

Self-Knowledge and Interpersonal Reasoning

BENJAMIN WINOKUR

doi:[10.48106/dial.v76.i4.02](https://doi.org/10.48106/dial.v76.i4.02)

Benjamin Winokur. 2022. “Self-Knowledge and Interpersonal Reasoning.” *Dialectica* 76(4): 547–570.
doi:[10.48106/dial.v76.i4.02](https://doi.org/10.48106/dial.v76.i4.02).



Self-Knowledge and Interpersonal Reasoning

BENJAMIN WINOKUR

Many philosophers contend that we often possess “privileged” and “peculiar” self-knowledge of our mental states. Self-knowledge is privileged insofar as it is systematically more secure than the knowledge that others have of one’s propositional attitudes, and it is peculiar insofar as it is systematically obtained in a way that is only suited for delivering self-knowledge. Focusing on privileged and peculiar self-knowledge of propositional attitudes like beliefs, I offer an account of its instrumental value. On my account, privileged and peculiar self-knowledge of one’s propositional attitudes enables one to be a more efficient and reliable interpersonal reasoner.

Self-knowledge of one’s current mental states often seems interesting—if not outright puzzling—for at least two reasons. First, such self-knowledge often seems to be *privileged*, for it seems to be systematically (though not universally) more secure than the knowledge one has of others’ mental states. Second, it often seems to be *peculiar*, for it seems to be systematically (though, again, not universally) obtained in a way that is only suited for delivering self-knowledge, hence, *not* by whatever means enable one to acquire knowledge of *other* minds (Byrne 2018, 4–9). The standard project in contemporary theorizing about self-knowledge is to vindicate these appearances by unearthing the special security and sources of self-knowledge. However, others have argued that we do not actually possess any privileged and peculiar self-knowledge (hereafter “PPSK”), at least when it comes to self-knowledge of propositional attitudes like belief (Gopnik 1993; Carruthers 2011; Cassam 2014). These PPSK-skeptics typically understand self-knowledge and other-knowledge of propositional attitudes as on a par in terms of their security, source, or both.

In reply, some PPSK-realists have offered competing interpretations of the putative evidence against realism about PPSK of propositional attitudes (Parent 2017; Keeling 2019a, 2019b; Andreotta 2021; Valaris 2018), while

31 others have pushed back against the non-privileged and non-peculiar accounts
32 of self-knowledge that are favoured by many skeptics (Coliva 2016; Keeling
33 2018; Marcus and Schwenkler 2019; Andreotta 2022). The stakes of these
34 debates are hard to grasp if we are unsure “what, if anything, of value we fail
35 to possess if these skeptics are right” (Peterson 2021, 365). For this reason, I
36 will argue that PPSK of one’s propositional attitudes—chiefly, our beliefs—
37 is instrumentally valuable for the efficiency and reliability of a widespread
38 activity in our social-epistemic lives, that of *interpersonal reasoning*. Some
39 readers may interpret my arguments as providing further support for PPSK-
40 realism if they believe that interpersonal reasoning is in fact a highly efficient
41 and reliable activity in our actual lives. Other readers might reach the more
42 modest conclusion that interpersonal reasoning is a more effective enterprise
43 *to the extent that we possess PPSK*, whether or not we really possess PPSK, and
44 hence whether or not interpersonal reasoning is a particularly efficient and
45 reliable activity for us to undertake. Either way, the significance of debates
46 between PPSK-skeptics and PPSK-realists can be better appreciated in light
47 of what follows.

48 Here is the layout for my paper. In section 1 I draw initial inspiration from
49 two earlier accounts of PPSK’s instrumental value. The first, due to Sydney
50 Shoemaker (1988, 1996), concludes that social cooperation in general requires
51 each of us to possess PPSK of many of our propositional attitudes. The second,
52 due to Charles Siewert (2003), concludes that PPSK is indispensable to social
53 cooperation *whenever this depends on justifying one’s actions to others*. Justify-
54 ing one’s actions to others can be one way of reasoning with others, that is,
55 reasoning interpersonally. But it is *only* one way of reasoning interpersonally.
56 I thus consider, in section 2, whether *all* interpersonal reasoning might benefit
57 from PPSK. My argument is that PPSK does indeed play beneficial roles in all
58 interpersonal reasoning. In section 3 I address objections to my account. In
59 section 4 I consider another recent account of PPSK’s instrumental value, one
60 that emphasizes its role in our capacity for “epistemic control,” and I show
61 how my account complements that account. In section 5 I conclude.

6.1 Cooperation and Privileged, Peculiar Self-Knowledge

63 Is PPSK instrumentally valuable? Some philosophers have argued that it is.¹
 64 Indeed, some have argued that it is instrumentally *indispensable*. Here is
 65 Shoemaker, who writes of self-knowledge by “self-acquaintance” instead of
 66 privileged and peculiar self-knowledge:²

67 When one is engaged in a cooperative endeavor with another, it
 68 is essential to the efficient pursuit of the shared goal that one be
 69 able to communicate to the other information about one’s beliefs,
 70 desires and intentions [...] When in such circumstances one conveys
 71 one’s beliefs to another, this is not merely for the purpose
 72 of conveying what one takes to be information about the world,
 73 namely, the contents of the beliefs; it is also for the purpose of
 74 giving him information about oneself which will assist him in
 75 predicting one’s behavior and so in coordinating his own behavior
 76 with it, and also to enable him to correct those of one’s beliefs he
 77 knows to be mistaken [...] And here the utility of self-knowledge
 78 depends crucially on its being acquired by self-acquaintance; if
 79 I had to figure out from my behavior what my beliefs, goals, in-
 80 tentions, etc. are, then in most cases it would be more efficient
 81 for others to figure this out for themselves than to wait for me to
 82 figure it out and then tell them about it. (1988, 185–186)

83 Shoemaker argues that PPSK is indispensable for efficiently cooperating with
 84 other human beings. For, if others could know one’s mind in the same way
 85 and as reliably as one knows one’s own mind, one would be far less efficient
 86 at soliciting others’ cooperation. This is because it would just as often be
 87 up to others to figure out one’s mind, and to decide on this basis whether
 88 cooperation was worthwhile. As a result, one would frequently fail to solicit
 89 others’ cooperation of one’s own accord.

1 Peterson (2021, 365) thinks that the question of PPSK’s value has been ignored by epistemologists working on self-knowledge. While I myself hope to contribute an answer to this question, I think that this assessment of the extant literature is somewhat exaggerated given the views that I discuss in this section, among others (see, e.g., Burge 1996; Nguyen 2015; Sorgiovanni 2019; Winokur 2021a, 2021b). Peterson’s own account of PPSK’s instrumental value is discussed in section 4. He also discusses the potential *intrinsic* value of PPSK, a topic that I do not broach here.

2 These, I submit, are just notational variants.

90 Reflecting on Shoemaker’s argument, Siewert wonders whether rational
 91 animals “could engage in cooperation and assistance-seeking behaviour, even
 92 by generally acting in an attitude-revealing fashion, without representing
 93 their own minds to themselves” (2003, 139). In a different idiom: couldn’t
 94 there be creatures that are exceptionally adept at *expressing*—i.e., showing,
 95 manifesting, displaying—their attitudes to their fellow creatures without also
 96 possessing *PPSK of the attitudes expressed*, and couldn’t this enable equally
 97 efficient cooperation?³ Contra Shoemaker, Siewert supposes that there could
 98 be such creatures. Still, he is optimistic about a nearby argument:

99 For whether or not there can be social animals that act in a usefully
 100 self-revealing fashion while oblivious to their own psychologies,
 101 they could not engage in the practice of *justifying* such acts, with-
 102 out being able to represent, in their justifications, relevant facts
 103 about their own desires and beliefs [...] Now, if the reasons we
 104 would offer did not have us acting in ways revealing our actual
 105 beliefs and desires to others, we would be much less effective in
 106 securing others’ cooperation and assistance in the satisfaction of
 107 our desires than we in fact are. (Siewert 2003, 139)

108 On this argument, it is not that efficient cooperation always requires *PPSK*.
 109 Rather, such self-knowledge is required for cooperation *whenever such cooper-*
 110 *ation also depends on justifying one’s actions to one’s would-be cooperators*. For,
 111 lacking *PPSK*, our actions would often fail to cohere with the attitudes that
 112 we self-ascribe. In turn, we would be worse at justifying our actions because
 113 we would be worse at appealing to the actual beliefs, desires, and intentions
 114 that underwrite them. These inconsistencies might be noticed by others, and
 115 this might diminish their trust in us.

116 More recently, Jon Greco has written that:

117 *Of course* thinking about one’s first-order mental states is essen-
 118 tial to activities involving coordination and cooperation [...] In
 119 particular, *giving one’s reasons*, both epistemic and practical, is es-
 120 sential to various activities in which one must defend one’s beliefs
 121 and actions, and having a grasp on such mental states oneself is
 122 essential to reporting them to others. (2019, 52)⁴

3 This is my gloss on Siewert’s argument. Like Bar-On (2004), I use “express” here to denote actions that express mental states, though I denote another sense of expression in section 3.

4 See Müller (2019, 6) for a similar view.

123 Like Siewert, Greco claims that self-knowledge is essential for justifying one's
124 actions (and beliefs) to others. However, Siewert argues that PPSK is indispens-
125 able to our widespread success in these matters, whereas Greco claims that
126 “this kind of metacognitive activity can tolerate the same fallibility that we
127 experience in cognition generally” (2019, 53).⁵ He thus denies the importance
128 of *privileged* self-knowledge (and is silent about *peculiar* self-knowledge). But
129 he does not consider Siewert's argument, and so it is hard to know whether
130 his position would change upon further reflection.

131 This difference between Greco and Siewert set aside, notice that they both
132 focus on a certain kind of interpersonal reasoning. Here, ‘interpersonal reason-
133 ing’ denotes exchanges of assertions between interlocutors, or exchanges
134 of questions and assertions, toward a discursive end. For instance, one might
135 reason interpersonally in order to acquire new rational attitudes, or to subject
136 one's already-held attitudes to the scrutiny of other agents, or to persuade
137 other agents to adopt one's already-held attitudes. I say that Siewert and Greco
138 are focused on a certain kind of interpersonal reasoning because they only
139 focus on cases in which agents reason interpersonally about one another's
140 actions or attitudes. In other words, neither philosopher focuses on cases
141 in which agents aim to justify “agent-neutral” propositions to one another,
142 these being propositions whose contents do not refer to any particular agent's
143 actions or attitudes. One such proposition is:

144 Runaway climate change is a worsening phenomenon.

145 It is to be contrasted with the sorts of propositions that Siewert and Greco
146 focus on, namely, “agent-specific” propositions like:

147 I should continue to be vigilant about my fossil fuel consumption.

5 Greco also ventures a response to the possibility of efficiently cooperative animals lacking self-knowledge: “One might object that non-human animals are also social in a sense that implies coordination and cooperation, and they manage their social lives without citing their mental states in explanations to themselves or their cohorts. But this objection misses the point that human social agency is also rational agency. It involves rationalizing one's thoughts and actions by means of giving one's reasons—i.e., overtly giving one's reasons—to oneself and to others” (Greco 2019, 53). This too is reminiscent of Siewert's view. But while Greco denies that cooperation among non-human animals involves rational agency, Siewert thinks that non-human animals could count as rational agents in so cooperating.

148 The latter proposition, but not the former, requires agents to provide self-
 149 referential information, this being information that justifies *the agent herself*
 150 to act in such-and-such a way or have such-and-such an attitude. Such infor-
 151 mation will naturally include “relevant facts about their own desires and be-
 152 liefs” (Siewert 2003, 139) whereas reasoning about agent-neutral propositions
 153 simply requires providing first-order evidence about the agent-independent
 154 world, e.g., evidence of rapidly melting arctic ice. But because reasoning about
 155 either sort of proposition can be conducted interpersonally, Siewert and Greco
 156 will have shown—at most—that self-knowledge matters for interpersonal
 157 reasoning about agent-specific matters (whether such self-knowledge is privi-
 158 leged and peculiar as Siewert claims, or not, as Greco claims). I note this here
 159 because I will argue in section 2 that PPSK plays a role in both agent-specific
 160 and agent-neutral interpersonal reasoning.

161 Before I get there, I want to make two preliminary points. First, the reader
 162 may have wondered whether that Shoemaker’s and Siewert’s arguments estab-
 163 lish what they purport to since, on close inspection, they seem to emphasize
 164 the importance of privileged access but not, in addition, peculiar access. This
 165 is because each argument insists that the special security of agents’ self-
 166 knowledge is what facilitates cooperation with other people, and yet this does
 167 not obviously entail that agents must exploit a peculiar means of achieving
 168 such security. In what follows I will provide arguments for the importance of
 169 peculiar self-knowledge as well, thereby going beyond the arguments consid-
 170 ered thus far.

171 Second, it should be noted that some philosophers deny that interpersonal
 172 reasoning of *any* kind (i.e., whether about agent-neutral or agent-specific
 173 propositions) requires self-knowledge of *any* kind (i.e., whether privileged
 174 and peculiar, non-privileged and non-peculiar, or any other combination).
 175 For example, Robert Brandom writes that there is “nothing incoherent in
 176 descriptions of communities of judging and perceiving agents, attributing
 177 and undertaking propositionally contentful commitments, giving and asking
 178 for reasons, who do not yet have available the expressive resources *I* pro-
 179 vides” (1994, 559). If these communities lack articulate use of the first-person
 180 singular, then they cannot self-ascribe and hence self-know their attitudes.⁶

6 See also Stribos and de Bruin (2012). The importance of this claim depends on assuming that self-knowledge requires linguistically articulate self-ascriptive thought, and some friends of “tacit” self-knowledge might dispute this (e.g., Boyle 2011, 2019). Alternatively, it could be granted that there is such a thing as self-consciousness that does not involve linguistically articulate self-ascriptive thought (cf. Musholt 2015, chap. 4). Even if this is a tenable view, I am focusing

181 Similarly, Ladislav Koreň claims that we can reason interpersonally by exer-
182 cising a “practical competence” with linguistic devices like “no,” “but,” and
183 “so,” thus manifesting a “sensitivity” to rational connections between claims
184 without having “metarepresentational” beliefs about the rational connec-
185 tions between one’s own attitudes or one’s interlocutor’s attitudes (2023, 853).
186 Finally, Annalisa Coliva offers the following thought experiment:

187 Take a subject who is able to judge that P, give evidence in favour
188 of it and withdraw from it if required and, therefore, has the first-
189 order belief that P based on judgement. Suppose you ask her “Do
190 you believe that P?” and she is unable to answer. You conclude
191 that she does not have the concept of belief. (2016, 191)

192 This is a situation in which one interlocutor reasons interpersonally while, *ex*
193 *hypothesi*, lacking the conceptual wherewithal to self-ascribe the attitudes that
194 her assertions express. Coliva adds that any such agent will at least possess
195 “the *ability* to differentiate between, for instance, believing P and P’s being the
196 case, by being *sensitive to the fact that* her point of view may be challenged [...]”
197 (2016, 192, emphasis mine). On my reading, the emphasized terms suggest
198 that such an agent utilizes *pre-metarepresentational* capacities in the service of
199 interpersonal reasoning; these abilities and sensitivities enable her to reason
200 with others without forming second-order beliefs about her first-order beliefs
201 or her interlocutor’s first-order beliefs.

202 These philosophers clearly reject Greco’s claim that “thinking about one’s
203 first-order mental states is essential to activities involving coordination and
204 cooperation,” given that interpersonal reasoning is itself a coordinated and co-
205 operative endeavour. But do they extend this rejection as far as to deny that in-
206 terpersonal reasoning *with an aim to justifying one’s own actions and attitudes*
207 requires self-knowledge or, at the very least, some form of self-representation
208 like a self-belief? As Steven Levine makes clear in a response to Brandom, it
209 is hard to see how they could cogently deny this. Levine begins by acknowl-
210 edging the possibility of agents who reason interpersonally insofar as the
211 assertions at issue are first-order assertions of the form “that-P,” these being
212 expressions of agent-neutral propositions in the sense described above. As
213 regards assertions of these propositions, “the performer can justify the state-
214 ment without explicitly claiming that it is he who is justifying the statement

on what epistemologists in this area ordinarily focus on, i.e., *explicit* self-knowledge involving linguistically articulate self-ascriptive thought (*pace* also those who view self-knowledge as an *ability*—cf. Campbell 2018).

215 [...] because this assertion concerns an objective state of affairs that can be
 216 justified by agent-neutral reasons” (Levine 2009, 111). However:

217 [...] is this the case when the assertion that is being challenged
 218 concerns an agent’s *own* action or perception? Here what is being
 219 challenged is, for example, one’s entitlement to perform an action
 220 or one’s entitlement to claim that one’s perception is veridical.
 221 In either case, the justificatory reasons offered cannot be agent-
 222 neutral in the way that reasons justifying the assertion “that-P”
 223 are. (Levine 2009, 111)

224 So Levine is in league with Siewert and Greco in arguing that, when one’s own
 225 actions are challenged by an interlocutor, one cannot merely avail oneself of
 226 agent-neutral reasons. Instead, one must avail oneself of agent-specific rea-
 227 sons, which will include facts about one’s own psychology. The only question
 228 is whether Levine would side with Siewert in understanding these exchanges
 229 as requiring PPSK on the part of whoever seeks to justify her own attitudes,
 230 or with Greco in denying any indispensable role for such epistemically high-
 231 grade self-knowledge.

232 As aforementioned, I will soon argue that PPSK plays important roles in
 233 interpersonal reasoning about both agent-specific and agent-neutral proposi-
 234 tions. But how can I be headed in this direction, having just traced a dialectic
 235 that only acknowledges a role for self-knowledge in interpersonally defending
 236 agent-specific propositions about one’s own actions or perceptions? In other
 237 words, if it is conceded to Brandom and others that agents can reason inter-
 238 personally about agent-*neutral* propositions without so much as a capacity
 239 for self-belief, then isn’t it foolish to contend that PPSK—let alone any other
 240 sort of self-knowledge—matters for such activity? Fortunately, there is no
 241 real problem here. My argument will be that PPSK contributes to interper-
 242 sonal reasoning *for agents who in fact possess the capacity for representing*
 243 *themselves and their beliefs in higher-order thought*. This focus allows me to
 244 grant Brandom, Koreň, and Coliva their contention that some agents can
 245 reason interpersonally despite lacking this metarepresentational capacity.⁷

7 There are other ways to dispute the indispensability of self-knowledge for interpersonal reasoning. For example, Roelofs (2017) argues that no such knowledge is required by interpersonal reasoners who are “evidentially unified” with and “cognitively vulnerable” to one another. Evidentially unified agents are automatically attuned to one another’s evidence without having to explicitly share it, while cognitively vulnerable agents can rationally cause changes in one another’s minds through cognizing this unified evidence (they can induce such changes as *basic actions*).

246 What I will argue is that agents who *do* possess this capacity, such as most
 247 cognitively developed adult human beings, are systematically vulnerable to
 248 certain deficiencies in interpersonal reasoning if they lack **PPSK**.

249 **Interpersonal Reasoning and Privileged, Peculiar** 250 **Self-Knowledge**

251 Oftentimes, cognitively developed adult human beings have knowledge—or
 252 at least beliefs—about their own attitudes, and they often have further beliefs
 253 about how their attitudes converge with or diverge from their peers. It is
 254 often these higher-order states of mind that motivate agents to reason with
 255 one another in the first place. After all, if one agent believes that there is a
 256 discrepancy between what she believes and what her interlocutor believes,
 257 this can help to explain why she bothers to try and settle the discrepancy
 258 through an interpersonal exchange of reasons.

259 For a hypothetical example, consider two interlocutors: Maya and Roman.
 260 Maya might aim to convince Roman that climate change is an existential
 261 threat to human civilization (note that this is an agent-neutral proposition: I
 262 emphasize the importance of this fact near the end of this section). My claim
 263 now is this: Maya would be in a precarious epistemic position, one that might
 264 undermine the efficiency of her reasoning with Roman, or one that might
 265 even make it better for her to *not* try to reason with Roman about this issue, if
 266 she did not possess **PPSK**.

267 Why so? It is easiest to begin by focusing on *privilege*. Here is the basic
 268 idea: if Maya were not in a systematically superior epistemic position regard-
 269 ing her beliefs about her attitudes than Roman was concerning his beliefs
 270 about Maya's attitudes, then Roman could more easily—i.e., with better epis-
 271 temic grounds—convince Maya that her attitudes already align with his. In
 272 convincing Maya of this, Roman would be providing second-order grounds
 273 for skepticism about Maya's belief that she believes climate change to be an

Evidential unity and cognitive seem conceptually possible, and they might even be achieved by actual agents who are wired to one another's brains in the right sorts of ways. The upshot is that neither party must have "I"-thoughts about their selves and attitudes in the course of interpersonal reasoning nor, for that matter, thoughts about others' selves and attitudes. Instead, by focusing strictly on first-order reasons, they can automatically adjust one another's attitudes. However, Roelofs admits that, for us, "it seems very unlikely...that such a close rapport could persist for very long, or cover very many topics" (2017, 17). We are simply not wired to one another in these ways, at least not with any real consistency. Accordingly, what I say below applies to ordinary agents who lack evidential unification and cognitive vulnerability.

274 existential threat to human civilization. As a result, Maya would not even
 275 bother to reason with Roman about the first-order discrepancy, because her
 276 self-belief would change in such a way that she no longer took there to be
 277 any such discrepancy. Roman might alter Maya's self-belief in good faith by
 278 providing evidence that it is mistaken. But in other cases, Roman might operate
 279 in bad-faith by knowingly supplying Maya with misleading grounds for
 280 the same conclusion. Indeed, if Roman's testimony is a source of evidence all
 281 on its own then, given Maya's lack of privileged access to her own belief, her
 282 epistemic situation upon receiving Roman's testimony is immediately altered
 283 even if Roman supplies no independent evidence in favour of his testimony.
 284 In such cases we could say—perhaps somewhat overdramatically—that Maya
 285 has been taken as Roman's *epistemic hostage*. As an epistemic hostage, Maya
 286 succumbs to Roman's efforts (good faith or otherwise) to convince her that her
 287 self-ascribed attitudes are not really her own. Maya, being falsely convinced
 288 of this, is even cut off from opportunities to reason with agents *other than* Roman
 289 about climate change, given that she has been pre-emptively convinced
 290 that she does not disagree with those—like Roman—who are climate science
 291 deniers.

292 We might construe these situations as threats to Maya's epistemic auton-
 293 omy. I say this because, plausibly, epistemic autonomy is at least partly a
 294 matter of being able to navigate various interpersonal reasoning contexts
 295 without having one's self-conception co-opted too easily by others. Indeed,
 296 this matters even if we are sometimes duped about the *first-order* issues by
 297 clever interlocutors who supply us with misleading evidence at *that* level
 298 of discourse (e.g., misleading statistics suggesting that climate change—of
 299 the anthropogenic variety, at least—is not taking place). An agent who is
 300 convinced by a clever interlocutor that the evidence for climate change is bad
 301 is still an agent who has assessed those reasons for herself and hence has been
 302 mislead on a basis that still deploys her own rational faculties to some degree.
 303 And while it is true that Maya might also deploy her own rational faculties in
 304 assessing Roman's claim that her *self-belief* is wrong, perhaps because Roman
 305 supplied her with good reasons (by *her* lights, at least) to do so, the *result* is
 306 that Maya lacks the self-knowledge that she needs in order to recognize that
 307 there is a discrepancy between her belief about climate change and Roman's
 308 belief about it, and *this* undermines her epistemic autonomy for reasoning
 309 with Roman about climate change itself.

310 Now, as aforementioned, this account of PPSK's instrumental value is
 311 most clearly geared toward *privileged* self-knowledge, since it is an argument

312 about what happens when the epistemic security of Maya's self-beliefs is,
313 as a general matter, no better than that of Roman's perspective unto Maya's
314 mind. But the account can extend to *peculiarity* as well, at least if we construe
315 the relationship between privilege and peculiarity in such a way that Maya's
316 privilege *is due to* the peculiar way in which she knows her own mind (cf.
317 Peterson 2021). For, if her self-beliefs are not generally acquired by a peculiar
318 means that is generally available to her, then nothing prevents individuals
319 like Roman from seizing upon the very same means to acquire knowledge of
320 Maya's mind, and this makes it harder to understand why Maya's self-beliefs
321 are, in general, so epistemically secure that Roman's contrary claims or beliefs
322 do not give Maya strong reason to change what she believes about herself.

323 To bring this point into sharper relief, we can consider a putative foil for my
324 account, namely, Quassim Cassam's *Inferentialist* account of self-knowledge.
325 According to Cassam, both self-knowledge and other-knowledge of agents'
326 attitudes are acquired through inferences. On his view, there remains an
327 epistemic asymmetry between self-knowledge and other-knowledge, but this
328 asymmetry simply "boils down to a difference in the kinds of evidence that are
329 available in the two cases" (Cassam 2014, 150). More specifically, the evidence
330 that one has about one's own attitudes is superior to the evidence that one has
331 about others' attitudes because it includes sensations, memories, and other
332 non-attitudinal mental goings-on that are not so easily accessed by one's peers.
333 Applying this view to interpersonal reasoners like Maya, we might say that
334 Maya's self-knowledge of her attitudes is privileged to some degree even if the
335 same method—inference—is used by both Maya and Roman in coming to
336 form beliefs about Maya's attitudes. So there is nothing peculiar about Maya's
337 route to self-knowledge. But now one might insist that Maya cannot be easily
338 taken as an epistemic hostage by Roman, even though she lacks a peculiar
339 way of knowing herself, simply because she has especially good evidence
340 about herself.

341 However, it could happen that such additional evidence is unavailable to
342 Maya in any number of cases, for what reason can be given for thinking that
343 Maya will always have access to special evidence, given that access to evidence
344 in general is a contingent matter of one's epistemic position relative to a body
345 of information? Peter Carruthers—another prominent Inferentialist—takes
346 it that we have privileged self-knowledge of non-propositional-attitudinal
347 mental states (2011), and contends that this can be used as a basis for inferring
348 our propositional attitudes. However, privileged access to these other mental
349 states can only provide a basis for inferring our propositional attitudes *when*

350 *we are in such mental states*, and yet this itself is a contingent matter. Moreover,
351 even stipulating that Maya has systematically better evidence about herself
352 than Roman has about her, we would also need a general assurance that Maya
353 infers the correct conclusions from this systematically superior evidence. It
354 could happen that Maya has privileged access to the evidence about what she
355 herself believes but cannot reliably *utilize* this evidence. At the very least, it
356 could happen that she is, in general, no better at utilizing this evidence than
357 Roman is at utilizing *his* evidence about Maya's attitudes. Indeed, philosophers
358 like Carruthers seem to embrace this point when they claim that Inferentialist
359 views best explain failures of self-knowledge.

360 Finally, Inferentialist views are vulnerable to what I call an *efficiency concern*
361 and a *gridlock concern*. The efficiency concern is that, absent peculiar access, it
362 could be generally appropriate for Roman to ask Maya to supply the grounds
363 for her self-beliefs, and for Maya to ask Roman to do the same, just to be
364 sure that they were operating in a case where Maya really did have (and had
365 effectively utilized) this superior evidence. Engaging in this second-order
366 interpersonal reasoning would significantly slow down their efforts to get to
367 the first-order issues, thus rendering interpersonal reasoning about first-order
368 issues a less efficient activity. The gridlock concern is that the second-order
369 issue might not get resolved at all whenever both parties fail to reach a verdict
370 about what Maya believes. One might attempt to circumvent these concerns
371 by arguing that Maya's inferences are subpersonal or non-conscious, such that
372 she cannot be expected to articulate them to Roman. But inferences that are
373 not available for peer-review are also inferences that Maya might be required
374 to lower her trust in, thus calling her self-beliefs into question all over again.
375 To be sure, if some sort of Inferentialism is true, it may follow that agents like
376 Maya often have better evidence and draw better inferences about their own
377 attitudes than their interlocutors can draw about her attitudes, but the points
378 I have been making suggest that such access will be *worse* for Maya than any
379 form of access that renders the special epistemic security of her self-beliefs a
380 non-contingent matter.

381 Now, even though I have been critiquing an Inferentialist rejection of pecu-
382 liar access, I want to reiterate a general lesson for all would-be skeptics about
383 such access. The lesson is that, if the same method—whether inferential or
384 otherwise—is used for acquiring both self-knowledge and other-knowledge,
385 then epistemic privilege will seem to be highly contingent. For, if two agents
386 can come to know one agent's mind by the same means, then there need be
387 no systematic barrier to their doing with equal epistemic pedigree. In the

388 context of my account, this would mean that there is no strong assurance
389 that agents are systematically warranted in retaining their self-beliefs when
390 challenged by their interlocutors. And this, in turn, would mean that there
391 is no general assurance that interpersonal reasoning *about the world*, rather
392 than about the interlocutors' minds, can proceed smoothly. The efficiency and
393 gridlock concerns also generalize: if Maya and Roman share the same method
394 for arriving at a view about Maya's mind, then Roman might endeavour to
395 interrogate Maya about whether their current context is one in which she
396 has exercised the method more effectively, whether the method is inferential
397 or not. This would slow down and (potentially) gridlock the discourse at the
398 second-order level. Crucially, though, I am not claiming that the systematic
399 protection provided by PPSK against these concerns is universal in scope.
400 For my purposes, PPSK's instrumental value will have been demonstrated if
401 it is our standard sort of self-knowledge. This would ensure that one is not
402 *systematically, generally, or universally* vulnerable to innocent-yet-erroneous
403 self-belief change, bad-faith epistemic hostage-takers, or to the efficiency
404 and gridlock concerns, thus improving interpersonal reasoning's reliability
405 and efficiency as a tool in our social-epistemic toolkit for understanding our
406 shared world.

407 So goes my account. If correct, it shows that PPSK is instrumentally valuable
408 for interpersonal reasoning, at least among those who are in a position to
409 form beliefs about their own attitudes in the first place (again, a child who
410 has yet to acquire the concept of belief cannot be erroneously convinced that
411 she *shares a belief* with someone else). Notably, the account applies whether
412 we imagine interpersonal reasoners as aiming to debate an agent-neutral
413 proposition or an agent-specific one. I initially described Maya as aiming to
414 convince Roman that climate change is an existential threat—this being an
415 agent-neutral proposition—whereupon Roman steers the discourse to the
416 second-order level in order to convince Maya that she does not really believe
417 this in the first place. But the content of the proposition was incidental to the
418 example. Had the proposition's content been agent-specific, e.g., about Maya's
419 particular climate-focused actions or the belief-desire pairs that rationalize
420 her actions, Roman might have proceeded in the same way. So, my account has
421 a broader scope than Siewert's: it applies to agent-specific *and* agent-neutral
422 interpersonal reasoning.

423 Addressing Objections

424 In this section I reinforce my account by addressing four objections. The
 425 first objection is that legitimate challenges to our self-knowledge are in fact
 426 quite frequent, and that this provides evidence against the claim that PPSK
 427 frequently serves as an epistemic shield against erroneous self-belief change
 428 in our actual interpersonal reasoning practices. The second objection is that
 429 PPSK does not suffice to ensure that interpersonal reasoning is a reliable
 430 route for rational attitude adjustments.⁸ The two final objections are specific
 431 defenses of the claim that factors beyond PPSK can protect interpersonal
 432 reasoners against erroneous self-belief change in interpersonal reasoning
 433 contexts.

434 The first objection turns on familiar cases of self-deception. Self-deception is
 435 ordinarily taken as a failure of self-knowledge in which an agent self-ascribes
 436 an attitude that she in fact lacks. Those who take us to have privileged self-
 437 knowledge surely ought to say something about this familiar phenomenon.
 438 If one does not take privileged access to be universal in scope, then it is at
 439 least logically possible to accommodate such cases. Alternatively, one might
 440 deny the ordinary view of self-deception by arguing that it does not involve
 441 false self-ascriptions (Bilgrami 2006; Coliva 2016). The apparent trouble for
 442 my account, however, is that accusations of self-deception are frequent and
 443 potentially epistemically legitimate in many cases, and yet these might be
 444 precisely the moves that our interlocutors use in order to convince us that
 445 our self-beliefs are false. If accusations of self-deception are epistemically
 446 legitimate and widespread, and if these accusations can spur agents to adjust
 447 their self-beliefs, then what protection does PPSK really provide here?

448 To begin my response, I want to reiterate a point from my introductory
 449 remarks about the dialectical ambitions of this paper, namely, that readers
 450 need not be convinced that I have unearthed PPSK's actual functional role
 451 for interpersonal reasoners at this world. Secondly, when I say that PPSK
 452 provides an epistemic shield against challenges to one's self-beliefs in inter-
 453 personal reasoning, I do not deny that people might sometimes fail to take
 454 advantage of this shield—PPSK offers *epistemic* protection that may not be
 455 *psychologically* appreciated. Beyond these somewhat concessionary responses,
 456 the devil must reside in the details, since any further response depends on
 457 how we understand the cases at issue. Thus, consider a case in which Maya

8 These first two objections were put to me by Rachel Cooper.

avows a love of comic books and Roman replies: “you do not love comic books; you’ve just tricked yourself into thinking that loving comic books makes you interestingly different.”⁹ What might bring Maya to accept this accusation? Well, Maya might fixate on the thought that her interlocutor has better evidence about her mind than Roman has about it. If she wondered about her own evidence, and wondered about its inferential role in supporting her self-beliefs, she would be supposing her own self-ascription to be vulnerable to the same epistemic standards that Roman uses to evaluate her self-beliefs. If her self-knowledge is peculiar, however, she will not fixate on this thought, because her self-ascription is *not* based on the same epistemic standards.

In fact, our actual manner of proceeding tends to bear this out: one’s interlocutor judges one to be self-deceived about one’s love of comic books, and one responds *not* by attempting (and possibly failing) to offer higher-quality evidence *about what one believes*, but by offering reasons about *why comic books are loveable*. Indeed, one possible *explanation* of privilege and peculiarity is that one’s own take on the reasons for or against adopting some attitude (typically) determine one’s adoption of it. And if one self-ascribes this attitude with full knowledge of the first-order reasons that one takes to support it, one is entitled to make this self-ascription even if other people have evidence contravening one’s self-ascription (Bilgrami 2006; Coliva 2016).¹⁰

Moreover, if we have PPSK, *other* challenges to our self-beliefs may also be illegitimated, these being challenges where other agents do not accuse us of being self-deceived but, rather, as having made innocent (or “brute”¹¹) errors about ourselves—errors that could only be made on the basis of innocent inferential or observational mistakes.

Here is another, final sense in which the devil is in the details. The objection under consideration is that accusations of self-deception are common, and that these accusations might frequently lead to (reasonable) changes in one’s self-beliefs. However, while such cases may indeed be common, they may only be common in the sense that *all of us* are *occasionally* susceptible to them. On this explanation of their commonality, no single agent is liable to be the reasonable target of an overwhelmingly large number of self-deception accusations. There is something suspicious about anyone, even one’s therapist,

⁹ I owe this example to Rachel Cooper.

¹⁰ Compare Schwengerer’s verdict on two cases he discusses (2021, 77). What I may owe my interlocutor, in this case, is an explanation of how my actions fail to live up to my self-ascribed attitude, *not* an explanation to the effect that the evidence shows that I have this attitude.

¹¹ For the operative notion of brute error, see Burge (1996) and Bar-On (2004).

491 who would unrelentingly accuse one of self-deception across myriad cases
 492 by saying things like “you do not believe that-P, nor hope that-Q, nor desire
 493 to φ , nor love S....” This suspicion may well reflect a fact about us: that we
 494 have enough PPSK to be reasonable in *not* giving in to too many accusations
 495 of self-deception—accusations which, if legitimate, would force us to change
 496 our self-beliefs.

497 The second objection to my account is that PPSK does not improve the
 498 reliability of interpersonal reasoning even if it provides us with epistemic
 499 warrant to ignore (many) accusations of mistaken self-belief. Cases in favour
 500 of this objection are easy enough to set up. For example, maybe Maya claims
 501 that climate change is an existential threat to human civilization and Roman
 502 gives insufficient epistemic uptake to her assertion because he is prejudiced
 503 against women. Indeed, in this case, Maya may be the victim of a “testimonial
 504 injustice” (Fricker 2007). But I want to offer two observations. First, although
 505 the factors preventing Roman from reasoning with Maya have nothing to do
 506 with Maya’s self-knowledge or Roman’s beliefs about Maya’s self-perspective,
 507 this does not change the fact that Maya would have an *additional* problem
 508 on her hands if Roman were generally in an epistemic position to make
 509 Maya erroneously change her self-beliefs. Second, to the extent that Roman’s
 510 prejudiced behaviour does not prevent Maya from knowing herself, she is still
 511 in a position to congregate with less prejudiced individuals and to reason with
 512 *them* (or even to reason with Roman *indirectly* by reasoning with someone
 513 that Roman is *not* prejudiced against, and getting that individual to convey
 514 Maya’s reasons to Roman). This point also applies to another concern, namely,
 515 that Roman might simply say that *he* agrees with Maya when *he* does not (this
 516 being an inverse version of the epistemic hostage-taking tactic). Maya may
 517 not be able to rationally challenge this claim if Roman has PPSK, unless she
 518 has reason to deem him insincere, since she will then have strong reason to
 519 take Roman at his word. Once again, though, this would not put Maya in the
 520 position of being made to form a false belief about what she herself believes
 521 about the world, and hence she would not be prevented from discoursing
 522 with other agents about the contents of her beliefs about the world.

523 I now address two objections to the effect that something other than PPSK
 524 can explain why we are protected against epistemic hostage-taking. According
 525 to the first objection, what *really* protects Maya against Roman’s nefarious
 526 machinations is the same thing as what explains her *first-person authority*,
 527 where what explains *this* is something other than PPSK. Roughly, “first-person
 528 authority” denotes two claims: (1) it is epistemically rational to presume

529 the truth of speakers' present-tense self-ascriptions of mental states, and (2)
530 it is typically epistemically irrational to interrogate the epistemic grounds
531 of speakers' present-tense self-ascriptions (hereafter just 'self-ascriptions').¹²
532 Now consider an "expressivist" explanation of first-person authority which
533 contends that speakers' self-ascriptions ought to be presumed true and be
534 insulated from requests for epistemic support because they express and thus
535 *show* the self-ascribed mental states to one's hearers (Bar-On 2004). This
536 explanation is available even if speakers do not *also* possess PPSK of the
537 mental states that their self-ascriptions express. The objection, then, is that
538 Maya's first-person authority gives Roman a strong reason not to challenge
539 most of her self-ascriptions, such that PPSK is explanatorily superfluous
540 in explaining why Roman is not likely to give Maya an erroneous basis for
541 changing her self-beliefs.

542 Now, for all I have said, Maya's self-ascriptions may be first-person authori-
543 tative in virtue of what they express, whether or not Maya also has PPSK of
544 what they express. Nevertheless, I argue that without *also* possessing PPSK,
545 Roman could purposefully *ignore* the first-person authority of Maya's self-
546 ascriptions in a bid to convince her that her attitudes converge with rather
547 than diverge from Roman's. He might (rightly) take Maya to have expressed
548 her first-order belief through a self-ascription but still claim that her self-
549 belief is false. Hence, PPSK protects Maya against being manipulated by bad
550 faith interlocutors who ignore her first-person authority, *however* that is to be
551 explained, because PPSK ensures the general (and systematically superior)
552 reliability of her self-beliefs relative to Roman's beliefs about her attitudes.
553 PPSK is what gives Maya an epistemic warrant for holding steadfast against
554 his machinations, even if he was already unjustified in challenging her self-
555 ascription challenged her self-ascription in the first place.¹³ Moreover, PPSK
556 protects Maya against erroneous self-belief change even if Roman, innocently,
557 fails to recognize that her self-ascription expresses the very attitude that it is
558 about.

559 The final objection to my argument is that Maya can get away with merely
560 *assuming* that she generally has PPSK, such that she is generally entitled to
561 not defer to interlocutors who challenge her self-beliefs (whether in good
562 or bad faith). More substantively, it might be argued that Maya possesses a

12 See Doyle (2021) and Winokur (2022) for more precise articulations of these claims.

13 I take expressivism to have brighter prospects than Schwengerer (2021) does, though I also agree with him that not *everything* epistemically interesting about mental state discourse can be explained by first-order phenomena, hence the account given in this paper.

563 distinctively strong *practical* warrant for holding steadfast when faced with
 564 accusations of mistaken self-belief, even if she lacks a distinctively strong
 565 *epistemic* warrant for doing so.

566 The trouble with this objection is that it is hard to see what could ground
 567 Maya's practical warrant for holding steadfast if it is not really, at bottom, the
 568 same as (or itself grounded by) epistemic warrant for doing so. This is because a
 569 *merely* practical warrant here would go against her epistemic wellbeing in any
 570 number of cases. Specifically, if she did not systematically know herself better
 571 than others know her, then she *would* often—perhaps even typically—have
 572 an epistemic reason to discourse with others about whether her self-beliefs
 573 are true, and this would be in tension with her practical warrant for avoiding
 574 such discourse. In other words, it is only if Maya really has PPSK, thus having
 575 epistemic warrant for holding on to her self-beliefs, that holding steadfast
 576 against her interlocutors' countervailing assertions does not inadvertently
 577 prevent her from indulging many epistemically legitimate disagreements
 578 about what her attitudes are. It is only if she really has PPSK that *not* entering
 579 these disagreements is by and large good for her to do.

580 It might now be complained, relatedly, that I have merely established the
 581 importance of an especially strong epistemic warrant for our self-beliefs, but
 582 that this need not amount to PPSK. In other words, Maya might have an
 583 especially strong epistemic warrant for her self-beliefs, but these self-beliefs
 584 need not be especially reliably *true*.¹⁴ Indeed, such warrant may also suffice
 585 for avoiding the efficiency and gridlock concerns described in section 2. But
 586 I think a similar response applies here. For, if Maya has especially strong
 587 epistemic warrant for her self-beliefs but this warrant does not amount to
 588 self-knowledge in at least most of the cases in which she possesses this war-
 589 rant, then in any number of cases she will still miss out on an epistemic
 590 good—that of *true* warranted self-belief—whenever she declines to engage
 591 with interlocutors who claim that her self-beliefs are false. Moreover, it is
 592 hard to understand how she could possess this special epistemic warrant
 593 for her self-beliefs if she did not actually possess self-knowledge in most of
 594 those same cases. After all, this would be tantamount to having epistemic
 595 warrant for self-beliefs that were not correspondingly likely to be true, and

14 Thanks to an anonymous reviewer for this objection.

596 this systematic mismatch between truth and warrant would call the warrant
597 itself into question.¹⁵

594 **Interpersonal Reasoning and Epistemic Control**

599 In sections 2–3 I argued that **PPSK** provides us with widespread (even if not
600 universal) protection against situations in which others provide epistemic
601 reasons for us to change our self-beliefs, whether our interlocutors are operat-
602 ing innocently or in bad faith, and that this helps to ensure the efficiency and
603 reliability of interpersonal reasoning. In this section I show that my account
604 complements another recent account of **PPSK**'s instrumental value.

605 According to Jared Peterson, **PPSK** is instrumentally valuable because it
606 facilitates “epistemic control,” which is a matter of being able to “keep private
607 or disclose particular facts about one’s mind to others” (2021, 368). Take
608 privacy first. If you have **PPSK**, then you can reliably conceal your attitudes
609 from others. For example, a teacher might fail to motivate a student’s learning
610 if the student knows that the teacher is pessimistic about the student’s progress.
611 But if the teacher has **PPSK** of her pessimism, then she has greater epistemic
612 control over whether the student discovers this. Therefore, the teacher has
613 greater control over the student’s motivation to continue studying. For an
614 example about disclosing rather than concealing one’s mental states, Peterson
615 says that “[a]n estranged lover might want a former partner to know in a
616 highly epistemically secure manner that she still loves him” (2021, 369).

617 He also says that epistemic control:

618 [...] allows societies to function in a much more productive, orga-
619 nized, and amicable way. When we accomplish group objectives
620 in an efficient and peaceful manner we do so in large part by
621 keeping private that which would be counterproductive to the
622 group’s efforts, and/or revealing our thoughts, beliefs, desires, etc.
623 that are valuable for other members of a group to know. (Peterson
624 2021, 371)

625 Peterson and I are both happy to emphasize the social importance of **PPSK**.
626 I am also willing to say that **PPSK** provides a way to reliably disclose one’s

15 This response is similar in structure to one pursued by Davidson (1991) regarding perceptual belief warrant, though I believe that the strength of our warrants for perceptual beliefs and self-beliefs differ.

attitudes to others. However, Peterson does not also acknowledge the additional possibility of, e.g., *expressing one's love itself* as a reliable way of putting one's former partner in a secure position with respect to one's mind, where this expressive capacity may or may not depend on an agent's self-knowledge.

More significantly, though, I submit that PPSK's role as a shield against erroneous self-belief change is independent of its role in enabling one to disclose or conceal one's attitudes from others. To be able to better conceal one's attitudes is to prevent others from discovering what attitudes one has, but this may not matter to interlocutors who do not care (or are simply mistaken) about the facts and, instead, aim to convince you of a certain belief about yourself. Similarly, having an especially epistemic secure way of disclosing your attitudes is something that interlocutors could ignore (as argued in section 3 when discussing first-person authority). Thus, one may be tempted to deny that the instrumental value of PPSK for interpersonal reasoning is a matter of epistemic control.

However, one might be just as easily inclined to regard this as an instance of epistemic control after all, since my account claims that agents with systematically superior knowledge of their self-beliefs thereby exercise greater control over their social-epistemic lives as interpersonal reasoners. Accordingly, the lesson to be drawn may be that we ought to broaden our view of PPSK's contribution to epistemic control, such that epistemic control encompasses (i) control over which attitudes one discloses to others,¹⁶ (ii) control over which attitudes one conceals from others, *and* (iii) control over which attitudes one is able to self-attribute, with especially secure epistemic warrant, in the face of disagreement about one's attitudes, while attempting to reason with others.

5 Conclusion

I have argued that privileged and peculiar self-knowledge contributes to our capacity for interpersonal reasoning about the world around us. To the extent that agents possess PPSK of their attitudes, interpersonal reasoning is a more reliable route to discursively navigating our shared world, and this explains one way in which PPSK is instrumentally valuable.

For the record, I happen to believe that phenomena like epistemic hostage taking are not widespread, and I regard PPSK as at least a partial explanation of this fact. I take myself, therefore, to have contributed to the debate

¹⁶ Again, if this particular capacity requires PPSK at all.

661 between PPSK-skeptics and PPSK-realists not *merely* by illuminating the de-
 662 bate's stakes, but *also* by taking an anti-skeptical stand within that debate.
 663 This being said, I reiterate that others may not be persuaded to go as far as me
 664 in this regard, such that the core contribution of this paper is best viewed as
 665 an account of how being a PPSK-skeptic or PPSK-realist should affect one's
 666 corresponding conception of our interpersonal reasoning competencies.*

667 Benjamin Winokur

668  0000-0002-0845-9460

669 University of Macau

670 bwinokur@um.edu.mo

671 References

- 672 ANDREOTTA, Adam J. 2021. "Confabulation Does Not Undermine Introspection for
 673 Propositional Attitudes." *Synthese* 198(5): 4851–4872, doi:10.1007/s11229-019-
 674 02373-9.
- 675 —. 2022. "More than Just a Passing Cognitive Show: A Defence of Agentialism About
 676 Self-Knowledge." *Acta Analytica* 37(3): 353–373, doi:10.1007/s12136-021-00492-y.
- 677 BAR-ON, Dorit. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford:
 678 Oxford University Press, doi:10.1093/0199276285.001.0001.
- 679 BILGRAMI, Akeel. 2006. *Self-Knowledge and Resentment*. Cambridge, Massachusetts:
 680 Harvard University Press, doi:10.2307/j.ctv1nznfgcn.
- 681 BOYLE, Matthew. 2011. "Transparent Self-Knowledge." *Proceedings of the Aristotelian*
 682 *Society, Supplementary Volume* 85: 223–241, doi:10.1111/j.1467-8349.2011.00204.x.
- 683 —. 2019. "Transparency and Reflection." *Canadian Journal of Philosophy* 49(7):
 684 1012–1039, doi:10.1080/00455091.2019.1565621.
- 685 BRANDOM, Robert B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive*
 686 *Commitment*. Cambridge, Massachusetts: Harvard University Press.
- 687 BURGE, Tyler. 1996. "Our Entitlement to Self-Knowledge." *Proceedings of the Aris-*
 688 *totelian Society* 96: 91–116. Reprinted in Burge (2013, 68–87), doi:10.1093/aristote-
 689 lian/96.1.91.

* Thanks are due to Adam Andreotta for his comments on an earlier version of this paper, and to Rachel Cooper for reading and commenting on a shortened version of it at the 2021 American Philosophical Association Pacific Meeting. Thanks are also due to the other audience members at that meeting, to audience members at the 2021 Canadian Philosophical Association Meeting, and to audience members at the 2022 Meeting of the International Social Ontology Society. Finally, special thanks are due to three anonymous reviewers at *Dialectica* for their many useful comments.

- 690 —. 2013. *Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning,*
 691 *Reflection. Philosophical Essays, Volume 3.* Philosophical Essays n. 3. Oxford: Oxford
 692 University Press, doi:10.1093/acprof:oso/9780199672028.001.0001.
- 693 BYRNE, Alex. 2018. *Transparency and Self-Knowledge.* Oxford: Oxford University Press,
 694 doi:10.1093/oso/9780198821618.001.0001.
- 695 CAMPBELL, Lucy. 2018. “Self-Knowledge, Belief, Ability (and Agency?).” *Philosophical*
 696 *Explorations* 21(3): 333–349, doi:10.1080/13869795.2018.1426779.
- 697 CARRUTHERS, Peter. 2011. *The Opacity of Mind: An Integrative*
 698 *Theory of Self-Knowledge.* Oxford: Oxford University Press,
 699 doi:10.1093/acprof:oso/9780199596195.001.0001.
- 700 CASSAM, Quassim. 2014. *Self-Knowledge for Humans.* Oxford: Oxford University Press,
 701 doi:10.1093/acprof:oso/9780199657575.001.0001.
- 702 COLIVA, Annalisa. 2016. *The Varieties of Self-Knowledge.* Innovations in Philosophy.
 703 London: Palgrave Macmillan, doi:10.1057/978-1-137-32613-3.
- 704 DAVIDSON, Donald. 1991. “Epistemology Externalized.” *Dialectica* 45(2/3): 191–202.
 705 Reprinted in Davidson (2001), doi:10.1111/j.1746-8361.1991.tb00986.x.
- 706 —. 2001. *Subjective, Intersubjective, Objective: Philosophical Essays, Volume 3.* Oxford:
 707 Oxford University Press, doi:10.1093/0198237537.001.0001.
- 708 DOYLE, Casey. 2021. “There’s Something About Authority.” *Journal of Philosophical*
 709 *Research* 46: 363–374, doi:10.5840/jpr2021816169.
- 710 FRICKER, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford:
 711 Oxford University Press, doi:10.1093/acprof:oso/9780198237907.001.0001.
- 712 GOPNIK, Alison. 1993. “How We Know Our Minds: The Illusions of First-Person
 713 Knowledge of Intentionality.” *Behavioral and Brain Sciences* 16(1): 1–14, doi:10.1
 714 017/S0140525X00028636.
- 715 GRECO, John. 2019. “The Social Value of Reflection.” in *Thinking about Oneself: The*
 716 *Place and Value of Reflection in Philosophy and Psychology*, edited by Waldomiro
 717 J. SILVA-FILHO and Luca TATEO, pp. 45–57. Philosophical Studies Series n. 141.
 718 Cham: Springer, doi:10.1007/978-3-030-18266-3_4.
- 719 KEELING, Sophie. 2018. “Confabulation and Rational Obligations for Self-Knowledge.”
 720 *Philosophical Psychology* 31(8): 1215–1238, doi:10.1080/09515089.2018.1484086.
- 721 —. 2019a. “The Transparency Method and Knowing Our Reasons.” *Analysis* 79(4):
 722 613–621, doi:10.1093/analys/anz031.
- 723 —. 2019b. “Knowing Our Reasons: Distinctive Self-Knowledge of Why We Hold Our
 724 Attitudes and Perform Actions.” *Philosophy and Phenomenological Research* 102(2):
 725 318–341, doi:10.1111/phpr.12655.
- 726 KOREŇ, Ladislav. 2023. “Have Mercier and Sperber Untied the Knot of Human Rea-
 727 soning?” *Inquiry* 66(5): 849–862, doi:10.1080/0020174X.2019.1684988.
- 728 LEVINE, Steven M. 2009. “Expressivism and I-Beliefs in Brandom’s *Making it Explicit.*”
 729 *International Journal of Philosophical Studies* 17(1): 95–114, doi:10.1080/096725
 730 50802614786.

- 731 MARCUS, Eric and SCHWENKLER, John. 2019. "Assertion and Transparent
732 Self-Knowledge." *Canadian Journal of Philosophy* 49(7): 873–889,
733 doi:10.1080/00455091.2018.1519771.
- 734 MÜLLER, Andreas. 2019. "Reasoning and Normative Beliefs: Not Too Sophisticated."
735 *Philosophical Explorations* 22(1): 2–15, doi:10.1080/13869795.2018.1523449.
- 736 MUSHOLT, Kristina. 2015. *Thinking about Oneself: From Nonconceptual Content to the*
737 *Concept of a Self*. Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpre
738 ss/9780262029209.001.0001.
- 739 NGUYEN, A. Minh. 2015. "What Good is Self-Knowledge?" *Journal of Philosophical*
740 *Research* 40: 137–154, doi:10.5840/jpr2015111656.
- 741 PARENT, Ted. 2017. *Self-Reflection for the Opaque Mind: An Essay in Neo-Sellarsian*
742 *Philosophy*. London: Routledge, doi:10.4324/9781315618449.
- 743 PETERSON, Jared. 2021. "The Value of Privileged Access." *European Journal of Philo-*
744 *sophy* 29(2): 365–378, doi:10.1111/ejop.12594.
- 745 ROELOFS, Luke. 2017. "Rational Agency without Self-Knowledge: Could 'We' Replace
746 'I'?" *Dialectica* 71(1): 3–33, doi:10.1111/1746-8361.12169.
- 747 SCHWENGERER, Lukas. 2021. "Beliefs Over Avowals: Setting Up the Discourse on
748 Self-Knowledge." *Episteme* 18(1): 66–81, doi:10.1017/epi.2018.56.
- 749 SHOEMAKER, Sydney S. 1988. "On Knowing One's Own Mind." in *Philosophical Per-*
750 *spectives 2: Epistemology*, edited by James E. TOMBERLIN, pp. 183–209. Atascadero,
751 California: Ridgeview Publishing Co. Reprinted in Shoemaker (1996, 25–49),
752 doi:10.2307/2214074.
- 753 —. 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge Uni-
754 versity Press, doi:10.1017/cbo9780511624674.
- 755 SIEWERT, Charles. 2003. "Self-Knowledge and Rationality: Shoemaker on
756 Self-Blindness." in *Privileged Access: Philosophical Accounts of Self-Knowledge*,
757 edited by Brie GERTLER. Ashgate Epistemology and Mind Series. London:
758 Routledge, doi:10.4324/9781315245997.
- 759 SORGIOVANNI, Ben. 2019. "The Agential Point of View." *Pacific Philosophical Quarterly*
760 100(2): 549–572, doi:10.1111/papq.12263.
- 761 STRIJBOS, Derek W. and DE BRUIN, Leon C. 2012. "Making Folk Psychology Explicit:
762 The Relevance of Robert Brandom's Philosophy for the Debate on Social Cognition."
763 *Philosophia* 40(1): 139–163, doi:10.1007/s11406-010-9288-z.
- 764 VALARIS, Markos. 2018. "Self-Knowledge." in *The Philosophy of Knowledge: A History,*
765 *Volume 4: Knowledge in Contemporary Philosophy*, edited by Stephen Cade HET-
766 HERINGTON and Markos VALARIS, pp. 155–174. London: Bloomsbury Academic,
767 doi:10.5040/9781474258814.ch-008.
- 768 WINOKUR, Benjamin. 2021a. "Critical Reasoning and the Inferential Transparency
769 Method." *Res Philosophica* 98(1): 23–42, doi:10.11612/resphil.1967.
- 770 —. 2021b. "Inference and Self-Knowledge." *Logos & Episteme* 12(1): 77–98, doi:10.584
771 0/logos-episteme20211214.

- 772 —. 2022. “There is Something to the Authority Thesis.” *Journal of Philosophical Re-*
773 *search* 47: 115–132, doi:[10.5840/jpr202298189](https://doi.org/10.5840/jpr202298189).

PROOF