

² One might object that the jump from one action to the other has more barriers in place than the jump from one thought to the other. For example, as a referee pointed out to me, Smith (in the previous example) may have a lock on his door, he may see a police officer as he drives, etc. All of these are physical objects will likely deter Smith moving from action X to action Y (murder). It seems less clear that there are similar barriers for thoughts. This suggests that there is more control (and thus more responsibility) when Smith moves from action to action than when he goes from thought to thought. I’m not convinced that this is always the case. There are many thoughts that can, so to speak, unlock doors in our minds. Suppose that Smith is a loving spouse who cares very much about fidelity. Smith never fantasizes about his attractive coworker, because he knows that if he were to start doing so, this would unlock a door in his mind to fantasizing about many other women in his life. In the same way that the lock on the door deters Smith from leaving the house, the concern for fidelity deters Smith from fantasizing. I would guess that the concern for fidelity is a more effective deterrent than a locked door. However, even if the reviewer is correct that this difference makes thoughts more permeable than actions, I think it would only show a difference in degree, not in kind. In principle, we can be responsible for thoughts which likely lead to other thoughts and actions which likely lead to other actions. But, it may be the case that this principle applies fewer times to thoughts than to actions.

This is questionable. In the next section, I argue that one's private thoughts can be intrinsically wrong and that they can affect the well-being of others, regardless of whether those thoughts are about shadow people.

III: Reasons Against WWM:

Here, I develop several objections to WWM that Sher does not consider.

i. What's Special About Our Minds?

Consider the following cases:

Hate in the Head: Smith, who hates Jews, privately thinks that the Holocaust was morally justified.

Hate in an Empty Room: Smith, who hates Jews, is in an empty room, on top of an empty mountain, etc., and publically states his belief that the Holocaust was morally justified.

The only difference between the cases is whether the hateful sentence is uttered out loud or merely in Smith's head. We can stipulate that, in both cases, nobody will find out about it, it won't make Smith more likely to do something bad, etc. Sher's view initially seems to suggest that there is a moral difference between these cases. That seems implausible.

Sher might respond that, in Hate in an Empty Room, Smith's words have the capacity to cause harm, while in Hate in the Head, Smith's thoughts lack such a capacity. But, this is not the case. I've stipulated that nobody will hear Smith. We can imagine a version of Hate in the Head in which someone reads Smith's mind and is harmed by his thoughts. Thus, the capacity to harm is not a genuine difference between these cases.³

These cases suggest that the mental realm does no work in Sher's account. If the mental qua mental were doing any work, then we should regard the above cases as morally different.

³ One might wonder whether Sher's view entails that certain thoughts become impermissible only when someone is in an MRI machine and the contents of their thoughts can be discerned. This would be a strange implication.

But, there seems to be no reason for doing so. Thus, on Sher's view, the mental does no intrinsic work.

Contrast these theses:

WWM: for any purely private thought, T, that is expressed an agent's, S's, mind, T is not morally wrong.

Permissibility of Unheard Utterances (PUU): for any utterance, U, it is not wrong for S to utter U, in her mind or in speech, so long as nobody is ever aware that S uttered U.

Sher intends to endorse WWM. But, if he cannot offer a principled reason for regarding Smith's action as permissible in Hate in the Head but wrong in Hate in an Empty Room, then he is really committed to PUU. In other words, when we press Sher's account, it seems that he cannot hold that our thoughts are not wrong *because they are in our mind* but rather must hold that our thoughts are not wrong *because others don't know what we are thinking about*. If so, Sher must revise his position to say that our utterances and feelings, mental or otherwise, are only wrong if they are made known to others.

This has two important implications. First, Sher's position, contrary to his opinion, is not about the mental at all. Second, as I will argue in the coming paragraphs, PUU is false. If Sher is committed to PUU, then his view is false.

Again, PUU says that our utterances, mental or verbal, are not wrong if others don't find out about them. We can see the implausibility of this thesis by considering the following cases:

Joke in an Empty Room: Smith, who is anti-Semitic, says a horribly offensive Holocaust joke in an empty room.

Causal Impotence Hitler Vote: voting is a kind of utterance. Smith lives in Germany in 1933. Smith knows that Hitler will win the election and thus knows that he is causally impotent over the outcome. And, Smith knows that nobody will ever find out about his vote. So, he votes for Hitler, because he hates Jews.

In both cases, we can stipulate that nobody will be directly harmed by Smith's action, and nobody will find out about Smith's actions. However, it seems intuitively clear that there is something morally wrong about both of these actions. The precise explanation of why these actions seem wrong, I contend, is that they involve an agent either endorsing an evil action or expressing and endorsing an evil belief. In *Joke in an Empty Room*, Smith expresses and endorses a morally repugnant belief, and in *Causal Impotence Hitler Vote*, Smith endorses an evil agent and his policies. This judgment can be summed up in the following thesis:

Endorsement: it is prima facie wrong to endorse morally wrong ideas or agents.⁴

The Endorsement principle seems like the best explanation of my intuitions in the aforementioned cases. Additionally, A further intuitive defense can be offered in favor of Endorsement. Consider these cases:

Smith: Smith is a typical person with typical beliefs, none of which are all that objectionable. He is generally nice to people in his life.

Nazi Jones: Jones is identical in all respects to Smith, but he also harbors horribly anti-Semitic beliefs. Although he never acts on these beliefs, Jones essentially subscribes to the Nazi political ideology.

By stipulation, both Smith and Jones will be generally nice people for most of their lives. So, the only difference between them is that Jones is a closeted Nazi. Intuitively, it seems clear to me

⁴ Corvino (2002) endorses a similar view. The argument for Endorsement is that it explains what I believe are clear intuitions in *Joke in an Empty Room* and *Causal Impotence Hitler Vote*; the defense of Endorsement does not rely on any reference to Unfelt Harms. The Endorsement Principle helps to clarify something seemingly paradoxical in my view. If it's wrong to have a certain thought, then isn't it wrong to hold the thought in one's head long enough to judge that it's wrong? According to the Endorsement Principle, what is wrong is not the mere having of a thought in one's head but rather then endorsement of its content. For example, if one were to write a paper about the wrong of rape, one would need to have the concept of rape in one's thoughts. This is permissible, because this person isn't endorsing rape.

that Jones is a morally worse person than Smith.⁵ And, the only explanation for this is that Jones endorses morally wrong ideas.

If this account is true, and if our intuitions in the above cases are correct, then not all unheard utterances are permissible. Thus, PUU is false. And, given that WWM reduces to PUU, WWM is false.⁶

One might object that the Endorsement Principle seems to rely on the claim that our beliefs are within our control, which sounds false. Thus, the Endorsement Principle never gets off the ground. I have several responses to this objection.⁷

First, although it is less commonly endorsed, one could be a doxastic voluntarist of a certain variety. As I see it, doxastic voluntarism is the view that agents exercise at least some control (either direct or indirect) over some of their beliefs. I think that it is beyond the scope of this paper to enter into the debate about the voluntariness of belief. But, what I can say here is that if it turns out that my view requires doxastic voluntarism to be true, this would be a less common, but still defensible position.⁸

Second, it seems clear to me that there is something *prima facie* wrong about endorsing morally wrong ideas or agents. Suppose that Smith had complete control over his political beliefs, and he chose to be a Nazi. This would be wrong. Of course, nobody has direct control over their beliefs. But, what this example shows is that there can be something *prima facie* wrong

⁵ Sher's view is not about one's character, so I don't mean this example to be mainly about character. Instead, I mean that the fact that Nazi Jones has a worse character than Smith is best explained by the wrongness of Jones' mental actions. So, the point here concerns a conduct-based evaluation, not a primarily character-based one.

⁶ Sher might respond that there is still a difference between Hate in the Head and Hate in an Empty Room, namely that verbal utterances have clear boundaries, while thoughts do not. Perhaps the precursors to a private thought are much more difficult to map out, while the precursors to a public utterance are clear. To this, I again reply that I think thoughts and actions are just as porous. Just as I cannot perfectly map out all of the chemical reactions that lead to my thoughts, nor can I do so for actions. Also, actions stem from our mental life, meaning that the permeability of thoughts would have to apply to actions as well.

⁷ Thank you to an anonymous referee for bringing this point to my attention.

⁸ For classic arguments, see Steup (2000) and Alston (1989)

about some endorsements. Now, modify the case to real life and add that Smith became a Nazi believer as a result of causes that he could not control. It seems like the wrongness of Smith's belief might be defeated by the fact that he could not control his beliefs. But, the important insight is that there is wrongness there to be defeated, meaning that endorsements can be *prima facie* wrong.

Why does this matter? This would show that many of our thoughts that involve bad endorsements are *prima facie* wrong but that the wrongness is defeated. This would still run contrary to Sher's position. As I read Sher, he wants to argue that there is no moral valence *at all* with our thoughts. As he says, the mental is "morality-free" (483). Beyond this, my position would also be able to maintain that any endorsements we make which are, in some way, within our voluntary control can be morally evaluated. I agree that many beliefs seem to not be under our direct control. But, many beliefs also seem to be under our indirect control. We can choose to look at (or ignore) evidence against our position, we can choose to seek out people who disagree (or choose not to), etc. These are all ways that we exercise some indirect control over what we believe. To the extent that Smith's Nazi beliefs and endorsements are the result of his indirect control, then he can be blamed for his endorsement of those beliefs.

ii. Unfelt Harms:

Here, I argue, based on the possibility of unfelt harms, that Sher's view is false. Consider two cases:

Chris: Chris lives a normal life, and all of his friends often think positively about him. Chris never finds about his friends' thoughts.

Alastair: Alastair lives a similar life to Chris. But, all his friends are constantly fantasizing about grotesque ways of killing him, stealing his money, etc. They won't ever do these things, but they still fantasize about them. They never speak to each other about these fantasies, and Alastair never finds out about any of this. Alastair desires that his friends think well of him.

Intuitively, it seems that Alastair's life is going worse for him than Chris's. It doesn't seem to matter that neither of them will find out about their friends' thoughts. Even with that stipulation in place, it seems obvious that Chris's life is going better than Alastair's.

The intuition that Alastair's life is going worse than Chris's can be explained by the concept of unfelt harm. Many philosophers have defended the view that individuals can be harmed by actions that never affect their subjective experience.⁹ Boonin advances a sustained argument for the possibility of unfelt harms. Although I lack space to outline Boonin's full argument, it is motivated by cases like this:

Adultery: "Bob wants his marriage to Carol to be monogamous and he believes that it is, but in fact Carol cheats on him regularly" (17).

Most people have the intuition that Bob is being harmed by Carol's adultery, even if he never finds out about it. Or, as Boonin puts it, "if Carol's acts really are harming Bob despite the fact that her acts are having no effect on Bob's mental states...then Carol is inflicting unfelt harm on Bob" (20). I lack sufficient space to launch a full-scale defense of the unfelt harm position, but I take it to be *prima facie* intuitive that, in cases like Adultery, unfelt harm is occurring. At the very least, the objector to this position must say that Bob is not being harmed at all and that our intuitions are being misled by something. This would be a *prima facie* counter-intuitive position.¹⁰

⁹ Feinberg (1984), Parfit (1984), Pitcher (1984), and Boonin (2019) have defended the possibility of posthumous harm.

¹⁰ One might wonder how the adultery case, which involves an action inflicting an unfelt harm, can be used to support the claim that a thought can inflict unfelt harm. The point of the adultery case is to illustrate that something which an agent does not know can still harm him. If I am correct that mental and verbal utterances are morally on a par, and if verbal utterances are actions, then it would follow that both external utterances and thoughts should be considered in the same category as physical acts. And, if physical acts can inflict unfelt harm, then external verbal speech acts and thoughts should be treated in the same way. Put more simply, as long as we agree that something an agent doesn't know about can harm him, I don't see a reason to think that unfelt actions inflict harm while unfelt thoughts or utterances cannot.

If unfelt harms are possible, then Alastair is being harmed by his friends' thoughts. If Alastair can be harmed without knowing it, and as long as we agree that the frustration of our desires can be harmful to us,¹¹ then it follows that Alastair is being harmed by his friends' thoughts. Alastair desires that his friends not engage in fantasies about killing him. Thus, when his friends engage in these fantasies, they frustrate his desires and harm him. Since well-being and harm are moral concepts, it follows that our private thoughts can be wrong in virtue of causing someone to have less well-being and to be in a harmed state.

One might immediately worry that even if we agree that unfelt harms are genuine harms, it needs to be argued that they can constitute wrongs. This is especially important, given that many of our thoughts are not within our control even if they are harmful.¹² My argument that the mental is morally laden is not meant to include involuntary thoughts. For example, intrusive thoughts, sudden thoughts, images popping into one's head, etc. would count as involuntary. On my view, these thoughts may be harmful in some sense, but the harm does not rise to the level of a wrong because the agent has an excuse, namely that the thought was not within her control. This is especially important for people with certain mental disorders, such as OCD, where the agent feels excessive responsibility and guilt over her thoughts. For such individuals, the realization that the majority of their thoughts are not within their control is freeing. I am only concerned with voluntary thoughts. Although many thoughts are not voluntary, the relevant ones are. I want to discuss thought categories like indulged fantasies, prolonged voluntary daydreaming, etc. Any time an agent gets a thought and chooses to indulge it and follow it is a voluntary thought.

¹¹ I am not claiming that desire frustration is *always* bad for us. My argument does not rely on such a strong claim. But, everyone agrees that the frustration of some clearly reasonable desires harms us. And, Alastair's desire for his friends to think well of him is a reasonable desire.

¹² Sher's argument does not rely on the claim that our thoughts are not within our control.

With that clarification, we can now assess the move from unfelt harm to unfelt wrong in the world of thoughts. The basic argument can be made in two ways. *First*, if we agree that unfelt harm is a genuine harm, and if we agree that harms of any kind are prima facie wrong, then we should agree that unfelt harms are prima facie wrong. The burden of proof shifts to the objector to say why unfelt harms are genuine harms but cannot move into the realm of being wrongs. *Second*, there are clear examples of unfelt wrongs. For example, suppose that my neighbor watches me shower every day through the window from his house. Given that he is on his property, he has not trespassed onto my land. He has done nothing that affects my subjective experience. Thus, the best explanation of why this peeping Tom is acting wrongly is some kind of unfelt harm. The burden would then be on the objector to explain this case as wrongful without reference to unfelt harm. One might object that in the previous case, the wrong can be explained by a privacy rights violation instead of unfelt harm. Still, it seems clear to me that the act is not just violating my rights; it also harms me. And, we can devise a case of an unfelt wrong that doesn't have this feature. For example, in the previous case of Alastair, it seems wrong for his friends to be constantly gossiping about him, but he has no right against this. Thus, unfelt harms leading to an unfelt wrong best explain the wrongness of their gossip.¹³

¹³ One might object that my examples jump between being examples of unfelt harms and unfelt rights violations. But, I've tried to lump those into one category and infer from unfelt rights violations to unfelt wrongs. However, it may be the case that unfelt rights violations are harmful and thus wrong, while unfelt harms (which don't violate anyone's rights) do not count as genuine harms. I have several responses to this objection. First, whether it is an unfelt rights violation or an unfelt harm, it still seems clear that, from the third person perspective, Alastair's life is going worse than Chris's. Perhaps unfelt rights impose a more substantial wrong on an agent. But, from the external point of view, the unfelt harm directed at Alastair (even if he has no right against it) makes his life go worse. To put it differently, his life would be going better for him if his friends were not gossiping about him. Both Boonin and Pitcher make a similar move in their defenses of unfelt harms. Second, it seems arbitrary to me to say that unfelt harms do not genuinely wrong an agent while unfelt rights violations do. I agree that unfelt rights violations might be worse, but they are not different in any relevant ways. Both unfelt harms and unfelt rights violations are unknown to the agent, don't affect her subjective experience, etc. It seems strange to say that one is morally relevant while the other is not.

Sher might respond by saying that the subject of the harm is not really Alastair; instead, shadow-Alastair is affected by the gossip. And, harms against shadow-Alastair don't matter. Suppose that an author needlessly kills her main character. This may be unnecessary, but it would be odd to say that it's wrong, because the character is not real. Shadow-Alastair is no different from this fictional character. Thus, we should not be concerned with harms done to shadow-Alastair. In a sense, this objection would allow Sher to agree with the possibility of unfelt harms while still responding to my objection, because he would be saying that the unfelt harm is taking place but is directed at a shadow person, whose interests are not morally relevant. While I see the motivation for this view, I'm skeptical of the ontological commitment involved in positing shadow people. Where do they exist? Do they come into existence whenever I conceive of them? Or, do we have a certain number of pre-existing shadow selves that exist? But, suppose that Sher had a good response to these worries, I still believe that the unfelt harms objection succeeds. If we were to ask Alastair's friends who they are directing their comments to, they would say that they mean them in reference to the real Alastair, not to shadow-Alastair. If Sher is to claim that the gossipers' comments are actually addressed at shadow-Alastair, then he is committed to a highly revisionary view about how these speakers use language. In other words, for Sher's view to succeed, the gossipers would have to be mistaken about who they are referring to as the object of the gossip. My view maintains that they are referring to exactly the person who they claim to be referring to.

iv. Action-Thought Composites:

Sher is keen to point out that his view is only about purely private thoughts, not about thought-act composites. A purely private thought is one that never leads to a corresponding action, while a thought-act composite is a thought that does lead to an action. Recall the cases of Smith and

Jones. In one case, Smith pushes Jones with the intention of saving his life, and in the other case, Smith pushes Jones in front of a train with the intention of killing him. Smith's thoughts clearly make a moral difference. One might be inclined to look at the cases of Smith and Jones and say that this is evidence that our thoughts have a moral valence. After all, the difference between an attempted murder and an attempted lifesaving is Smith's intentions. This might lead one to say that Smith's murderous thoughts are wrong, even if they are never put into practice. Sher denies this; as he says, "to warrant condemnation for maliciously injuring someone, a person must not only harbor the malice but actually inflict the injury. Thus, taken by itself, the claim that it is wrong to give public expression to private malice does not imply that there is anything wrong with simply entertaining the malice" (3).

This is counter-intuitive. Sher is saying that a malicious thought is not wrong on its own. But, the conjunction of a malicious thought and a malicious action is wrong. However, Sher presumably wants to say that the conjunction of a malicious thought and a malicious action is worse than just a malicious action. If Smith and Jones (who is Jewish) are boxing, and Smith hits Jones because he wants to win and because he hates Jewish people, this is clearly worse than Wilson (who is not anti-Semitic) hitting Jones during a boxing match only because he wants to win. So, Sher seems to be committed to the following position: the conjunction of a malicious thought and a malicious action is worse than just a malicious action, but the malicious thought is not wrong on its own. This is puzzling; if the malicious thought is not wrong on its own, then how does it add any wrongness to the malicious action? Sher seems to claim that a malicious thought, which is not independently wrong, somehow contributes to the wrongness of malicious thought-action composites.¹⁴ This view violates the following principle about wrongness:

¹⁴ One might object that there is no such thing as a malicious action without a malicious thought. While I agree in a sense that some actions depend constitutively for their wrongness on the motive behind them, there are certainly still

No New Wrongness: if an action, X, is not wrong, it cannot make some further action, Y, (to which it is conjuncted) more wrong than Y already was.

This principle seems hard to deny. Yet, Sher is committed to denying it, which puts him in a counter-intuitive position. Either he must admit that his view extends to both purely private thoughts and thought-act composites (which he does not want to do), or he must offer an account of the seemingly magical emergent wrongness of malicious thoughts that comes into being when they are added to malicious actions.¹⁵

Sher could object that No New Wrongness is false on the grounds that an organic unity is formed whenever a malicious thought and action are combined. In other words, even if there is nothing wrong about malicious thoughts, there might be something wrong in the state of affairs that combines malicious thoughts and actions, and this wrongness could be more than the wrongness of the malicious action on its own. While this would be a possible response, it doesn't seem more plausible than my view. What is gained by saying that an organic unity is formed when a malicious thought and intention are added together instead of saying that the thought is independently wrong? While I'm not skeptical of organic unities, it seems too convenient to posit them here. At the very least, they should be avoided when possible, and I have provided a way to avoid an organic unity here.

One might object that this argument misunderstands the connection between wrongfulness and the agent's mental states at the time of action. It could be that the person who

bad actions that have no thoughts attached to them. Killing someone without malice is still wrong. So, it seems perfectly reasonable to claim that there can be wrong actions independent of the thoughts that connect with them. But, the criticism above would not apply to actions like lying, because whether something is a lie or not depends on the intentions of the speaker. But, not all actions, or even most, are this way.

¹⁵ Consider the following potential counter-example: Smith swings a bat. This is not wrong. Now, Smith swings the bat at the same speed and intentionally hits a child. Here, an action that was not wrong is added to another action and increases the wrongness of that action. This is not a genuine counter-example. The two actions are X (swinging a bat) and Y (hitting a child). The wrongness of Y is not increased by X. Hitting a child is what makes Y wrong, not the method by which the hitting occurs.

unknowingly pushes Jones to his death is acting wrongly but is not blameworthy, due to ignorance, etc. Perhaps the agent's mental states are relevant to judgments of culpability but not relevant to judgments of wrongness. Thus, my argument in this section may rest on a false presupposition about how mental states and actions interact in terms of moral evaluation.

Although it is controversial, I am sympathetic to subjectivism about moral obligations, which I understand to be the view that actions are right or wrong depending on what the agent believes at the time of action. If this view is true, then attributions of culpability and wrongness go hand in hand. On other days of the week, I'm sympathetic to the ambiguity view, which I understand to be the view that "right" and "ought" can have both objective and subjective senses. So, on this view, there is some sense of rightness that goes with culpability and some sense that tracks the objective facts of the situation. All of this is to say that I think there is good reason to connect wrongfulness and culpability.¹⁶

But, I realize that the preceding is controversial. Given this, I think my view can succeed even if wrongness and culpability are completely separate. If this were the case, I could re-word the No New Wrongness Principle to be about culpability instead. It would then read: if an action, X, is not blameworthy, it cannot make some further action, Y (to which it is conjuncted) more blameworthy than Y already was. I think the same problem for Sher's view could be generated at the level of blameworthiness for thoughts, even if wrongness ends up not being the correct label. Sher would then need to explain how a thought which was not blameworthy suddenly becomes blameworthy when it is added to a wrong action.

IV. Conclusion:

¹⁶ For a helpful discussion of this debate, see Mason (2013), Graham (2010), and Zimmerman (2006).

Sher's argument for WWM fails. Contrary to Sher's view, the mind is not a wild west. There is a moral sheriff who governs our thoughts.

References

- Alston, William P., 1989, *Epistemic justification: essays in the theory of knowledge*, Ithaca, NY: Cornell.
- Boonin, David. *Dead Wrong: The Ethics of Posthumous Harm*. Oxford: Oxford University Press. 2019
- Corvino, John. "Naughty Fantasies." *Southwest Philosophy Review* 18, no. 1 (2002): 213-220.
- Feinberg, Joel. *The Moral Limits of the Criminal Law: Volume 1, Harm to Others*. Oxford: Oxford University Press. 1984.
- Graham, Peter A. "In Defense of Objectivism about Moral Obligation." *Ethics* 121, no. 1 (2010): 88-115.
- Mason, Elinor. "Objectivism and Prospectivism About Rightness." *The Journal of Ethics & Social Philosophy* 7, no. 2 (2013): 2-19.
- Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press. 1984.
- Pitcher, George. "The Misfortunes of the Dead." *American Philosophical Quarterly* 21 (1984): 183-188.
- Sher, George. "A Wild West of the Mind." *Australasian Journal of Philosophy* (2018): 1-14.
- Steup, Matthias, 2000, "Doxastic voluntarism and epistemic deontology", *Acta Analytica*, 15: 25–56.
- Zimmerman, Michael. "Is Moral Obligation Objective or Subjective?" *Utilitas* 18, no. 4 (2006): 329-361.