

## Research Statement

GRACE HELTON

I am a philosopher working in philosophy of mind, cognitive science, and epistemology, including social epistemology. Some of my work also touches on philosophy of AI, philosophy of film, and social philosophy. Much of my work concerns the psychology, epistemology, and normative significance of our representations of other people.<sup>1</sup>

### 1 The Psychology of the Social Mind

The first strand of my research concerns how we psychologically represent other people. This work focuses especially on how we perceive other people and how we feel about them, as opposed to how we *think* about them. In this way, this work emphasizes modes of representing others that are less ‘intellectual’ than thought.

One result in this strand of work is the view that in at least some cases, we literally visually represent certain of others’ mental states, just as we visually represent shape and color; I argue in particular that we sometimes see certain of others’ intentions, especially those of their intentions which figure in certain actions. For instance, in visually representing a squirrel fleeing a hound, we literally visually represent the squirrel as aiming to evade the hound (whether or not that representation is accurate). This result partially clears ground for the view that perception alone might help confer knowledge of others’ mental states, independently of inference to the best explanation, testimony, or other means of knowing about the inner lives of others (“Visually Perceiving the Intentions of Others,” *The Philosophical Quarterly*, see also “Recent Issues in High-Level Perception,” *Philosophy Compass*).

The result that we sometimes visually represent certain of others’ mental states draws on a proposal I develop about how our perceptual states and the judgments we form on the basis of them can be distinguished. Roughly, I argue that subjects aren’t necessarily psychologically capable of rationally revising their perceptual experiences in ‘real time’ in response to relevant counter-evidence, whereas subjects *are* necessarily psychologically capable of revising their beliefs in ‘real time’ in response to relevant counter-evidence. This capacity is often masked—for instance, by strong emotion—so, the claim that subjects can revise their beliefs in this way is consistent with evidence that humans often have conflicting or irrational beliefs (“If You Can’t Change What You Believe, You Don’t Believe It,” *Noûs*; “Against the Doxastic Theory of Perception,” under review).

A different piece in this strand of research focuses on how we represent our own and others’ subjective points of view in the experience of watching a film. In joint work, Sara Aronowitz and I argue that the subject who views a film sometimes represents a first-personal,

---

<sup>1</sup> An additional strand of my work, one not discussed here, explores ways in which results from philosophy of mind and cognitive science can help inform certain foundational issues in epistemology (“Amodal Completion and Knowledge,” with Bence Nanay, *Analysis*; “Against the Very Idea of a Perceptual Belief,” with Bence Nanay, *Analytic Philosophy*; “Thought Experiments as Tools of Theory Clarification,” *Seemings and Intuitions*, Routledge Studies in Epistemology).

multi-modal, emotionally centered experience—what we term a ‘subjectivity’—that is not indexed to the subject herself, nor to a character in the film, nor, indeed, to *anyone at all*. We argue that this framework locates film experience on a continuum with more familiar forms of cognition, such as mindreading and memory, at the same time that it resolves a longstanding puzzle in the philosophy of film (“Subjectivity in Film: Mine, Yours, and No One’s,” *Ergo*).

In a final work-in-progress in this strand, I consider the nature of *implicit biases*, which are roughly unreported attitudes. Focusing on those implicit attitudes which concern socially significant statuses, such as race, gender, and LGBTQ status, I draw on empirical results to argue that these biases are at least partly made up of relevant emotions (or else dispositions to have relevant emotions). For instance, on this view, anti-LGBTQ implicit bias is partly comprised of disgust (or else a disposition to feel disgust) in response to people who are LGBTQ. I further suggest that this result is important for developing hypotheses about how to ameliorate such biases and for understanding these biases in their broader social context.

## 2 The Social Mind and Issues of Value

The second strand of my work takes up various questions pertaining to the normative significance of our representations of others. The first piece in this strand concerns certain common and pervasive difficulties in predicting what others will want in a situation different than one’s own, what are sometimes called *empathy gaps*. One form of empathy gap concerns a deficit in figuring out what one oneself would want at a different moment. This is a kind of deficit in empathizing with one’s future self. For instance, empirical evidence suggests that subjects who are not currently hungry struggle to accurately predict how much food they will want when they are hungry.<sup>2</sup>

In joint work, Chris Register and I argue that this form of failure to ‘read one’s own mind’ casts doubt on the influential view that authentic choices should be based on a richly experiential grasp of one’s own values. We suggest instead that in some cases, choices might be fully authentic so long as they involve a kind of appreciation of how one’s values tie to relevant outcomes, even if that appreciation is not based on one’s own subjective experience (“Hot-Cold Empathy Gaps and the Grounds of Authenticity,” *Synthese*).

The second piece in this strand concerns the epistemic value of divergent points of view in philosophical discourse. I argue, in contrast to much recent work in metaphilosophy, that no *ideally* philosophically flourishing community will exhibit widespread and diachronically robust agreement on central philosophical questions. Moreover, this is so regardless of whether the community in question is at an earlier or rather a much later, even final, stage of inquiry. Thus, the common presumption that ideally, philosophical communities should ultimately move to rough convergence on persistent philosophical questions is a mistaken one. The argument I develop relies in part on a thought experiment concerning two different philosophical communities, one of which exhibits marked and stable agreement on the central questions and one of which exhibits marked and stable disagreement on those questions (“Viewpoint Convergence as a Philosophical Defect,” volume *Attitude in Philosophy*, OUP).

---

<sup>2</sup> Van Boven et al. (2013).

### 3 The Epistemic, Normative, and Metaphysical Significance of Social Knowledge

The third strand of my research considers, at some level of abstraction, the value of social knowledge. I argue that social knowledge—roughly, knowledge that is about or dependent on other minds—is both harder to vindicate and of broader epistemological, normative, and even metaphysical significance than has previously been appreciated.

First, I argue that social knowledge makes up a surprisingly broad swathe of our worldly knowledge and also that social knowledge cross-cuts intuitively natural domains. Such knowledge includes at least some knowledge of: geopolitical relations, historical facts, cultural traditions, religious practices, and artistic movements. Partly drawing on this result, I further argue that vindicating knowledge of other minds against skeptical worries is at least as epistemologically valuable as vindicating other forms of worldly knowledge against skeptical worries, such as knowledge of tables and quarks (“Epistemological Solipsism as a Route to External World Skepticism,” *Philosophical Perspectives*).

I further argue that social knowledge has unappreciated normative and even metaphysical implications. First, I argue that social knowledge plays a special role in the meaningfulness of our lives, a different role than does knowledge of tables and quarks (“The Simulation Hypothesis, Social Knowledge, and a Meaningful Life.” *Oxford Studies in Philosophy of Mind*). Second, I argue in a work-in-progress that social knowledge interacts in a surprising way with *the simulation hypothesis*, the recently influential claim that we and our environment are realized by a wide-scale computer simulation, one engineered by superintelligent AI. Specifically, if we have social knowledge, this reduces at least somewhat the odds that we are living in a simulation.

Having argued at length that social knowledge is of greater epistemic, normative, and even metaphysical significance than has previously been appreciated, I go on to suggest that social knowledge is much harder to vindicate against skeptical worries than one might have thought. For, at least some approaches which plausibly succeed against skepticism about things like tables and quarks do not vindicate knowledge of other minds (“Epistemological Solipsism as a Route to External World Skepticism,” *Philosophical Perspectives*). I make this point particularly vivid through an extended case study of *structuralism*—a recently developed, independently attractive approach to external world skepticism. I argue that structuralism does not vindicate social knowledge, even if structuralism vindicates other forms of knowledge (“On Being a Lonely Brain-in-A-Vat,” *Analytic Philosophy*). None of this is to suggest that social knowledge cannot be vindicated against skeptical worries; however, it might suggest that the vindication of social knowledge will require a different tact than the vindication of non-social knowledge.<sup>3</sup>

### 4 Next Directions

---

<sup>3</sup> I find the perceptual approach suggested by some of my other work a promising one, but there remain certain difficulties with this approach not mentioned here. These difficulties persist even if we acknowledge that we can (non-factively) perceive others’ mental states.

I am currently developing two new projects. The first of these introduces a new model of emotion, on which emotion is at least partly made up of a *disposition to attend* to relevant matters of value. For instance, on this view, grief is made up, in part, of a tendency to attend to a relevant profound loss. This attentional theory permits an improved understanding of several otherwise puzzling aspects of emotion, such as how it is that powerful emotions can resolve over time even when the situation which triggered the emotion remains the same. The answer is roughly that our *disposition to attend* to the relevant facts of value can diminish over time, even if our ability to experience those facts of value in a visceral way remains.

I build on this view of emotion to draw out some of the epistemic implications of treating conditions such as complicated grief and PTSD. I argue that treating these conditions isn't merely morally or practically permitted, but is also *epistemically* permitted. This is because treating these conditions can and typically does take the form of altering one's disposition to attend to the relevant loss or trauma, *without* thereby obliterating one's ability to emotionally grasp the loss or the trauma. So one needn't worry, for instance, that in taking conventional medication to recover from longstanding, unresolved grief, one is somehow denying the reality of one's loss.

The second project concerns the ethics of stereotypes. I argue that at least some stereotypes are inherently disrespectful in a certain way, even when the stereotype in question involves praiseworthy traits (e.g., 'women are nurturing') and even when the stereotype fits the target (e.g., even when a woman so stereotyped is nurturing). I focus in particular on stereotypes which attribute blameworthy or praiseworthy traits on the basis of non-blameworthy and non-praiseworthy group memberships, such as 'women are nurturing.' Being a woman is neither a blameworthy or praiseworthy trait. It is a neutral trait. But being nurturing is a praiseworthy trait. Stereotyping someone as nurturing simply because they are a woman deprives this person of some of the praise to which they are entitled for being nurturing and is thus, I suggest, degrading.

#### MY WORK CITED

- Aronowitz, S. & Helton, G. (forthcoming). Subjectivity in film: Mine, yours, and no one's. *Ergo*.
- Helton, G. (2016). Recent issues in high-level perception. *Philosophy Compass*, 11(12), 851-862.
- Helton, G. (2018). Visually perceiving the intentions of others. *The Philosophical Quarterly*, 68(271), 243-264.
- Helton, G. & Nanay, B. (2019). Amodal completion and knowledge. *Analysis*, 79(3), 415-423.
- Helton, G. (2020). If you can't change what you believe, you don't believe it. *Noûs*, 54(3), 501-526.
- Helton, G. (2021). Epistemological solipsism as a route to external world skepticism. *Philosophical Perspectives*, 35(1), 229-250.
- Helton, G. & Nanay, B. (2022). Against the very idea of a perceptual belief, *Analytic Philosophy*, 64(20), 93-105.
- Helton, G. (forthcoming). On being a lonely brain-in-a-vat: Structuralism, solipsism, and the threat from external world skepticism. *Analytic Philosophy*.
- Helton, G. (forthcoming). The simulation hypothesis, social knowledge, and a meaningful life. *Oxford Studies in Philosophy of Mind*.

- Helton, G. (forthcoming). Thought experiments as tools of theory clarification. *Seemings and Intuitions*. Routledge Studies in Epistemology. Eds. McCain and Stapleford.
- Helton, G. (forthcoming). Viewpoint convergence as a philosophical defect. *Attitude in Philosophy*. Eds. Goldberg & Walker.
- Helton, G. (unpublished manuscript). Against the doxastic theory of perception.
- Helton, G. & Register, C. (forthcoming). Hot-cold empathy gaps and the grounds of authenticity. *Synthese* special issue on Transformative Experience

SUPPLEMENTAL WORK CITED

- 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 *Advances in experimental social psychology* (Vol. 48, pp. 117-171). Academic Press.