



# Authority and Attribution: the Case of Epistemic Injustice in Self-Knowledge

Cristina Borgoni<sup>1</sup>

Received: 18 June 2018 / Accepted: 18 June 2018 /  
Published online: 27 July 2018  
© Springer Nature B.V. 2018

## 1 Introduction

Philosophy seems to progress very slowly and sometimes due to small steps that reveal blatant truths that nobody has noticed before. Coliva's *The Varieties of Self-Knowledge* stands as one such an advance in our investigations on self-knowledge. The central idea of the book—namely: given that our mental states are intrinsically different, a unitary account of self-knowledge is unlikely to succeed—is so convincing that it is hard to finish reading the book without being a pluralist about self-knowledge. Coliva's pluralist proposal sharply contrasts with a tradition centered on unitary explanations of self-knowledge.<sup>1</sup>

Besides the well-grounded defense and development of a proper pluralist view on self-knowledge, Coliva's book has other elements to praise. Coliva does an extraordinary job of including nearly all relevant accounts of self-knowledge, whereas it is not uncommon to find serious omissions in other attempts to characterize the terrain.

Another element of the book worth praising is the chapter on the variety of mental states (Chapter 2), which is also unusual in a book on self-knowledge. The deep relations between issues in philosophy of mind and issues on self-knowledge are then vindicated by Coliva's book. Of course, Coliva does not give—much less aims at giving—a treatise on each and every relevant mental state known in self-knowledge. With the exception of emotions—about which Coliva proposes her own view—she relies on other authors' accounts to map her view of the mental.

This paper, however, does not focus on the points of agreement with Coliva's treatment of self-knowledge. Rather, this paper raises some concerns regarding one of the pillars of her substantial account of first-personal self-knowledge: the notion of

---

<sup>1</sup>Many unitary accounts on self-knowledge recognize the limits of their proposals and in this way could be considered to be pluralists. Some accounts promise to apply only to beliefs, or more generally, to propositional attitudes. Others promise to explain only phenomenal states and so on. However, in contrast with Coliva's proposal, there is no real recognition that a proper account of self-knowledge demands a pluralist explanation.

✉ Cristina Borgoni  
cborgoni@gmail.com

<sup>1</sup> Department of Philosophy, University of Bayreuth, Bayreuth, Germany

authority. Against her view that authority is a necessary and apriori feature of self-knowledge, I argue that there are cases where the individual's self-knowledge is retained while her authority is undermined. Such counter-examples derive from cases of epistemic injustice.

It is constitutive of the phenomenon of epistemic injustice that others, wrongly and unjustly, raise doubts about the knower's assertions. Since epistemic injustice can also target mental self-ascriptions, doubts about such assertions are also possible. This fact highlights the attributional aspect of authority, which supports the idea that authority stands rather as a contingent and non-apriori aspect of self-knowledge. For that conclusion, I first explain Coliva's notion of authority and then propose cases of epistemic injustice regarding mental self-ascriptions. Finally, I reply to some possible responses to my argument and explore its consequences for constitutivism.

## 2 Authority

Coliva argues that authority is, along with groundlessness and transparency, one of the three central features of truly first-personal knowledge. In a nutshell, the three features amount to the following aspects: the notion of groundlessness refers to the idea that "self-knowledge is not the result of any substantial cognitive achievement, such as observing or inferring from a symptom to its likely cause" (Coliva 2016: 54); the notion of transparency—which differs from the extensively discussed notion found on Evans' works—refers to the idea that the occurrence of some of our mental states "are of a piece with one's awareness of them" (Coliva 2016: 57); finally, the notion of authority amounts to the idea that nobody can rationally cast doubts on your sincere avowals about your mental states, provided that you are competent in the use of the employed concepts (Coliva 2016: 62). The three notions are the topic of chapter 3 of Coliva's book, where she presents arguments for their necessity and apriority as aspects of first-personal self-knowledge. She holds that any good approach to first-personal self-knowledge should be able to account for such features, be it in their weak or strong formulations. Let us focus on the feature of authority.

Coliva acknowledges that many philosophers have recently raised doubts regarding authority as a necessary feature of self-knowledge. Snowdon (2012) has doubted that we are authoritative regarding our bodily sensations—we can envisage situations where we are uncertain about the location of a given ongoing sensation and thus, doubts about them are rationally welcome. Schwitzgebel (2008) has argued that we are very bad at knowing our inner feelings and emotions, and thus cannot be regarded as authoritative in their self-ascriptions. Additionally, Coliva enumerates a number of other examples of mental states on which we are arguably not authorities—e.g., confused and vague sensations; perceptual appearances, etc.

However, in relation to such supposed counter-examples to authority, Coliva concludes—correctly to my eyes—that such cases only show either that we are not authoritative in relation to certain aspects of those mental states (e.g., the location of a given sensation versus its very occurrence) or that the conditions for authority need to be better specified (such as being cognitively lucid versus having one's cognitive system impaired due to, for example, intense pain). More importantly, certain supposed counter-examples actually fit perfectly with Coliva's account. Dispositional mental

states are not known in a truly first-personal way due to their nature, and thus, it is an expected result that we are not authoritative regarding them. This is her line of response to cases of self-deception, which according to Coliva are “the most powerful counter-examples to authority” (Coliva 2016: 65).

After such an initial discussion of possible counter-examples to authority, Coliva proposes the following formulations of the notion of authority:

**Authority<sub>weak</sub>:** Given C-conditions (including concepts’ possession, cognitive well-functioning, alertness and attentiveness), if one judges to have a mental state M (save for dispositional ones or for the dispositional elements of some mental states), one will usually have it.

**Authority<sub>strong</sub>:** Given C-conditions (including concepts’ possession, cognitive well-functioning, alertness and attentiveness), if one judges to have a mental state M (save for dispositional ones or for the dispositional elements of some mental states), one will always have it. (Coliva 2016: 65)

Interestingly, the above formulations of the weak and strong notions of authority contrast in important ways with the intuitive and shared notion of authority that opens the subsection 1.3, namely, the notion that doubts about one’s competent and sincere avowals about a given mental state are unreasonable and generally misplaced. The proposed formulations seem rather to establish conditions under which mistakes occur rarely or are impossible. Perhaps authority—as the inadequacy of doubts—could be grounded in the unlikelihood of mistakes in self-knowledge. However, the formulations above do not directly establish what it means for one’s avowals to be *closed* to doubts. Rather, they indicate that avowals are not open to mistakes. Doubts about the person’s avowal could be raised—and thus her authority be brought into question—even though the formulations above were fulfilled. More specifically, one’s authority might be undermined even though one’s first-personal knowledge remains intact. Or so I will argue in the next section. In what follows, I will be employing the notion of authority in its intuitive version, as the feature of avowals of being closed to doubts.

### 3 Epistemic Injustice and One’s Status as Self-Knower

Recall the opening example of Fricker’s *Epistemic Injustice*:

In Anthony Minghella’s screenplay of *The Talented Mr Ripley*, Herbert Greenleaf uses a familiar put-down to silence Marge Sherwood, the young woman who, but for the sinister disappearance of his son, Dickie, was soon to have become his daughter-in-law: ‘Marge, there’s female intuition, and then there are facts.’ Greenleaf is responding to Marge’s expressed suspicion that Tom Ripley—a supposed friend of Dickie and Marge, who has curried much favour with Greenleaf senior—is in fact Dickie’s murderer (Fricker 2007, p. 9).

The example introduces the phenomenon—perhaps too well-known to some of us—of epistemic injustice, which is identified and analyzed by Miranda Fricker in her monograph. The phenomenon is one in which the person is wronged in her capacity

as a knower due to prejudice and thus, transmission of knowledge via testimony is impaired in such cases.<sup>2</sup> Testimonial injustice involves epistemic and ethical mistakes. In the example above, Marge is wronged in her capacity to figure out who the murderer is, due to sexism, and is mistakenly judged as forming the wrong judgment about it. She knows whom the murderer is, but she is not listened to because, according to sexist ideals, a woman's testimony is unreliable.

The example above concerns first-order propositions that amount to knowledge—i.e., “Tom Ripley is the murderer”—that are nevertheless incorrectly taken not to be a piece of knowledge due to a very specific phenomenon of injustice, namely epistemic injustice. However, it is not difficult to find parallel cases where second-order propositions involving mental concepts are the target of epistemic injustice. Consider the following case:

Imagine yourself being a woman (or someone belonging to a gender, race, or other target group of prejudice). You tell your colleague, after having considered a candidate's application portfolio for an open position in your department, “I disagree with your proposal. I believe that this is the person we should hire”. Your colleague, dismissing what you have just said and appealing to prejudicial ideas replies, “No, you don't believe it. Women just like to fuss with men's decisions.” In this imaginary case, your colleague is bringing doubts about the belief that you have just formed based on considered reasons. This type of mental state is a typical example of a ‘committal propositional state’ in Coliva's terminology, and exemplifies one of the most basic states you know in self-knowledge.

If this case is possible, it is a case where first-person authority—understood as being closed to doubts—is challenged. Your self-knowledge is retained—you in fact know what you believe. However, doubts about your mental self-ascriptions are raised. Your colleague suggests that you have ascribed the wrong mental state to yourself, i.e., that you don't have a belief about his proposal being bad, but rather a desire to fuss with him—and in case you do have a belief, it is a false one. Your own authority over your mind is challenged in such prejudicial scenarios.

Recall that, because this is a case of epistemic injustice, your colleague is not merely moved by subjective factors, such as his emotional attachment to his decision, or his tendency to respond negatively to criticisms of his actions. His response to your avowal is mainly motivated by the sincere but wrong idea that women (or members of other stigmatized groups depending on your chosen scenario) have poor judgment, including judgment about their own minds.

Let us now consider a more extreme case of epistemic injustice also involving mental self-ascriptions. Imagine an enslaved person speaking his mind in a slave-owning society. If his words can be said at all, they would very likely be dismissed, including this person's avowals about his wants and beliefs. Imagine this person saying that he wants to have his own home and to be free. Or that he believed it was wrong to be treated in the way he was. It is not difficult to picture a scenario where such expressions would be taken to be expressions of self-ignorance. Prejudice does not distinguish between levels of predicates.

<sup>2</sup> Fricker analyzes different forms of epistemic injustice in her book. In this paper, I'm focusing on cases of testimonial injustice.

Just as it happens in the first imaginary case, the individual belonging to a slave-owning society knows his own mind as well as you do. However, his first-person authority is systematically undermined. In a slave-owning society, doubts about slaves' avowals are not considered unreasonable. On the contrary, such doubts are part of the power relations that perpetuates such injustices. Such doubts are ethically and epistemically wrong, but they are spread out.

What I want to suggest is that these cases show that having first-personal knowledge cannot guarantee first-person authority, for the latter has necessarily an attributional element. In a slave-owning society, slaves are likely not recognized as authorities over their own minds. Similarly, women are not recognized as authorities over their minds in a sexist society. In such scenarios, there is no presumption of truth in what the target people of prejudice aver (pace Davidson 1984), nor it is the case that doubts about mental self-ascriptions are misplaced (pace Wright 1998). If there is a general and systematic failure in recognizing someone as an authority on her mind, she is thereby not an authority. However, lacking authority does not imply lacking self-knowledge. In the proposed cases, women and slaves know their minds. They are wronged in their status as self-knowers.

Let us take stock and summarize the claims made so far. As a working strategy, I have focused on the notion of authority as the idea that doubts about your competent and sincere avowals are unreasonable and misplaced; “avowals are closed to doubts”. Such an idea is found in Coliva's book and is shared by many other philosophers,<sup>3</sup> although it contrasts with her semi-technical formulation of the notion in chapter 3, to which we will soon return. I presented a couple of cases where such doubts are not unreasonable or misplaced. In scenarios where epistemic injustice takes place, doubts about your competent and sincere avowals are to be expected—even though they are epistemically and ethically wrong. In extreme prejudicial scenarios, one's authority can be totally undermined due to the lack of recognition of such an authority by one's peers. This result suggests that first-person authority necessarily has an attributional element. In contrast, first-personal knowledge does not have such an attributional element. In the proposed scenarios, the individual does have self-knowledge even if he or she is systematically denied such a status. Self-knowledge is, in principle, independent of attributional mistakes.<sup>4</sup>

One consequence of these results for Coliva's work is that the thesis that authority is a necessary and apriori feature of self-knowledge is shown to be hard to sustain. In contrast with common attacks on the notion of authority, the proposed cases of epistemic injustice keep the individuals' self-knowledge while undermining their authority, which thus supports the idea that authority is not a necessary feature of self-knowledge. In this sense, Coliva's argument that systematic failure of self-knowledge leads either to the conclusion that the subject lacks the

<sup>3</sup> Other formulations of the notion of authority are as follows: (i) there is usually a presumption that the speaker is not mistaken when she avers her beliefs, hopes, desires, intentions or feelings (Davidson 1984); and (ii) questions about the speaker's grounds to her mental self-ascriptions are normally misplaced (Wright 1998).

<sup>4</sup> At this point, one could raise a related question coming from developmental psychology, namely, whether one could develop abilities for self-knowledge if her authority were massively denied from early stages in one's life. Although interesting and important, this paper does not intend to answer this question. This paper disputes an apriori thesis about self-knowledge, namely whether authority is a necessary and apriori aspect of it.

relevant concepts or that her rationality is impaired does not concern authority, but rather self-knowledge itself.

#### 4 Possible Replies

One quick response by Coliva would be that the previous argument is constructed upon a different notion of authority than the one she offers in her book. Broadly speaking, Coliva proposes that authority holds when the following condition is fulfilled: if one judges to have a mental state *M*, one will usually/always have it. Such a condition is fulfilled in cases of epistemic injustice regarding one's own status of self-knower. Thus, the proposed cases are not counter-examples to authority.

This is a fair reply. However, allow me to insist, Coliva's proposed notion of authority is presented as being derived from the idea that doubts about avowals are unreasonable and normally misplaced. The proposed cases are counter-examples to such a notion. More importantly, Coliva's technical notion of authority identifies the conditions where mistakes are very uncommon (in the weak formulation of authority) or where mistakes are not possible (in the strong formulation of authority). However, the conditions under which mistakes are not possible, regarding cases of self-knowledge, are different from the conditions under which doubts can or cannot be raised. Cases of epistemic injustice are cases in which mistakes in self-knowledge are unlikely to occur whereas doubts about their expressions are expected. The discussed cases show that the conflation between the two notions—conditions where mistakes are not possible and where doubts cannot be raised—is mistaken. Ultimately, Coliva's discussion on first-person authority includes such a faulty conflation.

It would be unfair, however, to attribute this faulty conflation to Coliva alone. The two notions are significantly conflated in the literature on self-knowledge. It is, so to speak, a shared misconception about the notion of authority. A lot of the literature on self-knowledge has actually tried to explain self-knowledge via the explanation of the trait of first-person authority. The most common counter-examples to authority (see Section 2) are rather counter-examples to self-knowledge properly speaking, since they envisage situations where the individual is mistaken in their apprehension of their own minds. Thus, it is not a surprise that this unfortunate element of the tradition is preserved in Coliva's book.

Besides supporting the idea that the mentioned conflation is mistaken, cases of epistemic injustice show another problematic misconception about authority in our studies on self-knowledge. The misconception relates to an idealized picture of self-knowledge, where avowals about one's mind are made in a 'sterilized environment', where no social factors enter the scene. The supposition that doubts about people's avowals are obviously misplaced (including states that correspond to truly first-personal self-knowledge) is just false. But the supposition is false not because we can commit mistakes regarding the apprehension of such basic mental states. Rather, the supposition is false because authority is not granted unless our peers recognize it in our words.

In this sense, Coliva's view on authority as a central feature of self-knowledge is found wanting. It is not the case that the lack of first-person authority implies lack of rationality or lack of concepts. Cases of epistemic injustice are cases where one's

rationality, mastering of concepts, and self-knowledge are maintained while one's authority is undermined.

However, her discussion on authority does pick out something correct about self-knowledge: regarding some types of mental states, mistakes are not permitted on pain of the individuals' irrationality or lack of concepts. What I'm claiming is that this conclusion does not hold for the notion of authority as ordinarily understood. Nevertheless, the conclusion does hold for self-knowledge, and thus, for Coliva's strict definition of authority (in its weak and strong versions). In a way, I have argued that it is a mistake to call the feature picked out under such definitions 'authority'. The feature that is correctly picked out corresponds to the idea that first-personal self-knowledge does not accommodate systematic mistakes. I agree with Coliva that such a feature is necessary and *apriori* because:

For, if one were systematically proven wrong in one's psychological self-ascriptions, doubt would be cast upon one's possession of the relevant concepts. Alternatively, if one wished to maintain that a constantly mistaken subject (with respect to her own mental states) could still be said to have the relevant psychological concepts, perhaps because she retains the ability to apply them to other people's mental states, her persistent self-deception would impair her rationality. For she would avow certain mental states while she would systematically behave in ways which run contrary to them. Hence, we could no longer make sense of her linguistic and non-linguistic behaviour. (Coliva 2016: 65–66)

My argument is not, nevertheless, based on a mere terminological dispute. Calling such a feature 'authority' is part of a tradition that has ignored key aspects of the discussion on self-knowledge, for example, situations in which self-knowers are wronged in their epistemic status as such. Since our assessment of the phenomenon of self-knowledge is partly via our language practices, and such language practices include cases of epistemic injustice, we should acknowledge their implications. In extreme cases in which a self-knower is denied self-knowledge, he or she cannot be regarded as an authority on his or her mind. In order to be regarded as an authority on a given subject, one needs one's peers' recognition of such status. Lacking one's peers' recognition is compatible with, as a matter of fact, being a self-knower.

## 5 Authority and Constitutivism

Constitutivism plays an important role in Coliva's pluralist account of self-knowledge. Coliva holds a particular type of constitutivism—what she calls a “metaphysically robust constitutivism”—for a particular set of mental states—the committal propositional states. This final section analyses, in an exploratory way, some related concerns for constitutivism given the arguments raised in this essay.

It is at the heart of constitutivism to conceive self-knowledge in a non-epistemic way. Coliva reaches chapter 7, where she analyzes several types of constitutivism and proposes her own version, after having argued against epistemic treatments of self-knowledge. Despite the big differences between its varieties, constitutivism holds what Coliva identifies as the no-knowledge thesis, according to which:



the so-called self-knowledge is not a kind of cognitive achievement after all and it is somehow a misnomer to call it “knowledge” if knowledge is understood as the result of a, however minimal, cognitive endeavor (...) Rather, what we call “self-knowledge”—that is, the distinctive kind of authority we recognise in our fellow humans and in ourselves over our mental states as well as the distinctively groundless and transparent way in which we are aware of them—is guaranteed to hold a priori, as a matter of conceptual necessity. (Coliva 2016: 163).

The first concern derived from the arguments discussed in previous sections relates to such a thesis.<sup>5</sup> The cases of epistemic injustice targeting avowals suggest that there are two notions, mistakenly conflated under the notion of authority, that are employed in the discussions on self-knowledge: the impossibility of mistakes in self-knowledge and the impossibility of reasonable doubts regarding avowals. However, given the no-knowledge thesis, talking in terms of mistakes and their possibility in self-knowledge seems insignificant. The notion left for constitutivism to employ and explain is, thus, the notion of impossibility of reasonable doubts. However, if this is so, constitutivism seems to have a hard time explaining cases where the two notions split, since one of them has no significance in the model. In cases of epistemic injustice towards avowals, there are no mistakes in self-knowledge while doubts can be reasonably raised.

In fact, there is a sense in which the mistaken conflation between the two indicated notions works in favor of the constitutivist account. By conflating the two notions, one has the impression that by explaining authority one explains self-knowledge. But, if the previous arguments stand, constitutivism seems to have little resources to explain why the impossibility of mistakes holds in cases of first-personal self-knowledge. For that reason, the cases of epistemic injustice discussed in Section 3 may have a rather strong impact on the model.

However, the constitutivist could insist that, it does explain why mistakes cannot occur in avowals, and that explanation corresponds to the fact that the individual is the one who constitutes his or her mental states by avowing them. The constitutivist could insist that this fact also differs from the related feature of the impossibility of doubts regarding such avowals. This line of reply connects to a second concern about constitutivism given the cases discussed in this paper. This concern is about the underlying motivations of constitutivism as mentioned in the passage below:

These theories therefore inherit an aspect of Wittgenstein’s position [...] according to which, if it does not make sense to question one’s psychological self-ascriptions, at least in the normal run of cases, and thus there is no room for not knowing them, it means that they are not known either. It simply means that the kind of security they enjoy does not depend on the subject’s being in an epistemologically privileged position, which makes her right about her own mental states. Rather, it belongs to the “grammar”—that is to say, to the rules—of

<sup>5</sup> Holding the no-knowledge thesis does not make a position constitutivist. A constitutivist should hold additionally a positive view about self-knowledge. According to Coliva, all constitutivist theories agree on the following thesis as being a conceptual truth: “Constitutive thesis: Given C, one believes/desires/intends that P/to  $\varphi$  iff one believes (or judges) that one believes/desires/intends that P/to  $\varphi$ .” (Coliva 2016: 164)



our language game of making psychological self-ascriptions, that subjects are accorded authority over their own mental states. (Coliva 2016: 164)

Independently of whether explaining self-knowledge requires grasping the grammar of how we use psychological self-ascriptions, Wittgenstein and followers err in holding that “it does not make sense to question one’s psychological self-ascriptions, at least in the normal run of cases”. Cases of epistemic injustice are, unfortunately, part of the normal run of cases. It might be wrong and unjust to question one’s psychological self-ascriptions. But it happens and makes sense in unjust scenarios. Thus, the line of reasoning that says there is no room for not knowing our minds in avowing them *because* there is no room for others to doubt them is mistaken. Again, the conditions for the absence of reasonable doubts are different from the conditions for the absence of mistakes in self-knowledge. And in this matter, epistemic accounts may have a better chance to succeed.

To conclude, constitutivism (and other non-epistemic accounts of self-knowledge) seem to have a hard time accounting for the distinction between first-person authority and self-knowledge itself if they insist that our linguistic practices regarding avowals (at least of committal propositional states) block reasonable doubts. However, it remains an open question for future investigation whether a non-epistemic account of self-knowledge can handle the fact that we may have secure self-knowledge while being systematically denied first-person authority.

## References

- Coliva, A. (2016). *The varieties of self-knowledge*. London: Palgrave Macmillan.
- Davidson, D. (1984). First person authority. In D. Davidson, 2001, *Subjective, intersubjective, objective* (pp. 3–14). Oxford: OUP.
- Fricke, M. (2007). *Epistemic injustice: Power and ethics of knowing*. Oxford: OUP.
- Schwitzgebel, E. (2008). The unreliability of naïve introspection. *Philosophical Review*, 117, 245–273.
- Snowdon, P. (2012). How to think about phenomenal self-knowledge. In A. Coliva (Ed.), *The self and self-knowledge* (pp. 243–262). Oxford: OUP.
- Wright, C. (1998). Self-knowledge: The Wittgensteinian legacy. In C. Wright, B. Smith, & C. MacDonald (Eds.), 1998, *Knowing our own minds* (pp. 13–45). Oxford: Clarendon Press.