Virtual Reality as a Path to Self-Knowledge

Lukas Schwengerer

Penultimate Draft - Cite the published version at

https://link.springer.com/article/10.1007/s11229-023-04317-w

Abstract

I discuss how virtual reality can be used to acquire self-knowledge. Lawlor (2009) and Cassam (2014) develop inferential accounts of self-knowledge in which one can use imagination to acquire self-knowledge. This is done by actively prompting imaginary scenarios and observing one's reactions to those scenarios. These reactions are then used as the inferential basis for acquiring self-knowledge. I suggest that the imaginary scenarios can be in principle replaced with scenarios in virtual reality in a way that still provides an inferential basis for self-knowledge. Instead of internal prompting in imagination, I call this external prompting in virtual reality. I discuss the advantages and disadvantages of external prompting. On one hand, external prompting avoids some of the common biases that can intervene with internal prompting in imagination. On the other hand, external prompting comes with some challenges of its own. External prompting might be more time-consuming and might be open to a game-like approach of the agent leading to a different sort of distortion that gets in the way of self-knowledge. I suggest that these are practical challenges, but nevertheless, external prompting seems worthwhile for self-knowledge that is otherwise especially difficult to acquire.

Keywords: Self-Knowledge; Self-Belief; Introspection; Virtual Reality;

Introduction

In this paper, I argue that experiences in virtual reality (VR) can be used to gain self-knowledge.

Drawing on Lawlor's (2009) and Cassam's (2014) discussion of internal promptings as a source

of inferential self-knowledge I suggest that VR can be used in a form of *external prompting*

that can play a similar role to internal promptings as an inference basis for self-knowledge.

Agents can use the tools available with VR technology to generate evidence that allows them

to infer their own mental states, vices, virtues and values. Furthermore, I argue that external

promptings via VR have distinct advantages over internal promptings as a basis for inferential

self-knowledge, although they also come with their share of challenges.

I start with a quick overview of the role internal promptings can play in acquiring self-knowledge, before presenting the alternative of external promptings. I then provide reasons why external promptings might be preferable to internal promptings and consider difficulties for the external prompting route towards self-knowledge.

Inferential Self-Knowledge

For this paper, I understand self-knowledge in a very broad sense. It includes knowledge of one's mental states – as is usually the focus in work on self-knowledge – and knowledge of one's vices, virtues and values. This is a conception of self-knowledge along the lines of Cassam's (2014) notion of substantial self-knowledge. And just like Cassam, my interest in self-knowledge here is aimed at self-knowledge that is difficult to acquire – self-knowledge that requires work. For instance, Lawlor's (2009) paradigmatic example of inferential self-knowledge is about knowing one's desire to have another child. At least in many cases knowing whether one wants another child is anything but easy. Similarly, knowing one's vices and virtues can be difficult. I like to see myself in a good light, but that in itself might make it more difficult to be accurate in assessing myself. I am biased in favour of seeing positives about myself and might turn a blind eye to my vices. Some vices might even make their own detection more difficult (cf. Cassam (2018)). For instance, the arrogant person is likely to overvalue their own judgment and therefore might be unable to recognize their own faulty belief formation.

This focus on more substantial, difficult-to-acquire self-knowledge is based on epistemic considerations. There is no need to look for a new way of acquiring self-knowledge for cases in which our current methods already work well and require little effort. If I am in pain, I know that I am in pain. If I believe that it is raining I likely have no problem knowing that I have that

belief. There is no need to look for a new method of detecting my mental states in these circumstances. Only when self-knowledge is difficult the search for a new method is called for. Hence, I will focus on self-knowledge of mental states, vices, virtues and value that is difficult to acquire.

Before presenting the method of external promptings in VR it is worth looking at Lawlor's (2009) and Cassam's (2014) analysis of inferential self-knowledge via internal promptings because my proposal will draw heavily on their account. Throughout my discussion, I will work with their account. Given my aims of this paper and limited space, I will not defend their account against alternatives, such as agentialist (e.g. Moran (2001)) or expressivist (e.g. Bar-On (2004)) accounts. If one is sceptical of the inferentialist approach to self-knowledge one can read my argument as conditional: if inferentialist accounts akin to Cassam (2014) and Lawlor (2009) are true, then virtual reality can also be used to generate evidence for inferentialist self-knowledge.

Let me start with Lawlor's (2009) example of Katherine who is trying to find out whether she wants another child. The answer is not obvious to Katherine. Moreover, she cannot just consciously scan her mind for a relevant desire. She also cannot wait and use her behaviour to infer whether she wants another child. This would take too long and there is a good chance that at no point in the near future, her behaviour would provide sufficient grounds to answer the question of whether she wants another child. Katherine wants to find out now so that she can plan her actions accordingly. She cannot leave the question open. But what can Katherine do? She can, for instance, remember previous behaviour and feelings when she looks at early pictures of her first child. She can notice thoughts that come to her mind when interacting with children. And she can recognize how she feels when she talks to pregnant friends. These things can constitute evidence for Katherine. Perhaps previous behaviour and feelings point

towards her wanting another child. Sometimes that evidence is sufficient to tell what she desires. However, it is unlikely that this sort of evidence is decisive in Katherine's case. There is too little of it and it is not particularly strong evidence on its own. Something more is needed. Lawlor suggests that imagination can do the trick. Katherine can use imagination to generate more evidence in a routine of *internal prompting*. She can use active self-questioning and prompting to imagine situations that are relevantly connected to having another child. Katherine can imagine herself interacting with an infant again, waking up at night to soothe the crying child, putting on tiny shoes, or carrying the child with her throughout the day. And in imagining these situations she notices herself feeling some way. Feeling a way that makes her want to revisit these imaginings over and over. There is something about these imagined situations that draw her back. A feeling that is not itself stipulated in her imagination. She does not decide to imagine the interactions with a child to feel a specific way. She merely prompts the imagined situation and then allows the imagination to play out. And it turns out that putting herself into these situations comes with feeling a certain way. Thoughts come to her mind in these scenarios not because she consciously decides on these thoughts, but because they seem to come to Katherine as a natural result of being in these scenarios.

Lawlor (2009) suggests that this is exactly the sort of evidence that Katherine can use to answer the question of whether she wants another child. These imagined scenarios come with feelings, thoughts and imagined behaviour that can be used as evidence for her having a desire for another child. By imagining these scenarios Katherine generates evidence that can be used to self-ascribe that she wants another child. Lawlor's suggestion here is that these feelings, thoughts and behaviour in the imagined scenario are caused by Katherine's desire for another child. Hence, what Katherine does when she tries to find out whether she wants another child is a form of inference to the causal basis of all these feelings, thoughts and behaviours that

occur within her imagination when she performs the internal prompting routine. She feels, thinks and behaves a certain way within imaginary scenarios because she has the desire to have another child. And she imagines these scenarios in order to generate more evidence that allows her to infer that desire through its causal consequences. As Lawlor states: "[…] self-knowledge of desire is in routine cases a matter of self-interpretation of one's imagings, where that self-interpretation is a causal inference to the best explanation of one's inner life" (Lawlor, 2009, p. 62).

I want to emphasize that Lawlor's proposal is liberal with sources of evidence available that can be used to infer one's own mental states. Imagining something is not the sole path to self-knowledge. Anything that is caused by a mental state can be potentially used in inference to self-ascribe that mental state. It might be something that is under the agent's control (e.g. imagining something), but it can also be something merely happening to the agent (e.g. a thought that comes to one's mind apparently out of nowhere, or perhaps a particular dream that haunts the agent at night). However, for my purpose, the agent-led path to self-knowledge by imagining something is particularly interesting and can be used for my proposal of external promptings. What makes the imagination path interesting is that an agent seems to have a lot of control over that route to self-knowledge. One can decide to imagine suitable scenarios, almost regardless of the situation one is in. Hence, I focus on imagination within Lawlor's account and bracket other sources of evidence used for acquiring self-knowledge in her account.

For an internal prompting route to self-knowledge not any sort of engagement within imagination will do. It is important what aspect of the imagined scenario Katherine consciously determines and what parts of the scenario are left to play out on their own. Moreover, it is important to notice *how* these scenarios are left to play out on their own. It is Katherine's

imagination and she could make the rules of the imaginary world as she sees fit. To gain self-knowledge she has to refrain from doing so for epistemic reasons.

In general, for considerations in imagination to have a proper epistemic impact some restrictions are needed. These restrictions are meant to explain how we can intentionally better our epistemic position merely by imagining something (cf. Langland-Hassan (2016), Kind (2016)). Following Kind, I hold that imaginings that have proper epistemic impact need to satisfy a *reality constraint* and a *change constraint* to some degree*.* The reality constraint holds that the imagining has to capture the world as is and the change constraint holds that imagining a change in the world has to be guided by the logical consequences of such a change (Kind, 2016, p. 151).[1] When Kind talks about "capturing the world as is" for the reality constraint, she refers to considerations about the difference between the imagined scenario and the real world. When I respect the reality constraint I imagine a situation as closely to the real world as possible. To better illustrate let me take a slightly adjusted example from Kind & Kung (2016): I might imagine an impending alien attack. If my imagining satisfies the reality constraint I imagine a world that is very much like our real world, except for the impending alien attack. I imagine the earth's defences exactly as they are, not as I wish them to be. There are no unnecessary differences between the real world and the imagined scenario. The reality constraint is best captured by asking 'How different is the imaginary scenario from the real world?". And the constraint is respected if the difference is as small as possible, given what I want to imagine.

---

[1] Rarely are both constraints fully satisfied in human imagination. Kind takes them to be more of an aspirational ideal for imagination, but even when they are only partially satisfied imagination can justify beliefs (Kind, 2016, p. 153).

The change constraint on the other hand is concerned with the development of the imagined scenario. It is best captured by asking 'Does the imaginary scenario play out the same way as it would in the real world?'. Take again the example of the alien attack. I can imagine the scenario playing out in many different ways. However, some of them are closer to how the scenario would evolve in the real world, while others are far away from reality. For instance, Kind and Kung (2016) suggest that one could imagine the scenario playing out as humans developing powerful new defence systems very quickly. But that is something not at all within the initial content imagined. There is nothing in our current real-world situation that would lead to such a new defence system functional against an alien attack.[2] If I were to imagine the scenario evolving in this way the change constraint would not be satisfied. The change constraint demands that the scenario plays out realistically – mirroring what would happen in the real world from the relevant starting position. The starting position in the imaginary scenario of the alien attack is such that a sudden new defence technology is not something that would follow in the real world. In the real world, the scenario would not play out this way. The change constraint is violated.

Reality constraint and change constraint can come apart. For instance, I can imagine the alien invasion scenario in ways that violate the reality constraint without violating the change constraint in letting the scenario play out. I might imagine the world as being attacked by aliens and us humans as having futuristic laser weapons, but then the imaginary scenario plays out exactly how it would in the real world if there were both an alien invasion and laser weapons. On the other hand, I can also violate the change constraint without violating the reality constraint. That is the aforementioned case of imagining the impending alien invasion

---

[2] This is merely stipulated in the example.

as similar to the real world as possible, but then letting the scenario play out in a way that differs drastically from how it would play out in reality (e.g. sudden new defence technology).

The reality constraint for Katherine trying to acquire self-knowledge through internal prompting should be understood primarily as a constraint on imagining her mental states as similar to those in the real world as possible, given the imaginary scenario at hand. Of course, if Katherine imagines herself looking at tiny shoes, her perceptual states have to be different than in the actual world in which she currently does not. But to respect the reality constraint, her other mental states ought to be the same as in the real world. She ought not to also imagine herself as having some other mental states that she does not have in the real world.

The change constraint entails that Katherine's imagining has to be constrained by the logical consequences of any changes to the world set in the imagined scenario. If she changes the scenario, but the causal relations within her mental life stay fixed, then she has to imagine the scenario as evolving in a way that is logically consistent with that mental life and the causal relations within it. The way the scenario plays out cannot suddenly introduce new mental states, or new causal relations between mental states and actions. The scenario has to play out in a way that mirrors how it would play out in the real world.

There is some grey area in considering reality and change constraints for Katherine's case. Some cases can be described both as violating the change constraint and as violating the reality constraint. Suppose Katherine imagines an interaction with tiny shoes and the scenario plays out in a way that she begins to cry thinking about her first child. And suppose this is a case of some constraint violation. It could be a violation of the change constraint if the initial scenario was imagined as close to the real world as possible, but the scenario evolved in a way it would not in the real world. However, it could also be a violation of the reality constraint.

Suppose the initial scenario was imagined with an unnecessary change in dispositions related to her mental states. In that case, the scenario evolved exactly in the way it would in the real world, if that real world had mental states with very different dispositions. Whether the case ought to be considered as a violation of the reality or the change constraint seems to be sensitive to the details of how we describe the case. Often cases can be underdescribed in this regard. For my purpose, it does not matter much which constraint is violated. Instead, I want to focus on what both these constraints taken together do for acquiring knowledge from imagination. Together they provide the basis for why imagination has to be *realistic* rather than *fantastic* regarding the causal structure of mental states involved (cf. Gauker (2022)). Let me illustrate that difference generally, before applying it to the particular case of interest.

Suppose I imagine a very thin glass slipping out of my hand and falling towards the ground. A realistic way for that scenario to play out would be for the glass to hit the floor and shatter. This is based on experiences of similar events in real-life situations. A realistic way of imagining the scenario to play out uses the experiences from real life and develops the scenario based on the natural laws and causal relations known from these real-life experiences. A glass slipping out of one's hand in everyday life will fall to the ground and shatter. So, a realistic way of imagining the glass falling towards the floor will also look like that. Realistic imagining includes the same laws and evolves the scenario strictly following those laws, so it will look very much the same. It satisfies the change constraint because the scenario plays out in a way that mirrors how it would play out in the real world. It also satisfies the reality constraint, because the scenario is imagined as similar to the real world as possible. There are no changes in the imagining of natural laws, or dispositions and powers of the objects involved. Neither is anything about the nature of the glass and its component different.

On the other hand, fantastic imagining takes the laws, causality or some object's dispositions known from everyday life and leaves them – or some of them – out of the imagination. Scenarios imagined fantastically evolve in a way they would not in the real world. It does not fully satisfy reality or change constraint anymore. I imagine the glass slipping out of my hand, but then before it shatters on the floor it starts floating in the air. That is clearly something imagined, but it is imagined fantastically. It is imagined in a way in which the laws and causal structure of the real world are not determining how the imaginary scenario plays out anymore. They are replaced by different, imagined laws and causal structures – potentially even with no consistency in how they affect objects. You would not be able to learn what would happen to the glass in the real world by considering fantastic imagination, as it fails to satisfy reality and change constraints.[3]

Imaginings used to infer one's mental states do not have to be realistic in every way, but they have to be in one aspect: the imagining for Katherine has to capture roughly the same causal relations of mental states as are present within the real world. When a mental state in the real-world causes Katherine to be emotional when she hears a crying baby, then that mental state ought to do the same in response to an imagined scenario. Otherwise imagined scenarios would be useless for self-interpretation as a causal inference to the best explanation of one's inner life. The imagined scenarios need to stay true to the causal connections that exist in the real world within one's inner life and between one's mental states and behaviours. The imagined scenarios have to satisfy the reality constraint about causal connections and dispositions within Katherine's mental life, and playing out the scenarios needs to satisfy the change constraint.

---

[3] This is also a case in which the details of how we describe the scenario might matter for which constraint is violated in which way and to which degree. I bracket this question here.

When Katherine imagines an interaction with a toddler, she could decide to imagine herself in the interaction as feeling miserable. She would consciously shape the imagination in a way that determines how she reacts to the situation within the imagination. She could intentionally violate the change constraint. But that would not be helpful if Katherine wants to use her reaction to the imaginary situation as evidence to infer her desire to want another child. When she uses imagination as an internal prompt, she needs to leave open how she reacts to that prompt. She needs the reaction to be a result of the mental state she has in the non-imagined world. This is the sense in which her imagination has to be realistic. The causal relations of the mental state have to be the same in the imagined scenarios as in the real world. She cannot build a different reaction and different causal relations into the imaginary scenario. As soon as she would imagine different causal relations of her mental states her reactions in the imaginary scenario stop being useful to infer a mental state for herself outside of the imaginary scenario. A scenario in which the causal relations of her mental states would be radically different would not amount to an epistemically relevant imagining.

Again, this does not entail that the imagined scenarios as such have to be completely realistic. The imagined scenario can be different in many ways. It might even have some different physical laws. But what has to be imagined realistically are the causal relations of her mental states. A mental state that causes reaction x in situation y, must do so in both the imagined and the real world. Whenever I refer to realistic and fantastic imagining, I will only speak about the imagining being realistic or fantastic with regard to the causal relations of the agent's mental states.

I have now provided a short description of inner promptings as a source of self-knowledge by considering Lawlor's example of Katherine trying to find out whether she wants another child. One can (though does not always) acquire self-knowledge by imagining scenarios and

observing one's feelings, thoughts, and behaviours within these scenarios. Those observations can then be used as evidence to infer a mental state that causes these feelings, thoughts and behaviours.[4] Whereas Lawlor is mostly interested in using this model as a means to gain self-knowledge of mental states, Cassam (2014) also suggests that it works to infer one's vices and virtues, or one's values. Suppose I want to know whether I value honesty highly. The internal prompting story suggests that I can imagine myself in various scenarios in which honesty plays an important role. Scenarios in which I am confronted with other people being dishonest, lying to me in different ways (e.g. directly to me, by omission, bald-faced lies, etc.). Moreover, I can put myself into imaginary scenarios in which I would be tempted to lie. Scenarios in which I would benefit from lying, or would escape bad consequences by some form of dishonesty. Just like in Katherine's case, I can observe myself in all those imaginary scenarios, see how I would feel, what I would do, and which thoughts would come to my mind. Based on that evidence, I can then infer whether I actually value honesty highly.

The model can be applied broadly to mental states, virtues, vices, and values. Recently it has also been applied to group agents acquiring self-knowledge (Schwengerer, 2023). In the next section, I suggest that the role of one's imagination in the model might also be played by simulations leading to virtual reality experiences.

Virtual Reality and External Prompting

In order to discuss my proposal, I have to first explain what exactly I am referring to when I talk about *virtual reality (VR)*. Many of us likely have some idea of VR. We think of special glasses to wear that show us a computer-generated world and special gloves (or some tool we

---

[4] In many cases, the evidence generated by prompting imaginary scenarios is also supplemented by real-world behaviour as evidence.

hold on to) that track our arm movement. The way I talk about VR will be more general, not limited to any specific tool that generates a VR experience. Following Burdea & Coiffet (2003) I use VR as a catch-all term for any simulation in which computer graphics is used to create a realistic-looking world that responds to a user's input in real-time through multiple sensory channels (visual, auditory, tactile, smell and taste). I remain neutral on the exact technological implementation of such an interactive simulation. However, VR can be achieved to varying degrees. A simulation can be more or less realistic-looking – capturing more or less details in the visual representation. It can include different sensory channels to varying extents. For instance, one VR implementation might only be able to simulate a limited range of tactile experiences, while a different implementation might be able to capture a lot more and subtly different tactile experiences.

VR is often characterized by the mnemonic of '*the three I's*': interaction, immersion, and imagination. The first two I's are easy to see within my rough characterization of VR as paradigmatically being interactive and aiming for immersion by being both realistic-looking and allowing for real-time interaction with the simulated environment.[5] The third I – imagination – refers to the developers rather than the user or the VR experience itself. What VR is capable of, which scenarios can be simulated and which problems can be engaged with in VR, is up to the developers. The scope and application of VR are therefore limited by the developers' imagination (Burdea & Coiffet, 2003).[6] All three I's are relevant to my aims in this paper, but for now, these remarks will suffice to give a rough understanding of VR.

---

[5] Immersion and interaction nowadays reached a point at which we replicate real-life scenarios well enough to even allow for rather efficient surgery training in VR (see for instance Spicer & Apuzzo (2003), Thijssen & Schijven (2010), Kim et al. (2017), or Pottle (2019)).

[6] Although the borders can be blurred when the users themselves can impact the creation of a scenario. E.g. when a scenario is generated by AI based on a user prompt.

With this short overview of VR technology as a tool to simulate realistic-looking, immersive and interactive real-life scenarios in place I can now turn to the use of VR in the acquisition of self-knowledge.

My main proposal is a simple idea: if constrained imaginings of particular scenarios can lead one to internal promptings – to one's thoughts, feelings and behaviours in response to that scenario, then it should also be possible to bring a user into a VR scenario that leads agents to observe their thoughts, feelings and behaviours in that VR scenario in a similar way. And if observing one's reactions in imagined scenarios can be used as evidence to infer one's own mental state, then the observations of reactions in VR can also be used in the very same way.[7]

Let me take a look at Katherine again to illustrate this. In Lawlor's example, Katherine engages in the process of acquiring self-knowledge by imagining situations in which she has another child, wakes up at night to feed the baby, puts tiny shoes on the baby's feet and so on. She observes herself in these imagined scenarios. How does she feel when she gently puts the baby back to sleep? What is her immediate reaction and first thought when she wakes up to a crying baby? How does she feel when she has to cancel her plans over and over because of the baby's immediate needs? All her reactions, thoughts and feelings in these imaginary scenarios are then used as evidence to infer whether she wants another child.

I suggest the very same story could be told by replacing Katherine's imagination with the right sort of VR simulation. I call this *external prompting* to contrast Lawlor's story of internal prompting. Consider the following scenario: Katherine is engaging in a fully immersive and

---

[7] There is an alternative line of thought available based on real-world experiences as evidence. If real-world experiences can be evidence for inferential self-knowledge, and VR experiences are similar enough to real-world experiences, then VR experiences should also work as a basis for inferential self-knowledge. If one were to take this path, then one needs to argue about the similarity of VR experiences to real-world experiences. A good discussion of this question can be found in Chalmers (2022). Given the focus on imagination and internal promptings in Cassam (2014) and Lawlor (2009) I will bracket this alternative approach.

interactive VR world designed to help her find out whether she wants another child. Within VR she lives through a scenario in which the baby keeps crying over and over. And she can observe her own thoughts and feelings in that scenario. Perhaps she feels worried about the child and happy when she manages to calm them down. Perhaps she enjoys the moment, even though it is a little stressful. Perhaps she notices that she actually would not even mind these stressful moments that are often taken to be a downside of having a baby at home. This is all evidence that Katherine can use to answer the question of whether she wants another child. And just like in imaginary scenarios, Katherine can go through multiple VR scenarios related to having another child to see how she would react, feel, and think in these scenarios. She can gather more and more evidence. At some point, she likely can tell whether she wants another child.

There is nothing special about imagination itself in the inferential story towards self-knowledge in Lawlor (2009). She already accepts that in some cases internal promptings and the route through imagination are not even needed, when (for instance) real-world interactions provide enough evidence to know one's own desire (Lawlor, 2009, p. 57). Often real-world interactions will not be enough, but they can be. Katherine might visit a friend with a toddler and use that interaction to help her find out whether she wants another child. That by itself shows that for Lawlor's self-knowledge story, it only matters that the subject has enough evidence to infer their own desire based on that evidence. Where exactly that evidence comes from is unimportant. It can stem from real-world interactions or imagined scenarios, or, as I suggest, interactions in VR. What matters is that the evidence allows for an inference to the best explanation to a cause of the evidence – the cause of the feelings, thoughts and behaviours of the subject. Whether these feelings, thoughts and behaviours occur in real-world interactions, in imaginary situations, or in VR does not make a difference.

Cassam (2014) has argued that Lawlor's internal prompting story also works for knowledge of one's own values, vices and virtues. As mentioned earlier, I can use internal promptings to find out whether I am honest. I suggest that this is also true for external promptings. Different VR scenarios could be designed specifically with honesty in mind. Scenarios in which I am lied to and experience consequences because of those lies. Scenarios in which I am tempted to lie for some benefit. Scenarios in which I am tempted to omit a truth because of bad consequences if I were to tell. Whatever scenario one can imagine in which honesty or dishonesty plays an important part, the scenario could very likely be simulated in VR. And if I go through such a VR scenario I can observe my feelings, thoughts and behaviours in that scenario. I can gather evidence to use for inferring whether and to what degree I value honesty. Hence, external promptings seem to be a path towards self-knowledge of values. I take it, that the same sort of story also transfers to vices and virtues. I can infer whether I am open-minded from interacting in VR scenarios just as well as from imaginary interactions. More than that, I suggest that in many cases of vices, virtues and values VR scenarios are a better way to generate evidence than one's imagination. This is what I aim to show in the next part.

In Favour of External Prompting

Imagination is a powerful tool for acquiring self-knowledge. It allows us to create scenarios in which we can look at the causal consequences of our mental states without the limits of the actual world. I do not need to wait to be confronted with an infant to get some evidence relevant to the question of whether I want another child. I can just put myself into that situation in my imagination and I can control exactly how to put myself into such a situation. This allows inner promptings to generate evidence quickly and easily. However, imagination

is also a flawed tool for the very same reason. If I control the scenarios I imagine, I can also bias these scenarios. I can bias them in two distinct ways. First, by cherry-picking imaginary scenarios; and, second, by tampering with the causal structure of myself within the scenarios. I go through them in turn.

Cherry-picking scenarios is straightforward. I can control internal prompting by deciding which imaginary scenarios I put myself into. For the internal prompting routine to lead to self-knowledge of x the scenarios have to be sufficiently related to x. For instance, if Katherine wants to find out whether she wants another child, the scenarios need to be such that they elicit reactions that are potentially related to the desire of wanting another child. However, scenarios do not only have to be related, they also need to be chosen adequately. They need to capture a wide array of situations in which a desire of wanting another child might manifest in some way. Moreover, they also need to capture situations in which a desire not to have another child might manifest in some way. Too few scenarios[8], or a bad balance in the choice of scenarios will prevent an inference to the best explanation from working properly.

The internal prompting routine involves an inference from Katherine's reactions in imaginary scenarios to a desire that causes those reactions. If the scenarios are badly chosen then the inference cannot successfully lead to self-knowledge, because the inference basis is not good enough. For instance, if Katherine would only go through two scenarios that would clearly underdetermine the attitude Katherine has. Too many attitudes are compatible with the reactions in just those two imagined scenarios. Suppose Katherine considers the following two

---

[8] A reviewer points out that knowledge based on causal inferences might not need a lot of data. It seems to me that at least in the case of inferring a mental state, one usually needs a substantial amount of data. Lawlor herself points out how difficult it is for Katherine to find out whether she wants another child. Lawlor also presents us with similar mental states that need to be ruled out to finally infer the right mental state (e.g. wanting the same child again). The general question of how much data is required is an important one, that I do not have the space for.

scenarios: one in which a younger, toddler-aged, version of her son tries to put on his socks; and one in which a younger, toddler-aged, version of her son snuggles up to her and falls asleep. In both scenarios, she might react with a feeling of happiness and comfort. Perhaps she says to herself "I want that". Even when the reactions to both scenarios point in a very similar direction, it is left open whether she wants another child, or to experience something similar again with her son who outgrew the toddler age. The reactions to these two scenarios are not enough to decide between these options. Katherine needs more evidence.

A similar problem of choosing scenarios occurs if Katherine only goes through scenarios in her mind in which having another child is almost universally considered stressful. Because she is not going through any scenarios which would lead to happy experiences for most parents, it seems that any inference will be unlikely to grant her self-knowledge.[9] Katherine needs to imagine a variety of scenarios, such that her reactions in these scenarios will suffice as a basis for an inference to her desire for another child.

These constraints become especially important when one attempts to gain self-knowledge of attitudes, character traits or values that fall under normative considerations. Suppose I want to find out whether I am humble. Following Cassam (2014), one can use an internal prompting routine to go through scenarios in which one's humbleness – or lack thereof – shows in one's reactions, thoughts, or feelings. The problem of an internal prompting story here is that humbleness is generally thought of positively. I want to be seen as humble and I want to see myself as humble. And the natural worry is that I might bias the evidence that I am aware of, or bias my evaluation of such evidence, in favour of self-ascribing humbleness. I want to see

---

[9] One might suspect that the question of which scenarios someone chooses is already also a guide to what attitude the person has. I will bracket this.

myself as humble, so I might look only at the evidence that makes me seem humble, ignoring or misinterpreting all other evidence.

This behaviour has been well documented in various forms of a self-serving bias (see for instance Miller & Ross (1975), Bradley (1978), Pronin et al. (2002) Mezulis et al. (2004)). For internal prompting routines, this bias might cause me to only go through imaginary scenarios that are compatible with an inference to my humbleness. I simply do not put myself into imaginary scenarios in which I would not be humble at all. My self-serving bias leads to cherry-picking an inappropriate basis for any inference to my humbleness (or lack thereof). This is one of the reasons why self-knowledge of virtues and vices is so difficult to acquire. The inner prompting routine is volatile and can be influenced by epistemic vices and virtues. Some vices are *stealthy* (Cassam, 2018) such that the vices themselves get in the way of the inner prompting routine that could detect them. But even self-knowledge of attitudes or traits that are not stealthy by themselves is difficult because biases and vices can always impede the inner prompting routine by selectively choosing the scenarios one goes through in imagination.

Even if the right imaginary scenarios are chosen, there will still be room for biases to get in the way of self-knowledge. Not only the choice of scenarios is under one's control, but the way these scenarios proceed is also within one's control. As discussed earlier, to lead to self-knowledge, they have to proceed in a realistic way, not in a fantastic way (cf. Gauker (2022)) as far as the causal relations of mental states are concerned.[10] This has important implications for internal promptings as a means to gain self-knowledge. Any inference to a real-world mental state or character trait based on one's reactions to imagined scenarios can only work

---

[10] As mentioned earlier, this does not entail that the imagined scenario has to be fully realistic. All that is required is a realistic imagination about the causal relations of my mental states.

if these imagined scenarios capture the laws and causal structure of the mental state or character trait in the real world. These inferences require realistic imagining.

Sometimes it is easy to tell whether I imagine something in a realistic or a fantastic mode. The floating glass is obviously fantastic. However, there are more subtle ways of imagining fantastically that are a problem for internal prompting routines. Take the example of wanting to know whether I am honest – not in a particular instance, but honest in general. I can imagine scenarios in which I am given the choice of being honest and then look at my responses to those scenarios. Moreover, I can imagine scenarios in which I am dishonest and observe my feelings and reactions to these scenarios. And then I take all these reactions and responses as evidence to infer whether I am an honest person. Can I be sure that I imagine these scenarios realistically? How would I be able to tell whether my response in the imaginary scenario is similar enough to how my response would be in an actual real-life scenario of that kind? I might be able to identify some very out-of-character reactions in imaginary scenarios as fantastic, but it seems difficult to argue that I would be able to tell for all imaginary scenarios whether I imagine them realistically or fantastically. It does not seem far-fetched that my self-serving bias steers me towards fantastic imagination in a way that makes me look like an honest person to myself. When I imagine a scenario in which it would be beneficial to lie, I might build into the imagined scenario that I would react by refusing to lie on principle. I can do that, even though I would not respond to a similar scenario in real life the same way. In a real-world scenario, I would lie. What happens here is that I imagine the scenario with a particular built-in response that is *fantastic*. It is a response that does not fit with the causal relations of my actual dispositions and mental states. I can deceive myself by using fantastic imagination instead of realistic imagination, without noticing. The internal prompting routine

has no safeguard against this subtle way of imagining fantastically and the common self-serving bias makes it likely that at times I imagine fantastically.

External prompting avoids both problems completely. The VR scenario is not under the control of the potential self-knower. Here the choice of scenarios can be regulated from the outside. This could be done either by other people playing a role in selecting the scenarios, or (in a somewhat futuristic idea) by using some form of artificial intelligence as a scenario selection mechanism. Moreover, because external prompting uses immersive VR scenarios instead of imagination, there is no danger of using fantastic imagination at all. The immersive VR experience should trigger the same dispositions as real-life scenarios and therefore provide a good basis for the inference required to gain self-knowledge. External prompting avoids one's imagination and outsources the role of the imagination to the VR set-up. Hence, it avoids the problems that stem from one's control over one's imagination.

Furthermore, because at least behavioural reactions in VR are observable by others and recordable, it is in principle possible for other people to also infer the state of the potential self-knower. If I want to find out whether I am a humble person and I go through VR scenarios to generate evidence about my humbleness, then much – though not all[11] – of the evidence is also in principle accessible by other people. They can confirm my self-ascription of humbleness, or challenge me to reconsider. Perhaps they ask me to go through more external promptings if they take my assessment to be faulty.

Overall it seems that external prompting can help us avoid some of the difficulties with internal prompting. However, even though external prompting has these advantages over

---

[11] Some of my reactions to VR situations will be internal to me, e.g. my feelings in a particular situation.

internal prompting concerning biases getting in the way of self-knowledge, external prompting also has its share of problems and challenges. I go through these in the next part.

Against External Prompting

I argued that we should worry about biases getting in the way of internal prompting routines, and that external prompting gives us a better chance of detecting and avoiding these biases in looking for self-knowledge. Unfortunately, something might also get in the way of a proper external prompting routine. The main worry for external prompting is that reactions to situations in VR are not similar enough to the reactions one would have in situations in real life. In a sense, this is very similar to the worry about fantastic imagination for internal promptings. Fantastic imagination is not similar enough to the real world to allow for inferences to a real-world attitude or vice. One should worry that VR has the same problem because an agent in a VR scenario knows it is VR and might intentionally act differently because of that knowledge. This would not be all that surprising, given that we know from VR games that people can put themselves into a different role, divorcing the VR experience from any ordinary real-world experience. I can act as if I were someone else in VR. If one creates this gap between the VR experience and one's regular self in this way, then the VR experience will not provide one with any evidence of one's actual mental states outside of VR.

Of course, an ideal self-knower would not create this separation between themselves and the VR experience, but an ideal self-knower would also not be biased in internal promptings, so this is not going to alleviate the worry. However, there is a difference between the two cases of interference. The worry for internal prompting is that biases interfere. Biases are unconscious and difficult to detect. On the other hand, the worry about external prompting is that a game-like attitude of the VR user will interfere. Game-like attitudes are usually

conscious and not difficult to detect at all. I know when I play a different role instead of being just myself. This by itself already gives the advantage to external prompting, because it is easier to tell whether I perform external promptings to properly assess myself. Of course, merely being able to detect that I play a role rather than be my usual self might not be enough. Being able to tell and stopping a game-like attitude are two different things. Nevertheless, the easier detection seems to be a point in favour of external promptings – even though not a decisive one.

An additional point in favour of external promptings is the possibility for other people to observe the potential self-knower and potentially detect foul play. This is limited to some degree. Outsiders can observe the behaviour of an agent in VR, but they cannot observe the feelings and thoughts of the agent. Nevertheless, in internal promptings, outsiders have nothing to work with to confirm that the potential self-knower is self-ascribing correctly. External prompting allows for the outside perspective as a safeguarding mechanism. That additional security was not available for self-knowledge from internal promptings.

There are some additional ways of preventing VR users from treating the experience as a game-like experience. First, sufficiently high immersion and presence in a real-world-like experience help. The closer the VR experience is to real life, the more difficult it will be to treat VR experiences as something distinct from real-world experiences. Likely it is still possible, but it takes more effort and time.

VR with high immersion that fosters a strong feeling of presence is already used successfully in psychological research, even when the subjects are fully aware that they are merely in a virtual world. Immersion captures the degree to which the VR system delivers experiences that are (among other things) vivid, interactive, and extend to multimodal sensory stimuli (cf.

Slater et al. (1996)). Presence captures a psychological feeling of "being there" in the virtual world (Slater et al., 1996, p.165). Subjects who show a high degree of presence in an immersive VR environment respond to virtual stimuli directly and quickly. Such reactions seem difficult to circumvent based on beliefs about it merely being a VR situation. Mel Slater describes this in a commentary as follows:

> Of course no one, not even when they are standing by a virtual precipice with their heart racing and feeling great anxiety, ever believes in the reality of what they are perceiving. The whole point of presence is that it is the illusion of being there, notwithstanding that you know for sure that you are not. It is a perceptual but not a cognitive illusion, where the perceptual system, for example, identifies a threat (the precipice) and the brain-body system automatically and rapidly reacts (this is the safe thing to do), while the cognitive system relatively slowly catches up and concludes 'But I know that this isn't real'. But by then it is too late, the reactions have already occurred. (Slater, 2018, p. 432)

Even recognizing that it is merely a virtual environment does not stop the subject from reacting to that environment, if the right sort of presence is established for the subject. Of course, Slater's example of perceiving a threat is one in which the reaction is especially quick. One might be worried that most reactions used in inferring one's mental states are unlike the example of reacting to a threat. However, the examples that Lawlor uses to illustrate her account are very similar to immediate reactions. For instance, Lawlor uses the example of memories coming to Katherine's mind when she puts her son's now-too-small clothes away. There is an immediate reaction that suddenly occurs in the interaction with the clothes. The reaction is there, even if she then discards that reaction shortly after. There is that moment of genuine reaction that can be used as evidence. In a similar vein, a reaction to a VR scenario

can be used as evidence to self-ascribe a mental state, even when shortly after that reaction one thinks that it is just a simulation. A sufficient feeling of presence ensures that there will be genuine reactions and gives the subject the evidence they are looking for.

High immersion and presence also give us a distinctive advantage of VR over the use of other external tools as providing prompts useful for self-interpretation. For instance, a narratively presented scenario in a book, or an engaging scene in a movie could also be used as some form of prompt akin to an imagined scenario. Katherine might watch a movie with a particularly emotional scene about a baby. When watching the scene, thoughts and emotions come to her mind. Perhaps she starts thinking about soothing the baby if she were there. And she might find out that she wants another child at that very moment by observing her reaction to the movie scene. This is a legitimate path towards self-knowledge. However, reading a story, or watching a movie is rarely able to create immersion or a feeling of presence that can match modern VR scenarios. And if immersion and presence are lacking, then genuine reactions that constitute evidence are likely also lacking. I am not ruling prompting with texts or movies completely out. However, VR seems to provide a better chance of generating adequate evidence.

Closely connected is the second way to prevent game-like attitudes from influencing reactions in VR. VR scenarios can be intentionally built such that the reactions of VR users are as spontaneous as possible. As Slater states, if we present an agent with well-designed VR scenarios "the cognitive system relatively slowly catches up" when the reactions have already occurred. The idea here is to prompt quick reactions. This can be done by ensuring that a user does not know the scenario in advance such that they cannot prepare a response. It can also be done by creating a sense of urgency in the user with an appropriately designed scenario,

such that they spend less time on thinking through their actions.[12] Scenarios designed in that way make it more difficult for game-like attitudes to determine the user's reaction. Giving the user less time to consciously interfere with reactions gives them less chance to get in the way of a genuine reaction. This idea is well-known from attempts at assessing attitudes, for instance in the implicit association test (cf. Greenwald et al. (1998); Greenwald et al. (2003), Brownstein et al. (2020)). Prompting more spontaneous responses is also no guarantee to avoid game-like attitudes completely, but it reduces the risk of such attitudes.

A reviewer worries that agents can adopt very general game-like attitudes about the VR experience. The idea is that I am going with a game-playing attitude into VR and whatever happens in there is automatically something I will have a game-like attitude towards. However, adopting a game-like attitude in this way towards whatever one might experience next without any restrictions seems rather difficult. It seems difficult because game playing always happens with the ability to stop playing if required. Agents need to be able to quit their game-like attitudes when the circumstances require it. And dangerous circumstances are identified in part in virtue of the agent's experiences. If I were to have a game-like attitude towards whatever I experience next, without any restrictions, then I would be insensitive to all kinds of dangers and emergencies that require me to stop the game-like attitude. Hence, game-like attitudes need some restrictions based on reasonable expectations of what a game experience would be, contrasted to a non-game experience. This necessity to escape game-like attitudes works directly against a blanket game-like attitude towards everything that happens in VR. This also fits with Slater's description of the cognitive system catching up. The perceptual system presents the agent with something that is not from the get-go perceived in

---

[12] It might also be done by simply instructing users to react as quickly as possible.

a game-like attitude. The cognitive system has to get involved and – perhaps – tells us that it is just a game quickly after.

Both ways to answer the worry from game-like attitudes build on an assumption that the immediate reactions to VR scenarios are evidence of one's mental states. In the background here is a picture in which mental states either are – or at least come – with a set of dispositions, some of which are triggered in the VR scenario. One might respond that this is not the right way to think about the mental states in question. This can be illustrated with the implicit association test. One might argue that the quick responses seen in such a test are not measuring attitudes. Instead, the attitudes can be found in reflective judgments that agents make after taking time to think. I take it, that there is a difficult question here and I am not able to answer it sufficiently. However, I do not have to. Lawlor's model already gives evidential weight to immediate reactions – to such phenomena as thoughts just showing up in one's mind, or feelings prompted by a memory. Those are not slow, deliberate, reflective processes. Hence, if my starting point is Lawlor's account, it seems justified to assume that immediate responses can be evidence of one's mental states.

A second worry about external prompting concerns the influence the designers of VR scenarios have. Which VR scenarios are part of the external prompting routine depends on the ability and imagination of the people creating these VR experiences. For internal prompting, I argued that a bias in selecting the imaginary scenarios can prevent the internal prompting routine to function properly as a guide to self-knowledge. The very same problem is present when the VR scenarios are designed by biased programmers, or a flawed algorithm (in case of scenarios created by artificial intelligence). There is no fully satisfying solution to this problem. The best available option is for a diverse group of people to check the scenarios or code for any sign of bias. On one hand, this seems still better than the situation in internal

prompting. For external prompting, we have a way of detecting badly chosen VR scenarios, whereas for internal prompting we do not have any way to detect badly chosen imaginary scenarios at all. On the other hand, it leaves one vulnerable to other people in a way that internal prompting does not. Someone with malicious intent might manipulate the external promptings, whereas it seems difficult to imagine that internal promptings are maliciously manipulated in the same way. That creates a challenge for external prompting routines, but one that can be answered. It raises the stakes for the selection of suitable programmers for external prompting scenarios, but nonetheless, it seems possible to find appropriate ones.[13] Perhaps more difficult, but still possible.[14]

A third worry about biases creeping into the external prompting routine relates to biased attention.[15] Plausibly, any scenario we are in – both in reality and in VR – is too complex to attend to everything available to our perception. Agents make decisions – mostly unconsciously – on what to attend in a scenario. Biases might determine what we attend to in a way that prevents us from attending to objects or events that would lead to evidence for having a particular mental state. If Katherine had a bias that prevented her from attending to baby clothes, then she would not have particular responses to those clothes and would lack the evidence constituted by those responses. Without the evidence, she might be worse off

---

[13] A friend of internal promptings can propose here that the same goes for virtuous internal prompters. It might be difficult to be unbiased in internal prompting, but it is not impossible. The problems with internal prompting and the problems with external promptings are on par and we should not prefer external over internal prompting routines. There is truth in that proposal. Internal promptings are not doomed to be biased and perhaps one can learn to be a better internal prompter. I do not deny that possibility. All I suggest is that external prompting is an alternative option that we should explore because all things considered it might turn out to be easier to deal with the practical challenges of external promptings in VR than to learn how to be a virtuous internal prompter. For instance, it seems easier to deal with a self-serving bias by involving other people in the construction of VR scenarios than to learn how to imagine without a self-serving bias. Thank you to a reviewer for pointing to this possible defence of internal prompting.

[14] At this point, I do not have a concrete proposal for how exactly we filter between better and worse VR designers for this purpose. I leave this practical challenge for future discussion.

[15] Thanks to a reviewer for suggesting the worry.

in trying to infer her desire for another child. She might fail to make the inference because the evidence is missing.

I think this worry is right and I cannot rule out biased attention in VR scenarios. However, it is no different in internal prompting routines. When I imagine scenarios, the same kind of biases also influence how I imagine the scenarios – what part of them I attend to. Prima facie it seems that these attentional biases are built into inferential accounts of self-knowledge in general. VR scenarios might even be in a better position than competing ways to find evidence, because the VR scenario can be designed by someone else, such that the design combats attentional biases. For instance, if the designers have a suspicion that a subject will likely not attend to feature x, the designer can adjust the scenarios to make feature x more prominent in the scenario. They cannot rule out attentional biases, but they can make it more difficult to ignore parts of the scenario due to biases.

A fourth worry is that external prompting requires a lot more time than internal prompting. This worry comes in two different forms. First, it takes a lot of time to develop and design suitable VR scenarios. Once developed they can be reused, but the first design takes time and effort. Ideally, the scenarios can be generated procedurally by artificial intelligence, although this does not seem to be an option currently.[16] And, second, it takes a substantial amount of time to go through VR scenarios to create reactions that can then be used as a basis for inferential self-knowledge. However, this is less of a problem than it might initially look like. Self-knowledge by internal prompting can also take a long time. Of course, some self-knowledge comes quickly and easily, but that is not the self-knowledge I am interested in. The whole point of looking into external promptings is to find an alternative for cases in which self-

---

[16] However, the success of ChatGPT already indicates that it could be possible in the somewhat near future.

knowledge is difficult to acquire with our usual methods. And in those difficult cases using internal promptings can also take time until one acquires self-knowledge. Take Katherine again, who wants to know whether she wants to have another child. It seems plausible that finding out that she wants another child is not done by imagining a couple of situations quickly. She has to perform inner promptings over and over again[17] – perhaps while being in different moods. It is difficult to tell whether she wants another child and hence it takes time to find out. Going through external promptings over a longer time is therefore by itself a challenge, but not a decisive argument against external prompting routines. It merely gives us reason to be selective when we choose an external prompting route. Not every self-knowledge has to be based on external prompting. My suggestion is rather that in some cases of difficult to acquire self-knowledge the external prompting routine is a good choice. A choice worth the time and effort required.

The final worry concerns the privacy of one's own mind. External promptings are in principle accessible by other people in a way that internal promptings are not. As I argued, this comes with some advantages. Other people can be a safeguard, preventing biases or conscious interferences from getting in the way of self-knowledge. However, it also comes with a very serious problem. Many substantial propositional attitudes concern very personal affairs. Katherine's desire to have another child is not something that she likely wants to be shared with just anyone. Agents often have and want control over who knows about their mental states – or at least some of their mental states. It should be up to the agent to decide who gets access to designing the scenarios for external promptings and who gets to observe the agent's reactions in these VR scenarios. Hence, there need to be safeguards in place to prevent

---

[17] This would still be a lot faster than waiting for real-life situations to arrive in which a desire for having another child would manifest in her reactions in some ways.

people to access one's external prompting routine without one's permission. While designers

of VR scenarios need to have access to the programming code, they do not need access to

recordings or data that include an agent's reactions in VR scenarios, unless the agent wants

them to have that access. This is something that ought to be enforced by technological means

and sanctions for any access without permission. But again, it seems to be more of a practical

challenge than a principled problem with the external prompting route.[18] All the worries

raised here are such that they can be answered if we design the tools for external promptings

properly.

Conclusion

I have argued that external prompting as a source of self-knowledge is an alternative option

to internal promptings. Moreover, it is an option that has some advantages in avoiding biases

that might get in the way of a proper internal prompting routine. As I have argued, external

prompting still comes with challenges. Nevertheless, external prompting is overall a viable

alternative to internal promptings for substantial self-knowledge that is otherwise difficult to

acquire. External prompting routines can be performed better or worse. Ideal external

prompting routines are based on VR scenarios that are as interactive and immersive as

possible. Moreover, they capture a wide variety of scenarios relevant to the mental state,

character trait, or value the agent wants to detect in themselves. The external prompting

routine has to elicit reactions in a way that makes it difficult for agents to consciously

---

[18] Though, as a reviewer correctly observes, it is not clear where and how exactly the line between practical and principled problems should be drawn. I do not have a general answer for the border between them. However, I take a practical problem in the context of this paper to be one that can be resolved by implementing suitable and already available processes. We can pick and train programmers for VR scenarios that can do what we aim for in external prompting. A principled problem on the other hand is one that cannot be resolved by implementing suitable and already available processes. The self-serving biases being problems for the internal prompting routine is principled, because at the moment there is no suitable process available to circumvent those biases. Self-serving biases are so ingrained in our psychology that it is difficult to see what could remove them from the internal prompting routine.

manipulate their reactions, e.g. by creating scenarios in which agents react spontaneously. Moreover, the VR program has to be such that it is difficult (and punishable) for anyone other than the agent to observe the agent's reactions in the VR scenarios without permission from the agent. Hence, even though I suggest that VR technology can be the basis of an external prompting routine that helps one to acquire self-knowledge, it is only under certain constraints that VR can and should be used as a method to gain self-knowledge. The optimistic prediction is that many of the constraints on good VR-based self-knowledge can be satisfied. The development of artificial intelligence might allow for a quicker generation of relevant VR scenarios, and an easier way of giving control over the evidence generated in VR to the self-knowing agent. If few to no other people have to be involved in an external prompting routine it seems easier to keep the evidence generated in VR private to the self-knowing agent. However, at this point, these options seem far away. What is called for in the current stage of research is an experimental approach towards VR-based self-knowledge. In principle, all the tools needed are in place and it is worth testing the proposal in practice. And there is a need for new methods of self-knowledge that is otherwise very difficult to acquire. Many intellectual vices (and some propositional attitudes) are elusive when we use internal promptings. It is difficult to gain self-knowledge of one's arrogance, dishonesty, carelessness, dogmatism, or racism. And it is also difficult to accept outside testimony about these vices. Perhaps external prompting can be a way out that allows us to recognize our flaws in a better way.

Acknowledgements

Funding

References

Bar-On, D. (2004). *Speaking My Mind.* Oxford: Oxford University Press.

Bradley, G. W. (1978). Self-serving biases in the attribution process: A reexamination of the fact or fiction question. *Journal of Personality and Social Psychology, 36*(1), pp. 56-71.

Brownstein, M., Madva, A., & Gawronski, B. (2020). Understanding Implicit Bias: Putting the Criticism into Perspective. *Pacific Philosophical Quarterly, 101*(2), pp. 276-307.

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology* (2nd ed.). Hoboken, New Jersey: Wiley-Interscience.

Cassam, Q. (2014). *Self-Knowledge for Humans.* Oxford: Oxford University Press.

Cassam, Q. (2018). *Vices of the Mind: From the Intellectual to the Political.* Oxford: OUP.

Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy.* New York: Norton & Company.

Gauker, C. (2022). On the Difference Between Realistic and Fantastic Imagining. *Erkenntnis, 87*, pp. 1563–1582.

Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, pp. 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), pp. 197-216.

Kim, Y., Kim, H., & Kim, Y. O. (2017). Virtual Reality and Augmented Reality in Plastic Surgery: A Review. *Archives of Plastic Surgery, 44*(3), pp. 179-187.

Kind, A. (2016). Imagining Under Constraints. In A. Kind, & P. Kung (Eds.), *Knowledge Through Imagination* (pp. 145-159). Oxford: Oxford University Press.

Kind, A., & Kung, P. (2016). Introduction. In A. Kind, & P. Kung (Eds.), *Knowledge Through Imagination* (pp. 1-37). Oxford: Oxford University Press.

Langland-Hassan, P. (2016). On Choosing What to Imagine. In A. Kind, & P. Kung (Eds.), *Knowledge from Imagination* (pp. 61-84). Oxford: Oxford University Press.

Lawlor, K. (2009). Knowing what one wants. *Philosophy and Phenomenological Research, 79*(1), pp. 47-75.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is There a Universal Positivity Bias in Attributions. *Psychological Bulletin, 130*(5), pp. 711-747.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82*(2), pp. 213-225.

Moran, R. (2001). *Authority and estrangement.* Princeton: Princeton University Press.

Pottle, J. (2019). Virtual Reality and the Transformation of Medical Education. *Future Healthcare Journal, 6*(3), pp. 181-185.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The Bias Blind Spot: Perceptions of Bias in Self Versus Others. *Personality and Social Psychology Bulletin, 28*(3), pp. 369-381.

Schwengerer, L. (2023). Collective Vice and Collective Self-Knowledge. *Synthese, 201*(Article 19).

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*(109), pp. 431–433.

Slater, M., Linakis, V., Usoh, M., & Kooper, R. (1996). Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess. *VRST '96: Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pp. 163-172. doi:10.1145/3304181.3304216

Spicer, M. A., & Apuzzo, M. L. (2003). Virtual reality surgery: neurosurgery and the contemporary landscape. *Neurosurgery, 52*(3), pp. 496-497.

Thijssen, A. S., & Schijven, M. P. (2010). Contemporary virtual reality laparoscopy simulators: quicksand or solid grounds for assessing surgical trainees? *199*(4), pp. 529-541.