

# Hot-Cold Empathy Gaps and the Grounds of Authenticity

Grace Helton and Chris Register<sup>1</sup>

*Penultimate draft.*

Forthcoming in *Synthese*, Special Issue on Transformative Experience, Authenticity, and Rationality

*Key words:* empathy, hot-cold empathy gaps, authenticity, transformative experience, subjective knowledge, self-knowledge, rationality, understanding, philosophy of cognitive science, epistemology, seductions of clarity

*Abstract:* Hot-cold empathy gaps are a pervasive phenomena wherein one's predictions about others tend to skew 'in the direction' of one's own current visceral states. For instance, when one predicts how hungry someone else is, one's prediction will tend to reflect one's own current hunger state. These gaps also obtain intrapersonally, when one attempts to predict what one oneself would do at a different time. In this paper, we do three things: We draw on empirical evidence to argue that so-called *hot-cold empathy gaps* arise when one projects one's own current state into a simulation about another. Second, we argue that this process does not typically confer knowledge, even when the predictions it produces happen to be accurate. Third, we suggest that these results can be used to develop a challenge for L.A. Paul's view that authentic action sometimes requires subjective knowledge of one's own values and how these values relate to relevant outcomes. We then sketch an alternative view of the epistemic grounds of authenticity, one on which authenticity requires a kind of understanding. The relevant form of understanding can be achieved by subjective knowledge but can also be achieved otherwise, such as through testimony from a close friend about what one values.

It is a platitude that you shouldn't go grocery shopping while hungry. The thought is that if you do, you risk leaving the store with a cart brimming with more food than you could possibly eat, much of it of little nutritional value. As it turns out, empirical evidence vindicates the advice encapsulated by this platitude. It suggests moreover that the wisdom of that advice is likely explained, at least in part, by a broader phenomenon involving how subjects in *hot states*, i.e., affective states, make predictions about their behavior in relevant

---

<sup>1</sup> Both authors contributed equally.

cold states, i.e., non-affective states. In particular, subjects in relevant hot states tend to predict that, even were they *not* in that state, they would behave in a way that is congruent with being in that state. For instance: people who are currently hungry overestimate how much they would want to have spaghetti for breakfast.<sup>2</sup> Notice that this prediction is skewed ‘in the direction of’ the subject’s current state of hunger.

This effect also occurs in the opposite direction to a similar degree.<sup>3</sup> Subjects in cold states are poor at predicting how they would act if in some complementary hot state. For instance: when considering the distant possibility of performing an embarrassing dance, people tend to underestimate the effects of social anxiety and overestimate their willingness to perform.<sup>4</sup> Notice again that the ‘direction’ of the misprediction is skewed ‘in the direction of’ their current state.

The examples so far involve *intrapersonal* predictions about oneself at other times, but one’s current hot or cold state also tends to affect predictions about how *others* would act, i.e., *interpersonally*. For instance: people who are thirsty tend to overestimate how thirsty other people are,<sup>5</sup> and doctors who are not in pain underestimate how much pain their patients’ pain, perhaps under-administering analgesics as a result.<sup>6</sup>

Psychologists refer to this broad class of phenomena, wherein one’s current hot (or cold) states systematically skew one’s behavioral predictions about both oneself and others ‘in the direction’ of those states, *hot-cold empathy gaps*. This label is used to refer inclusively to both the intrapersonal and interpersonal variants and to versions where the subject making a prediction is in a hot state and to versions where she is in a cold state.<sup>7</sup> The effect has been studied most in the context of *visceral states*, which are roughly drive states which serve biological fitness, such as hunger, thirst, fatigue, warmth, sexual arousal, and pain. However, other results show similar effects for other felt states, such as curiosity, social anxiety, attachment to possessions, and social pain.<sup>8</sup> The label ‘empathy gaps’ reflects the common presumption that the gaps result from a failure to first-personally

---

<sup>2</sup> Gilbert, Gill, and Wilson (2002)

<sup>3</sup> Fisher and Rangel (2014).

<sup>4</sup> Van Boven, Loewenstein, Welch, & Dunning (2012).

<sup>5</sup> Van Boven and Loewenstein (2003).

<sup>6</sup> Loewenstein (2005a, 2005b).

<sup>7</sup> In some contexts, ‘hot-cold empathy gaps’ refers exclusively to versions of the phenomena wherein subjects in a hot state mispredict the behavior of subjects not in that state. We aim to make clear by context which of these usages we are employing. See also Van Boven et al. (2013).

<sup>8</sup> Respectively shown in: Loewenstein, Prelec, and Shatto (1998), Van Boven, Loewenstein, Welch, & Dunning (2012), Van Boven, Dunning, Loewenstein (2000), Nordgren, Banas, McDonald (2011).

model the target subject's psychological state.<sup>9</sup> On this presumption, the reason that (say) a hungry person fails to correctly predict how much food a full person would eat is that she fails to imagine 'from the inside' what it is like to feel full, rather than hungry. The predictive error is rooted, on this view, in a failure to accurately or fully render the subjective experience of some 'other', whether that other is oneself in a different scenario or another person altogether.

In this paper, we have three broad aims. The first is to develop a partial, empirically plausible account of hot-cold gaps. Specifically, we will argue that this process involves a subjective or first-personal encoding of one's current visceral states (§1.1) and that hot-cold gaps typically cannot be 'closed' by way of correcting one's simulation (§1.2). Our second aim is to argue that the predictions generated by this process very often do not amount to knowledge, even when they happen to be accurate (§2).

Our third aim is to build on the previous results to draw out some implications for a compelling and influential view of authentic decisions developed and championed by L.A. Paul, on which authentic decisions must, in certain cases, be guided by subjective knowledge of one's values (§3). We will argue that when it comes to many decisions concerning hot or cold states, subjective knowledge of one's values is not psychologically possible. As a result, the requirement that authentic action be guided by subjective knowledge of one's values is triggered in far fewer cases than one might have thought. In light of this result, we briefly motivate an alternative view of the epistemic grounds of authentic action, one on which authentic action should be based on a kind of *understanding*, where understanding can be achieved through subjective knowledge but can also be achieved otherwise, such as through testimony from those who know you best about what your values are.

Before proceeding to our main claims, it might be helpful to say something about the broader significance of hot-cold gaps. First, it might seem, from the cases we've glossed so far, that hot-cold empathy gaps are *generally a bad thing*, in the sense that they tend to stymie practical ends and, in some cases, moral ends. For instance, because of hot-cold gaps, the recovering alcoholic who is in the throes of a craving will underestimate how much she will regret her drinking later, when she is no longer craving alcohol, and this predictive difference might plausibly contribute to her choosing to drink later, with consequences she might sharply regret.<sup>10</sup> On the moral front, hot-cold gaps may play a role

---

<sup>9</sup> See, e.g., Van Boven (2013) for a view on which empathy gaps are due to two judgments.

<sup>10</sup> This suggestion is indirectly supported by Poggiolini (2019). The suggestion here is that hot-cold gaps might play some role in addiction; it is certainly not the claim that these gaps are the sole or primary drivers of addiction. For an interesting take on the contextual factors in addiction, see Pickard & Pearce (2013). For evidence that hot-cold gaps plausibly play a role in sub-optimal negotiation processes, see, e.g., Adler (1995); Van Boven et al. (2000)

in the unjust stigmatization of impulsive behaviors, such as overeating and addiction,<sup>11</sup> in the undertreatment of pain by doctors who are not currently in pain,<sup>12</sup> and perhaps even in the public's acceptance of policies that permit torture.<sup>13</sup>

However, we would resist the sweeping conclusion that empathy gaps are generally a bad thing, whilst acknowledging that empathy gaps sometimes stymie practical and moral ends. For, in some cases, empathy gaps help *facilitate* practical ends and even, potentially, moral ones. For example, someone's excitement in the moment might help them start a future project they would avoid were they to fully appreciate how bored it would make them by the end. Likewise, the addict who is not currently craving might be more willing to check themselves into rehab than they would be if they were to fully appreciate how powerful their craving will become after that decision is made. The decision to enter rehab is plausibly both a practical and (in some cases) a moral decision. Thus, while empathy gaps sometimes stymie practical or moral ends, in other cases, they seemingly facilitate them.<sup>14</sup>

A second point about the significance of hot-cold empathy gaps concerns the connection between these gaps and empathy deficits more generally. On the psychological view of hot-cold gaps we will defend, hot-cold gaps emerge when one 'projects' one's own hot or cold state into a first-personal rendering of someone who is in a different state. While this kind of projection *might* undergird empathy deficits in general, it is no part of our claim that they necessarily do. Rather, we are more inclined to see the process of projection at play in hot-cold gaps as merely one of a 'grab bag' of methods by which we 'read minds,' that is, attribute mental states of behavioral intentions to others. Other psychological strategies of mindreading plausibly include background theories, perceptual processes, emotional resonances, and heuristics.<sup>15</sup> In defending the view that hot-cold gaps are explained by a process of first-personally projecting one's own state into a model, we are neutral on whether this kind of projection also figures in (for instance) more complex

---

<sup>11</sup> Nordgren, Pligt, & van Harreveld (2007).

<sup>12</sup> Loewenstein (2005b).

<sup>13</sup> Nordgren, McDonnell, and Loewenstein (2011).

<sup>14</sup> For further discussion of the ways that empathy in general can be morally problematic, see for instance: Kate Manne's discussion of excessive empathy for men, what she terms "himpathy," Sukaina Hirji's discussion of the ways in which empathy for an abuser can make it difficult for the person abused to maintain a proper sense of herself, and Olivia Bailey's discussion of the ways in which empathy can warp testimonial trust in pernicious way (Manne (2017), Hirji (forthcoming). Bailey (2018). For an argument that the question is a fraught one of whether, all things considered, one should empathize with vicious perspectives, see Bailey (2021). For an argument that empathy in general can compromise the empathizer's authenticity, despite its moral and epistemic benefits, see Paul (2021).

<sup>15</sup> For helpful recent discussions of some of these myriad strategies in mindreading, see Spaulding (2018, 2020).

forms of empathizing, such as understanding a loved one's puzzling behavior in the wake of their having suffered a traumatic event or appreciating the ideological motivations of someone opposite you on the political spectrum.

### 1. Hot-Cold Empathy Gaps and the First-Personal Perspective

In this section, we will develop and briefly defend a partial psychological account of hot-cold empathy gaps.<sup>16</sup> On the partial view we will develop, subjects employ their own current visceral states in order to simulate some situation 'from the inside,' i.e., first-personally, and they thereby arrive at a prediction about the behavior or preferences of either some other person or else of themselves at a different time.<sup>17</sup> Moreover, subjects by default treat their current states as inputs to this prediction. Roughly, the prediction occurs when one 'projects' one's own current visceral state into some kind of subjective rendering of a scenario. This hypothesis we dub the *first-personal projection view*.

This account is neutral on further questions about how empathy gaps occur. In particular, it is neutral on the questions of whether the relevant first-personal process of prediction is inoculated from broader background theories, and whether it is encoded imagistically rather than propositionally, graphically, or otherwise.

Notably, we will often refer to the predictive process which underlies hot-cold gaps as one in which a subject *simulates* her experience, by which we mean she somehow models it 'from the inside' or with respect to a subjective 'I.' It does not matter for our purposes whether this process is achieved imagistically, but the process will tend to involve some 'felt' quality insofar as it involves a model constructed around one's current visceral state. To illustrate the view with an example: When Calvin makes a prediction about how much food he will want when very hungry, he does so partly on the basis of his current feelings of hunger when imagining that prospect, and not (merely) on the basis of third-personal information about what someone in his state would likely want.

---

<sup>16</sup> Elsewhere, we fill out this psychological view in more detail (Helton & Register, unpublished manuscript).

<sup>17</sup> In contrast to how we will use the term, simulations are not always assumed to be first-personal. Some approaches to understanding intuitive theories posit that subjects use intuitive theories to run simulations, akin to how a computer simulation runs on a physics engine (Gerstenberg & Tenenbaum 2017). That sort of simulation may not be accompanied by imagery presented in the first-personal mode. Moulton and Kosslyn (2009) argue that all mental imagery is part of some simulation process and that many kinds of simulation employ mental imagery. The view is similar to the simulation view in the broader debate about mindreading. However, unlike most species of the simulation view, we are neutral for present purposes on whether the first-personal predictive process proceeds by (perhaps implicit) theorizing and whether it is guided by mental imagery. For a helpful recent overview of this issue, see Barlassina & Gordon (2017).

### 1.1 Evidence For the ‘First-Personal Projection’ View

First, here is why we take hot-cold gaps to arise because of some form of first-personal process, as opposed to a process of third-personal prediction. Multiple studies have shown that, when asked about what someone else might be thinking or feeling, participants first-personally imagine the circumstances that the other person is in.<sup>18</sup> There is also evidence that asking participants to imagine the feelings of someone else increases the activation of the participants’ own self-focused thoughts, relative to a control.<sup>19</sup> We take these studies to show that people tend, by default, to represent others’ situations ‘from the inside’ when making predictions about them. Additionally, other results suggest that people first-personally project *at least* as frequently when making predictions about their own experiences as they do when making predictions about others’ experiences.<sup>20</sup>

Second, here is why we take empathy gaps to arise because of a process of projection, whereby subjects by default treat their own visceral states as inputs to a prediction. This ‘projection’ view can accommodate the fact that hot-cold predictions err in systematic ways. In particular, they tend to ‘skew’ in the direction of the predictor’s current visceral states. At the same time, there is evidence that subjects’ broader theories about how their current states map onto more general states are poorly equipped to explain empathy gap results.<sup>21</sup>

### 1.2 The Limits of First-Personal Projection: When Simulating Others’ States Is Not Possible

So far, we’ve suggested that hot-cold gaps emerge by way of a first-personal process or projection, wherein subjects treat their current visceral states as default inputs into a simulation. In some cases, these predictions are inaccurate. We now turn to evidence that subjects typically cannot ‘correct’ their simulations so as to avoid these erroneous predictions. This isn’t to suggest that there is *no* way by which they might correct the predictions, but the possibility of doing so is at least seriously constricted.

On its face, one might think that predictive gaps emerging from a process of simulation *should* be surmountable by a relevant process of counter-simulation; after all, one might think subjects can simply import *a different hot or cold state than their own* into

---

<sup>18</sup> Berntsen and Jacobsen (2008), Depow et al. (2021), Van Boven and Loewenstein (2003), Van Boven et al. (2013)

<sup>19</sup> Davis et al. (2004).

<sup>20</sup> Pronin and Ross (2006), Gilbert and Wilson (2007), Van Boven et al. (2013).

<sup>21</sup> Steinmetz (2018).

the simulation to yield a better prediction. But several results tell against this suggestion. In particular, for especially intense states and for states of certain kinds (e.g. pain), it is not obviously possible for subjects to simulate visceral states that are incongruent with their current actual states. For instance, it is not obviously possible to fully simulate intense pain when one is currently not in any pain at all.

A first result which suggests that subjects cannot use simulation to ‘correct’ their erroneous hot-cold predictions comes from Nordgren et al. (2006). In this study, fatigued participants were asked to assess some non-fatigued person’s behavior by pretending that they were also non-fatigued; yet, this instruction did not alter the predictive gap. The same result held of non-fatigued participants asked to assess a fatigued individual by pretending they were fatigued. These results led the study’s authors to conclude that “... efforts designed to help people overcome empathy gaps are likely to be unsuccessful.” (Nordgren, p. 638).<sup>22</sup>

A second set of results which suggests that subjects struggle to correct their simulation-based predictions comes from results concerning pain, especially extreme pain. Some evidence suggests that those who have undergone extreme pain but are not currently experiencing such pain—such as those who have previously given birth—struggle to accurately predict the behaviors of those currently in extreme pain. Barbara Montero (2020) argues at length on the basis of this and other evidence that the ‘felt’ component of pain is in principle inaccessible in memory.<sup>23</sup> If Montero is right, and if simulation of hot-cold gaps related to pain would require the reconjuring of pain on the basis of phenomenal pain memories, then gaps related to extreme pain cannot be bridged, as Montero herself also argues.<sup>24</sup>

Why would it be impossible for subjects to fully close intrapersonal empathy gaps, e.g., why would it be impossible for someone who is in the throes of an intense nicotine craving to simulate a feeling of not craving at all? While we will not fully endorse this hypothesis here, we are sympathetic to a suggestion from Loewenstein (1999) that intense visceral states significantly control a subject’s attention and motivation, with the result that

---

<sup>22</sup> Results from Loewenstein (1998) further shows that subjects cannot correct empathy gaps even when motivated to do so.

<sup>23</sup> See Christensen-Szalanski (1984) and Morley (1993) for examples of corroborating evidence.

<sup>24</sup> Montero (2020: 119-122). See also Read and Loewenstein (1999). An important challenge for our view is from Steinmetz et al. (2018), which might suggest that subjects can counter-simulate in some cases. We think that closer inspection shows that these results do not suggest that this is possible, especially for intense states. For one thing, the Steinmetz et al. (2018) result involved small effect sizes which were not measured against subjects’ baseline states, which opens up other interpretations of the results. We discuss this and other considerations in more detail in Helton & Register (unpublished).

they “crowd out” contrary states, making the simulation of such states either very difficult or psychologically impossible.

We conclude that there are good reasons to think that, in at least some cases, it is not possible for subjects to simulate hot or cold states different than their own, perhaps especially when the relevant states are dramatically different than one’s current state. To say that empathy gaps cannot be overcome by a process of counter-simulation is not, of course, to deny that the *behavioral* consequences of empathy gaps cannot be altered, for instance, by an agent’s choosing to abide by certain time-tested heuristics or rules. Consider: one who finds herself very hungry while at the grocery store, doing her food shopping for the week, might choose to stick to a pre-made shopping list, perhaps one made by looking at her average food consumption in the past, rather than relying on her current visceral state to make predictions about how much food she will need for the week. This choice might well prevent this agent from buying more food than she needs, but it *doesn’t alter her first-personal predictive process*; it merely inoculates her food decisions from that process, by treating an alternative source of evidence, one rooted in third-personal evidence about her past behavior, as a superior basis for action.

## 2. Hot-Cold Empathy Gaps and Knowledge

We turn, in this section and the next, to considering some of the philosophical implications of hot-cold empathy gaps. In this section, we will argue that the predictions which figure in hot-cold gaps do not typically amount to knowledge, even when they happen to be accurate. Ultimately, in the next section, we will draw on this result to sketch a challenge for L.A. Paul’s view that authentic action requires subjective knowledge of one’s values and how those values map onto relevant outcomes.

In some cases, the predictions generated by projecting one’s own hot or cold states into a simulation will not generate true beliefs; indeed, the psychological literature focuses on these cases. For instance, if Nadia is not currently craving a cigarette, she might underestimate the extent to which later, when she is craving one, it will be hard for her to resist lighting up, even if she has a long-term goal of quitting smoking. Or, if Nadia is exhausted from work, she might underestimate how much her well-rested partner will want to go out for dinner later.

However, in some cases, the process which underpins hot-cold gaps will result in accurate predictions. This will happen whenever the target of the prediction happens to be in the same visceral state as the predictor. For instance, suppose Claire is feeling energetic, i.e., extremely non-fatigued, when deciding whether to sign up for a mountain trek later, while on vacation. As a result of importing her current energetic state into a simulation of going on the trek, she concludes she should go on the trek. Further suppose that, as it turns out, Claire *will* happen to be feeling energetic during the scheduled trek,



so she will enjoy the excursion. In this case, the strategy of importing one's current hot state into a simulation of a future event yields a correct verdict.

So, the process which underpins hot-cold empathy gaps is sometimes inaccurate and sometimes accurate. But, we can further ask, when the process of projecting one's own states into a simulation generates accurate predictions, do those accurate predictions tend to amount to knowledge?

Here are two reasons to doubt that predictions generated from projection in a simulation typically constitute knowledge, even when those predictions are accurate. First, there is something inherently odd about the process, in the sense that, there is no obvious reason why one's current state should be a guide to an unknown other's states or to one's own states at a different time. Consider that the states in question—hunger, thirst, pain, fatigue, sexual arousal, and the like—tend to be highly variable across persons and across the same person at different times. So, there is something odd about the internal 'logic' of the process. Even if it works, it's at least not obvious why we would expect it to.

Second, the process itself will not typically be reliable, despite the fact that it sometimes yields accurate predictions. On some views, knowledge can *only* be generated by a reliable method. But even those who reject this view might well grant that reliability of some process is a *defeasible indicator* of whether that process generates knowledge.<sup>25</sup>

To illustrate that the process which underlies hot-cold prediction is often unreliable, consider in particular *safety*, where some process is safe just in case: in most or all near worlds where you form that belief on the basis of that method, that belief is true (Williamson 2000, Pritchard 2009, Sosa 1999). In contemporary work on the connection between knowledge and reliability, safety is the most often-discussed variant of reliability. So, we will take it that if the method of first-personal simulation is not safe, this is a good if defeasible reason to think that this method does not satisfy whatever kind of reliability (if any) which is either a condition on knowledge or a defeasible indicator of knowledge.

To see that the process of first-personal simulation is not safe, consider again Claire's true belief that she will enjoy the mountain excursion during her future vacation. Claire's judgment is formed because she imports her current energetic state into a simulation about what it would be like for her to undertake the excursion, and she exploits that simulation to derive the judgment that she should go on the excursion. Does this method produce true beliefs in most or all near worlds, as safety requires? Assuming Claire's states of energy and fatigue ebb and flow much like the rest of ours, based on a complex of biological and situational factors, the answer must be 'no.' For, in some near worlds, Claire is exhausted when she considers whether she should take the trip even

---

<sup>25</sup> One of the author's is a theorist of this stripe, having criticized reliability as a necessary condition on knowledge (Helton and Nanay, 2019) but nevertheless regarding reliability to be a good if defeasible proxy for whether some process is knowledge-conferring.

though—in that world—she would happen to be in a well-rested state before the excursion itself. In other near worlds, Claire is in an energetic state when she considers whether she should take the trip even though she will be worn out from travel and unexpected work obligations just prior to her trip.

We claim that the process of projection is typically unreliable, but we do not deny that it is ever reliable. Certainly, this process might be reliable for some individuals and some kinds of visceral states. For instance, two people who tend to eat together might tend to have similar cycles of hunger and satiation and thus, their first-personal-based predictions about the others' hunger might tend to be reliable. But in any context where such visceral states are not synched, the process of projection will not be reliable, and these contexts are extremely common. So, many such predictions will be unreliable in the way we might expect knowledge to be.

We conclude that the process which underpins hot-cold empathy gaps is (at least often) not safe. We take this to be a good if defeasible reason to doubt that this method is generally knowledge-conferring, even when its predictions happen to be accurate. Certainly, there is more to say on this matter. For one thing, there are other ways a process might be knowledge-conferring which are neither to do with its reliability nor its internal logic.<sup>26</sup> Going forward, we will help ourselves to the assumption that first-personal simulation across hot-cold gaps is at least often not knowledge-conferring, for the sake of drawing out what, if anything, this result tells us about the scope of an important notion of authenticity developed and defended by L.A. Paul.

Notably, to say that the hot-cold predictive process *itself* often does not generate knowledge is not to suppose that it cannot help yield knowledge in combination with other strategies. In particular, perhaps someone who is aware of the empirical reality of hot-cold gaps could exploit this theoretical knowledge to render her own predictions more reliable.<sup>27</sup> For instance, consider again Claire, who is deciding whether to book a mountain excursion on her next trip. If she also knows about hot-cold gaps, she might strategically decide to make this decision when she's very tired, as she anticipates that, after traveling to her destination, she will be tired. Deliberately waiting until she's in the 'right' state might help her make a good choice. We do not deny that this process, one which partly exploits first-personal prediction, might be knowledge-conferring. But, we note that this process requires more than first-personal prediction and will be difficult to implement in many

---

<sup>26</sup> For instance, the process might have a kind of *presentational phenomenology* which helps render it potentially knowledge-producing. See, e.g., Bengson (2015) and Chudnoff (2012). Or it might disclose truth-makers in a way that renders it especially epistemically valuable Johnston (2006, 2011). Or it might confer certain relevant capacities or competences, see, e.g., Schellenberg (2018) and Miracchi (2015).

<sup>27</sup> We thank Olivia Bailey for this suggestion and for helpful discussion on this point.

cases. In any case, we do not take it to threaten our claim that very many instances of first-personal projection will not be knowledge-conferring.

### 3. Hot-Cold Empathy Gaps, Authenticity, and the Limits of Subjective Knowledge

Intuitively, some of our actions reflect our values and some do not, either because those actions are neutral with respect to our values or else because they fly in the face of our values. For instance, if Calvin values honesty and, guided by this value, discloses something embarrassing but important about his past to his new romantic partner, his disclosure reflects his deep values in at least some way. Or, if Nadia values being a reliable friend but, due to an ongoing dependence on alcohol, struggles to keep promises to her friends, Nadia's actions do not reflect her deep values in at least some way.

While acting in a way that reflects one's values is not the only standard by which actions might be assessed, we take it to be an important one and one that many of us care about.<sup>28</sup> Following L.A. Paul, we will dub actions which reflect one's values *authentic* ones and those which reflect what one disvalues *inauthentic* ones. In making this terminological point, we do not mean to suggest that other views of 'authenticity' are inapt. What we care about is value-reflectance, and the choice of terminology is arbitrary. Those who prefer to employ the term 'authenticity' to pick out some different trait can freely substitute in some other term, such as 'value reflectance,' without any loss in the argument.

What is required for an action to be authentic in the sense of reflecting one's values? More particularly, what epistemic situation must one be in to carry out authentic action? Paul has argued that in at least some cases, authentic action involves knowledge obtained from first-personal experience. These experiences permit a kind of grasp of one's subjective values, a way of understanding them which is importantly different than other ways one might represent one's values, such as purely cognitively. The kind of grasp Paul is concerned with is an experiential kind, and Paul uses the 'Mary' thought experiment to draw the distinction: Before leaving the black-and-white room, Mary has a purely cognitive understanding of color experience, but after leaving it and seeing color for the first time, she attains an experience-based grasp of color experience, an understanding that is somehow distinct from her purely cognitive understanding of color vision, even if the facts which are understood are the same.

These remarks from Paul are representative of her view of authenticity and subjective knowledge of one's values:

---

<sup>28</sup> In addition to assessing an action along the familiar lines, such as with respect to its morality, its aesthetic qualities, or its practical rationality, we can also assess it on any other number of grounds, such as whether it exhibits spontaneous freedom (Gingerich 2022) or whether it exhibits shared improvisational agency (Bagley 2013, 2015). We pursue a thoroughgoing pluralism about the standards along which action might be assessed. Thanks to Sophie Dandélet for discussion on this point.

Authentic decision-making can require imaginative knowledge of what my future circumstances will be, where such imaginative knowledge carries with it a direct affective, emotional engagement that allows me to cognitively and emotionally empathize with my possible future selves.<sup>29</sup>

Notice two things about this passage. First, Paul suggests (by pragmatics) that authentic choice needn't require 'imaginative knowledge,' though it 'can require' it. Second, on Paul's conception, it is *knowledge* produced by imagination that is, in at least some cases, required for authentic decision-making.<sup>30</sup> As we are interpreting Paul, on the basis of this and other passages, her total view of the relation between authenticity and imagination is this:

AUTHENTICITY FROM SUBJECTIVE KNOWLEDGE

All else being equal, authentic action should be based on subjective knowledge of one's values.

For short, we sometimes refer to this view of Paul's as *the subjective knowledge view*. We will say more shortly about Paul's view of what else must be equal for this requirement to be triggered. For now, it is important that while Paul often speaks of subjective knowledge being achieved imaginatively, via a first-personal process of engagement, her fuller view is rather that this subjective knowledge can include knowledge of memories and also cognitive encodings of information, so long as this information is first-personally indexed and derived from some sampling of one's total experience during some duration.<sup>31</sup>

Paul motivates the subjective knowledge view of authentic decision-making in several ways, just some of which are these: First, she draws on certain compelling thought experiments to motivate the view. For instance, she suggests that were a person to decide whether to have a child by relying solely on the testimony of others about what parenthood is like, this would be odd and intuitively inauthentic.<sup>32</sup> Second, she suggests that the kind of grasp conferred by subjective knowledge of one's values permits a 'sense of control...' in one's choices.<sup>33</sup> Third, Paul suggests that authentic preferences should be formed on the basis of subjective values, where these values are partly constituted by, but not exhausted

---

<sup>29</sup> Paul (2015b: 810).

<sup>30</sup> On the first point, see, e.g., Paul (2020). For other remarks by Paul on the connection between knowledge and authentic experience, see, e.g., (Paul 2015b: 808, 810). For a metaphysics of the phenomenal feel which figures in simulation and experience, see Paul (2017b).

<sup>31</sup> Paul (2014: 106).

<sup>32</sup> For related points, see, e.g., Paul (2014: 75) and Paul (2015b: 809, 811-813).

<sup>33</sup> Paul (2014: 107).

by, aspects of experience.<sup>34</sup> Together, these claims characterize an intuitive ideal of knowledgeably navigating the world guided by one's own first-personal perspective.

We find these motivations to be extremely compelling ones, such that we think the thesis of authenticity from subjective knowledge merits serious consideration.<sup>35</sup> At the same time, we will suggest that considerations from hot-cold gaps show that the requirement that authentic action be governed by subjective knowledge is triggered in far fewer circumstances than one might have thought. We will also suggest, in a more polemical vein, that this result in turn should motivate us to re-evaluate Paul's view of the epistemic grounds of authenticity, with an eye to whether some broader theory might subsume this one.

To mount this argument, we will draw on three claims previously defended: First, hot-cold empathy gaps derive from a first-personal predictive process. Second, this process does not confer knowledge, even when the judgments it produces are accurate. Third, in at least some cases, subjects cannot 'correct' this process via a process of counter-simulation. As a result, in at least some cases, subjects cannot make choices concerning hot or cold states that are based on subjective knowledge, even where experience is broadly construed to include the sum of one's present perceptual, emotional, cognitive, and other processes. For, as a psychological matter, subjects often simply cannot simulate in a way that would produce this sort of knowledge.

For the sake of keeping things concrete, consider how these claims apply in specific cases. Recall Nadia, the smoker who is trying to quit. With respect to Nadia, our claims are these: first, when Nadia is in this non-craving state, she is likely to rely on this state in making a prediction about her later preferences. Second: this judgment does not constitute knowledge, whether or not it happens to be correct in this particular instance. Third, while Nadia is in this non-craving state, it is not psychologically possible for her to simulate otherwise, i.e., to simulate what it would be like for her to (say) walk past her cigarettes when she is craving a cigarette. So, Nadia's choice to (say) leave her cigarettes where they are, lying on the coffee table, is not based on knowledge from simulation.

The same claims apply to Claire, who, while in an energetic state, considers whether to book a mountain excursion for a later vacation. Because Claire's assessment of the trek is influenced by her energetic state, her choice is not based on subjective

---

<sup>34</sup> Paul (2014, 2015b, 2017a). See also Paul (2015b: "Reply to Campbell") for the view that these subjective values are partly individuated by worldly features.

<sup>35</sup> We are sympathetic to the thought—helpfully raised to us by Jane Friedman—that it isn't knowledge at all which ought to ground authentic action, but is rather some other kind of state altogether, such as belief, justification, or something else. The rival view we will sketch appeals to understanding, and we are neutral on the controversy over whether understanding is a form of knowledge, so for present purposes, we are neutral on the question of whether the epistemic grounds of authenticity ought to be knowledge or something else altogether.

knowledge of her values. This is because the kind of simulation she employs isn't knowledge-conferring, even if accurate. Moreover, it is not psychologically possible for Claire to make this choice in a way that is based on knowledge from simulation. She cannot 'correct' the first-personal predictive process to make it a knowledge-conferring process because, while in a highly energetic state, she cannot simulate the state of being fatigued.

Consider what these results mean for Paul's view of authenticity, on which, *ceteris paribus*, authenticity requires subjective knowledge of one's values. On this view, either Nadia's decision concerning her cigarettes is not the kind of choice that can *be* authentic or inauthentic, or else, that this is not the kind of decision which triggers the requirement of subjective knowledge. Likewise, this view entails either that Claire's decision about her mountain trek is not the kind of choice that can *be* authentic or inauthentic, or else, that this is not the kind of decision which triggers the requirement of knowledge from simulation.

We will argue this way: Both Nadia's decision about whether to leave the cigarettes out and Claire's decision about whether to make the trek are the kinds of choices which *can* be authentic ones. This isn't to say that such choices typically *will* be authentic, merely that it is humanly possible that they be so. What Paul's subjective knowledge view leaves us with, then, is the view that these choices are, for whatever reason, not the kinds of choices which trigger the requirement of subjective knowledge. While we think this result can be defended on non-*ad hoc* grounds, ones Paul herself develops, we will ultimately suggest that this result raises questions for a different reason, namely that it reveals the subjective knowledge requirement to be of very limited scope, since it is rarely triggered.

First, to defend the thought that choices about leaving cigarettes out or booking a mountain excursion can be authentic ones. For, one might reasonably wonder whether these micro-decisions can reflect on one's deep values in whatever way authenticity requires. On this view, these decisions are not assessable with respect to authenticity, perhaps because they are minute.

Certainly, these are not the kinds of choices which tend to animate discussions about authenticity. More commonly, this literature focuses on dramatic, one-off decisions, such as whether one should move to another country or whether one should have a child. Indeed, many of the common examples involve what Paul dubs *transformative experiences*, where these are experiences which alter oneself in a deep way and which are such that one cannot truly know what they are like until one has had them. The decision about whether to leave the cigarettes out or whether to go on a mountain hike are extremely unlikely to deeply change oneself in the way transformative experiences do.<sup>36</sup>

There are two things to say about the concern that the kinds of prosaic choices which often figure in hot-cold decisions are not significant enough to be evaluable with respect to authenticity. First, while some hot-cold decisions are not likely to be life-

---

<sup>36</sup> Paul (2014).

changing, others are potentially transformative. For instance, the choice about whether to enter rehab to treat one's addiction is potentially a choice made at least partly on the basis of hot-cold simulations, insofar as it is much easier to decide to enter rehab when not currently craving a substance. This choice is potentially transformative, since it might alter a person's deep self and might involve experiences one cannot fully grasp in advance.

Second, we would resist the suggestion that even prosaic, non-transformative hot-cold decisions are not sufficiently tied up with one's deep values to qualify as either authentic or inauthentic. For arguably, an authentic *life* is no less made up of the countless micro-decisions one makes a day than of those large decisions which keep us up at night. These small decisions are on their own of little consequence, but together they make up much of what a life is: what kind of diet you and your loved ones have, whether you travel, whether you smoke, whether you make plans with friends—with all of the attendant sequelae these choices have for susceptibility to depression and other illness, the quality of one's social relationships, the nature of one's accomplishments, and on. When it comes to an authentic life, these micro-decisions are not minutiae; they help make up the fabric of an authentic life, if not the whole of it.

Our interim conclusion then, is this: Decisions made at least partly on the basis of hot or cold states *are assessable with respect to authenticity*. Given this result, Paul's view of authenticity forces us to the view that these choices are not the kinds of choices which trigger the requirement of subjective knowledge.

At this point, Paul might reasonably object to our suggestion that hot-cold decisions cannot be based on subjective knowledge. While Paul's view is that knowledge obtained by a sampling of one's experience is sometimes required for authentic choice, recall that this sampling involves *all* subjective aspects of one's experience; it isn't restricted to one's visceral states. It can extend, for instance, to subjectively encoded memories or other first-personally presented forms of knowledge. So, Paul might maintain that, even if the hot-cold predictive process alone cannot generate knowledge of another person or one's future self, other elements of one's psychology can serve this role, even when it comes to hot-cold decision-making. For instance, Nadia might remember that whenever she has left out her cigarettes in the past, she has tended to light up later. So, her memory might serve as a form of subjective knowledge which can ground her authentic choice not to leave her cigarettes out, even if the hot-cold predictive process itself cannot play this role.<sup>37</sup>

We acknowledge that the epistemic limitations of hot-cold gaps do not, in their own right, create a difficulty for Paul's subjective knowledge view, precisely because Paul's view encompasses forms of knowledge which go beyond simulation from visceral states. At the same time, we think there are additional reasons to think that in many hot-cold cases, these additional forms of knowledge are either not subjective or else are inadequate

---

<sup>37</sup> We thank Laurie Paul for helpful discussion on this point.

to ground knowledge about the relevant choice. As a result, in at least some such cases (and perhaps most), the subject will lack subjective knowledge which might ground her decision.

For instance, consider the suggestion that memories which might inform one's hot-cold decision, such as Nadia's memory that whenever she leaves the cigarettes out, she tends to smoke later. If this memory is encoded third-personally—for instance, if Nadia merely remembers *that* she does this, without accessing the values which explained her behavior in a subjective way—then the knowledge is not subjective in the relevant way.

If, on the other hand, Nadia's memory encodes in a 'felt,' first-personal way the values which explained why she smoked in those cases, e.g., if she remembers the intense craving which accompanied that action, then that knowledge is subjective and is potentially the sort of knowledge which might ground an authentic decision in the way Paul's view requires. But, we think it empirically implausible that, whilst in a non-craving state, Nadia will be able to access this memory of the florid, first-personal feeling of craving. For, if the typical subject in this case were able to access this feeling of craving, it would be puzzling why she would tend to act in accordance with her current non-craving visceral state. And the literature on hot-cold gaps suggests precisely that she will tend to act in accordance with her current visceral state. This isn't to suggest that it is impossible that she access this memory or that there is no variance across subjects or cases, just that there are independent grounds for thinking that this is typically not the case.<sup>38</sup>

We conclude that choices involving hot-cold states are often made on the basis of simulations which are not themselves knowledge-producing. We also think that in at least many (and perhaps most) such cases, subjects lack access to other subjective grounds for making the choice. Together, we take these claims to force Paul to the view that either: the requirement that authentic action be grounded in subjective knowledge is often not triggered, as in many hot-cold decisions, or: these decisions are systemically inauthentic, as a matter of something like psychological necessity. We take the latter result to be highly counterintuitive and will focus the remaining discussion on the first possibility.

Might the requirement that authentic action be grounded in subjective knowledge be rarely triggered? Paul's view is at least in principle consistent with this result. She considers cases where subjective knowledge is not available and explicitly argues that in such cases, authentic action can be grounded in third-personal knowledge, such as testimony from experts.<sup>39</sup>

Thus, our complaint with Paul's view is not that hot-cold decisions constitute a counterexample to it; by Paul's own lights, the requirement that authentic decisions be based on subjective knowledge is not always triggered. But, we will argue that hot-cold decisions nevertheless raise a *puzzle* about Paul's view, a puzzle deriving from the ubiquity

---

<sup>38</sup> See Montero (2020) for an argument that these memories are deeply inaccessible.

<sup>39</sup> Paul (2015b: 811)



of hot-cold decisions in human experience: If a great many decisions can be authentic which do not trigger the requirement of subjective knowledge, should we be confident that it is subjective knowledge *per se* which bears the relevant connection to authentic knowledge? Or rather, and despite the suasive motivations for the subjective knowledge view, might the epistemic grounds of authenticity be something other than subjective knowledge? More specifically, might there be some more general epistemic condition on authenticity which both explains why subjective knowledge is required in certain contexts and also explains why it is not needed in a great many others?

Several possible answers present themselves. We will briefly sketch just one of these, what we dub *the understanding view of authenticity*, which we present for the sake of provoking inquiry, not as a matter of endorsement. On this view, authentic decisions should be informed by a kind of *understanding* of one's values and how those values fit with possible outcomes. At a minimum, the relevant form of understanding will involve a kind of appreciation of the explanatory relations between one's values and relevant outcomes.<sup>40</sup> This form of understanding can be achieved by first-personal, subjective forms of knowledge but can also be realized in other ways. For instance, one might come to appreciate one's values and how they connect to relevant outcomes via third-personal means, as when a close friend helps one better appreciate one's values by pointing out patterns in one's past behavior and what values might undergird them.<sup>41</sup> So, the understanding requirement is weaker than the subjective knowledge requirement: While subjective knowledge is a way of realizing the relevant form of understanding, understanding can be achieved without subjective knowledge.

While the understanding view is weaker than Paul's subjective knowledge view, the understanding view is more demanding than a natural view inspired by expected utility theory. On that view, authentic choice only requires knowing the utility values of the outcomes of a choice, such as on the approach taken by Pettigrew (2015). However, knowledge of utility values in a choice scenario does not include knowledge of how one's

---

<sup>40</sup> Some theorists treat understanding as a species of knowledge, whereas others treat knowledge and understanding as disjoint kinds. In construing understanding in terms of appreciating explanatory relations, we are meaning to be neutral on the question of whether the relevant form of appreciation is knowledge, grasp, or something else. We are also neutral on the question of whether understanding is a kind of cognitive map (Gopnik, Glymour, & Sobel 2002; Gopnik et al. 2004; Grimm 2016), a skill (Hills 2016; Khalifa 2017, Cf. Sullivan 2018), or a propositional attitude. For helpful overviews on recent work on understanding, see: Baumberger, Breisbart, and Brun (2017), Grimm (2021), and Hannon (2021). See McSweeney (forthcoming) for a view of understanding on which it involves a kind of phenomenal experience and a belief. In contrast to this usage, we are here employing 'understanding' in a factive way.

<sup>41</sup> Or, this process might be deeply dynamic, as in the kind of reciprocal form of conversation described by Dover (2022).

deeper values are linked to the relevant outcomes; so it does not satisfy the understanding requirement of authenticity.

Notice that the understanding view provides a powerful explanation of why authentic action sometimes requires subjective knowledge but does not always require it. In some cases, subjective knowledge will be the only form of understanding available to the agent, and in such cases, subjective knowledge will be required for authentic action. At the same time, the understanding view is inconsistent with the thought that subjective knowledge, where available, is a superior grounds of authentic action over, say, testimony from a close friend. On the view we are proposing, there is a kind of egalitarianism between subjective knowledge and other forms of understanding. All forms of understanding are on a par when it comes to grounding authentic action. This is so even though different forms of understanding differ in other respects, such as in the psychological mechanisms which produced them.

One might object to the understanding view on the grounds that it flies in the face of the powerful intuition that subjective forms of knowledge are in some sense better than other epistemic grounds of authentic action. Paul presses this point in many contexts, often returning to the example of Mary in her black-and-white room. Before leaving the room, Mary the vision neuroscientist might understand color experience in *some* sense, but the phenomenal grasp she gains upon leaving the room and seeing red for the first time is intuitively better, both on epistemic grounds and for some action-guidance purposes, than anything she might have learnt from a textbook.

We take seriously the suggestion that the kind of understanding which is generated by first-personal experience is intuitively superior to other forms of understanding one might have, for instance, understanding derived from a close friend's testimony about one. At the same time, we think it is worth taking seriously that this intuition might be incorrect. In a conjectural mode, we suggest how this intuition might be debunked: Perhaps there is something 'seductive' about subjective forms of knowledge, something which makes subjects trust them even when they are not trustworthy. If this is so, then perhaps in making judgments about the value of subjective knowledge, we ourselves are hostage to this seductive quality.

Along these lines, there is evidence that subjects treat simulations in particular as good sources of information, even in contexts where they are not reliable.<sup>42</sup> For instance, one study showed that participants' memories were superior as a basis for certain predictions than their simulations but that subjects nonetheless relied on their simulations in forming their predictions.<sup>43</sup> It turned out that what explained this reliance on simulation was the phenomenal vividness of their simulations as compared to their memories.

---

<sup>42</sup> Chituc et al. (2021); Kappes and Morewedge (2016); Levine et al. (2020).

<sup>43</sup> Levine et al. (2020).

Likewise, subjects routinely make hot-cold predictions on the basis of simulations despite the fact that these simulations are, we have argued, not typically knowledge-producing.

Together these results suggest that perhaps simulations confer a *feeling* of understanding, one which results in subjects' relying on those simulations, whether or not the simulations are reliable. When this feeling of understanding is spurious, for the reason that the simulations are inaccurate and thus cannot confer understanding, simulations generate what C. Thi Nguyen has termed a 'seduction of clarity.' They give their subjects a feeling of understanding which is, in at least some cases, not accompanied by actual understanding. On Nguyen's view, what is 'seductive' about this feeling is that it can prompt subjects to terminate a process of inquiry, potentially prematurely, and to trust the source of the feeling, whether or not the source is an epistemically good one.<sup>44</sup>

Now for the speculative bit, made not as an endorsement but for the sake of provoking further inquiry: Perhaps the fact that first-personal imagination in general tends to produce a strong feeling of understanding, one which is nevertheless at least sometimes not accompanied by actual understanding, explains the powerful intuition that subjective knowledge is superior to non-subjective knowledge. If first-personal imagination produces a 'seduction of clarity' in this sense, perhaps this felt sense *also* drives our own intuition that Mary's understanding of color experience is superior after she leaves the room than when she is still in it. We, being humans, find Mary's experience extremely compelling for the same reason that we find our own experiences trustworthy: they generate a powerful feeling of understanding, one rooted in the vividness of those experiences, whether or not the relevant experience is accurate, reliable, or otherwise epistemically worthy. If this is the case, then the total theory of the epistemic grounds of authenticity needn't accommodate the intuition that where available, subjective knowledge of one's values is superior to other forms of knowledge. For, this intuition might be inaccurate for reasons explicable in terms of established psychological mechanisms.

We present these remarks concerning the epistemic grounds of authenticity as invitations to further inquiry.<sup>45</sup>

## BIBLIOGRAPHY

Bagley, B. (2013). *Improvisational agency* (Doctoral dissertation, The University of North Carolina at Chapel Hill).

---

<sup>44</sup> Nguyen (2021).

<sup>45</sup> For helpful discussions on this paper, we are indebted to: Josh Armstrong, Olivia Bailey, Molly Crockett, Sophie Dandele, Daniela Dover, Ram Neta, Laurie Paul, and participants of the Transformative Experience workshop at Yale University.

- Bagley, B. (2015). Loving someone in particular. *Ethics*, 125(2), 477-507.
- Bailey, O. (2018). Empathy and testimonial trust. *Royal Institute of Philosophy Supplements*, 84, 139-160.
- Bailey, O. (2021). Empathy with vicious perspectives? A puzzle about the moral limits of empathetic imagination. *Synthese*, 199(3), 9621-9647.
- Barlassina, Luca and Robert M. Gordon. (2017). Folk psychology as mental simulation. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/folkpsych-simulation/>
- Baumberger, C., Beisbart, C., & Brun, G. (2017). What is understanding? An overview of recent debates in epistemology and philosophy of science. *Explaining understanding: New perspectives from epistemology and philosophy of science*, 1-34.
- Bengson, J. (2015). The intellectual given. *Mind*, 124(495), 707-760.
- Berntsen, D., Jacobsen, A.S. (2008). Involuntary (spontaneous) mental time travel into the past and future. *Consciousness and Cognition*, 17, pp. 1093-1104.
- Chituc, V., Paul, L., & Crockett, M. (2021). Evaluating Transformative Decisions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43.
- Chudnoff, E. (2012). Presentational phenomenology. *Consciousness and subjectivity*, 51-72.
- Depow, G. J., Francis, Z., & Inzlicht, M. (2021). The Experience of Empathy in Everyday Life. *Psychological Science*, 32(8), 1198-1213.
- Dover, D. (2022). The Conversational Self. *Mind*, 131(521), 193-230.
- Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational behavior and human decision processes*, 88(1), 430-444.
- Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, 317(5843), 1351-1354. <https://doi.org/10.1126/science.1144161>

- Gingerich, J. (2022). Spontaneous Freedom. *Ethics*, 133(1), 38-71.
- Gopnik, A., Glymour, C., & Sobel, D. (2002). Causal maps and Bayes nets: A cognitive and computational account of theory-formation. *The cognitive basis of science*, 117-132.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological review*, 111(1), 3.
- Grimm, S. R. (2016). Understanding and transparency. In *Explaining Understanding* (pp. 228-245). Routledge.
- Grimm, Stephen (2021). Understanding. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/understanding/>
- Helton, G. & Nanay, B. (2019). Amodal completion and knowledge. *Analysis*, 79(3), 415-423.
- Helton, G. & Register, C. (unpublished manuscript). The psychology of hot-cold empathy gaps.
- Johnston, M. (2006). Better than mere knowledge? The function of sensory awareness. *Perceptual experience*, 260-290.
- Johnston, M. (2011). On a neglected epistemic virtue. *Philosophical Issues*, 21, 165-218.
- Fisher, G., & Rangel, A. (2014). Symmetry in cold-to-hot and hot-to-cold valuation gaps. *Psychological science*, 25(1), 120-127.
- Hannon, M. (2021). Recent work in the epistemology of understanding. *American Philosophical Quarterly*, 58(3), 269-290.
- Hills, A. (2016). Understanding why. *Noûs*, 50(4), 661-688.
- Hirji, S. (forthcoming). Outrage and the Bounds of Empathy. *Philosophers' Imprint*.
- Kappes, H. B., and Morewedge, C. K. (2016) Mental Simulation as Substitute for Experience. *Social and Personality Psychology Compass*, 10: 405- 420. doi: 10.1111/spc3.12257.

- Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*. Cambridge University Press.
- Levine, L. J., Lench, H. C., Stark, C. E., Carlson, S. J., Carpenter, Z. K., Perez, K. A., Stark, S. M., & Frithsen, A. (2020). Predicted and remembered emotion: Tomorrow's vividness trumps yesterday's accuracy. *Memory*, 28(1), 128-140.
- Loewenstein, G. (2005a). Projection bias in medical decision making. *Medical Decision Making*, 25(1), 96-105.
- Loewenstein, G. (2005b). Hot-cold empathy gaps in medical decision making. *Health Psychology*, 24, S49-S56.
- Loewenstein, G., & Adler, D. (1995). A bias in the prediction of tastes. *The Economic Journal*, 105, 929-937.
- Loewenstein, G., Prelec, D., & Shatto, C. (1998). Hot/cold intrapersonal empathy gaps and the underprediction of curiosity. Unpublished manuscript, Carnegie Mellon University.
- Manne, K. (2017). *Down girl: The logic of misogyny*. Oxford University Press.
- McSweeney, M. M. (forthcoming). Metaphysics as Essentially Imaginative and Aiming at Understanding. *American Philosophical Quarterly*.
- Miracchi, L. (2015). Competence to know. *Philosophical Studies*, 172(1), 29-56.
- Montero, B. G. (2020). What Experience Doesn't Teach: Pain Amnesia and a New Paradigm for Memory Research. *Journal of Consciousness Studies*, 27(11-12), 102-125.
- Nguyen, C. T. (2021). The seductions of clarity. *Royal Institute of Philosophy Supplements*, 89, 227-255.
- Nordgren, L. F., van Der Pligt, J., & Van Harreveld, F. (2006). Visceral drives in retrospect: Explanations about the inaccessible past. *Psychological Science*, 17(7), 635-640.

- Nordgren, L. F., van der Pligt, J., & van Harreveld, F. (2007). Evaluating Eve: visceral states influence the evaluation of impulsive behavior. *Journal of Personality and Social Psychology*, 93(1), 75.
- Nordgren, L. F., Banas, K., & MacDonald, G. (2011). Empathy gaps for social pain: why people underestimate the pain of social suffering. *Journal of personality and social psychology*, 100(1), 120.
- Nordgren, L., McDonnell, M., & Loewenstein, G. (2011). What constitutes torture? Psychological impediments to an objective evaluation of interrogation tactics. *Psychological Science*, 22, 689–694.
- Paul, L. A. (2014). *Transformative experience*. Oxford University Press: Oxford.
- Paul, L. A. (2015a). Précis of *Transformative Experience*. *Philosophy and Phenomenological Research* 91 (3):760-765.
- Paul, L. A. (2015b). *Transformative Experience: Replies to Pettigrew, Barnes and Campbell*. *Philosophy and Phenomenological Research* 91 (3):794-813.
- Paul, L. A. (2017a). De se preferences and empathy for future selves. *Philosophical Perspectives*, 31(1).
- Paul, L. A. (2017b). Phenomenal Feel as Process. *Philosophical Issues*, 27(1), 204-222.
- Paul, L. A. (2020). The First Time as Tragedy, the Second as Farce. *Journal of Consciousness Studies*, 27(11-12), 145-153.
- Paul, L. A. (2021). The Paradox of Empathy. *Episteme*, 18(3), 347-366.
- Pickard, H., & Pearce, S. (2013). *Addiction in context* (pp. 165-189). Oxford: OUP.
- Read, D. and Loewenstein, G. (1999), Enduring pain for money: decisions based on the perception and memory of pain. *J. Behav. Decis. Making*, 12: 1-17.
- Schellenberg, S. (2018). *The unity of perception: Content, consciousness, evidence*. Oxford University Press.

- Spaulding, S. (2018). Mindreading beyond belief: A more comprehensive conception of how we understand others. *Philosophy Compass*, 13(11), e12526.
- Spaulding, S. (2020). What is mindreading? *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(3), e1523.
- Sullivan, E. (2018). Understanding: not know-how. *Philosophical Studies*, 175(1), 221-240.
- Steinmetz, J., Tausen, B. M., & Risen, J. L. (2018). Mental simulation of visceral states affects preferences and behavior. *Personality and Social Psychology Bulletin*, 44(3), 406-417.
- Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology*, 79, 66-76.
- Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin*, 29, 1159-1168.
- Van Boven, L., Loewenstein, G., & Dunning, D. (2004). Changing places: A theory of empathy gaps in emotional perspective taking. Unpublished manuscript, University of Colorado, Boulder.
- Van Boven, L., Loewenstein, G., Dunning, D., & Nordgren, L. F. (2013). Changing places: A dual judgment model of empathy gaps in emotional perspective taking. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (pp. 117-171).
- Van Boven, L., Loewenstein, G., Welch, N., & Dunning, D. (2012). The illusion of courage in self-predictions: Mispredicting one's own behavior in embarrassing situations. *Journal of Behavioral Decision Making*, 25, 1-12.