

Book review: COLIVA, Annalisa. *The Varieties of Self-Knowledge*. London: Palgrave Macmillan, 2016. 288p.

Alexandre de Borba

Federal University of Santa Maria
Department of Philosophy
Santa Maria, RS
Brazil
azdeborda@gmail.com

Article info

CDD: 121

Received: 07.09.2017; Accepted: 08.10.2017

DOI: <http://dx.doi.org/10.1590/0100-6045.2017.V40N3.AB>

Keywords:

Groundlessness

Transparency

Authority

First-personal Self-knowledge

ABSTRACT

In this review, I argue that Coliva's strong transparency admits three possible interpretations: (a) a Wittgensteinian interpretation, (b) a metaphysical interpretation, and (c) an epistemic interpretation. I favor (c) over (a) and (b).

Annalisa Coliva's recent monograph, *The Varieties of Self-Knowledge*, is, if not the major, one of the greatest contributions of the decade to the subject of self-knowledge in philosophy. In it, the philosopher defends the original thesis that the acquisition of knowledge about our own mental states admits a plurality of methods. In this review, I will focus on an idea that Coliva brings up in her work and explores in detail. More precisely, I will focus here on the criteria that, according to Coliva, demarcate the territory of the so-called "first-personal self-knowledge", as opposed to the third-personal self-knowledge—i.e., those instances of knowledge about oneself acquired by publicly accessible methods. According to Coliva, transparency, authority and groundlessness are necessary and *a priori* aspects of first-personal self-knowledge (COLIVA, 2016, p. 6). Before we proceed, let us make a general overview of the work.

The first chapter is an introductory chapter, in which the innovative thesis of the work is stated and there is a brief exposition of the content that we will have throughout the reading. The second chapter is dedicated to demarcate the territory of mental states, from sensations and perceptions, passing through the so-called “propositional” attitudes and ending in the emotions. Coliva provides us a geography of the mental states in which she will operate successfully. She sets aside, however, states of the mind such as moods and character traits—and this is, I think, justifiable in the context of philosophy because of the lack of a deeper literature on the subject. One of the major merits of this chapter is the fact that Coliva draws the distinction, widely retracted throughout the book, between propositional attitudes as *dispositions* and as *commitments*. With this distinction in mind, Coliva succeeds, already in chapter 7, following the philosopher Akeel Bilgrami, illuminating the phenomenon of self-deception, as well as, in the appendix of the work, illuminating what is *possible* and what is *paradoxical* in Moore’s paradox. It is, therefore, one of the most important distinctions outlined in the monograph.

The third chapter discusses the varieties of self-knowledge, distinguishing between two kinds of self-knowledge, namely, the first-personal self-knowledge and the third-personal one. It is here that Coliva demarcates the territory of first-personal self-knowledge as necessarily involving groundlessness, transparency, and authority. Coliva is convincing in proposing that the knowledge we have of our own mental states, being theses as varied as the previous chapter pointed out, I say, that self-knowledge admits different methodologies. Some cases of self-knowledge are acquired by inference, or observation, or any other method involving some minimal epistemic effort. However, Coliva proposes that cases of first-personal self-knowledge are not necessarily instances of cognitive achievement. Yet, this proposal forces us to take the theories that Coliva calls “epistemically robust accounts” of first-personal self-knowledge as false. In this way, Coliva seems to reject already in advance the accuracy of theories such as the inner sense theory, the inferentialist theory, and the simulation theory, which are explored in the fourth chapter.

The fourth chapter is dedicated precisely to the epistemically robust accounts of first-personal self-knowledge, in which the inner sense, the inferentialist, and the simulation theories are presented. These theories state that the knowledge we have about our own mental states can be acquired by introspection, observation, inference to the best explanation, simulation, etc.

The fifth chapter deals with the so-called epistemically weak accounts of first-personal self-knowledge, which include theories such as Peacocke's rational internalism, Burge's rational externalism, and Evans's transparency method, later developed in different ways by Fernández and Moran.

The sixth chapter presents the so-called expressivism, the result of some interpretations of Wittgenstein's work. It is in expressivism that for the first time we see the statement that first-personal self-knowledge is not exactly a kind of knowledge since it does not fit the Wittgensteinian criteria of knowledge self-ascription¹. The seventh chapter presents the so-called constitutivism, whose main representatives are Shoemaker, Wright, Bilgrami, and Coliva herself. As in expressivism, constitutivists declare that first-personal self-knowledge is not exactly the result of a *sui generis* epistemic achievement, and since it is not based on anything, we should conclude that to call it "knowledge" is a misunderstanding (p. 163). In contrast to expressivism, however, constitutivism appeals to metaphysical theses².

Finally, the eighth chapter is the chapter in which Coliva exposes pluralism about self-knowledge. According to her, the limits of constitutivism involve the scope of propositional attitudes as commitments. Concerning the basic emotions, sensations, and perceptions, Coliva promotes a meeting between constitutivism and expressivism. Finally, the knowledge we have of our complex emotions and our propositional attitudes as dispositions are genuine cases of knowledge that we obtain by methods that are publicly accessible—Coliva also includes in this scope the knowledge we acquire about our own personality³. The appendix deals with Moore's paradox.

The point I want to focus on in this review is the criteria that Coliva presents to demarcate the territory of first-personal self-knowledge. They are three: groundlessness, transparency, and authority. Both admit a weak and another strong variant. In a weak characterization, groundlessness admits some epistemic ground, although it dispenses the foundation as being of an observational or inferential kind. The so-called epistemically weak accounts, such as rational internalism and externalism, of Peacocke and Burge respectively, or even Evans's transparency method, such as developed by Moran, satisfy this criterion. It was not clear, however, how the transparency

¹ See p. 140.

² See p. 164.

³ See p. 239.

method as developed by Fernández satisfies weak groundlessness, since, in Fernández's account, a self-ascription of, let's say, a belief, is based on the same evidence of the first-order belief which is the object of the self-ascribed belief. If this is the case, then the self-ascription of a first-order belief based inferentially is based on the same kind of evidence of the first-order belief, a consequence which would hurt weak groundlessness.

Strong groundlessness holds that cases of first-personal self-knowledge are simply not grounded in anything. As stated by Coliva, strong groundlessness can be described as the idea that first-personal self-knowledge is neither observational nor inferential, *nor is it epistemologically based on one's previous awareness of one's ongoing mental states*. If this is the case, then first-personal self-knowledge is not exactly an instance of knowledge and, therefore, it is a terminological error to call it "knowledge" after all. Coliva ends up adopting exactly this perspective, which is consistent both with expressivism and constitutivism.

Weak transparency consists in the idea that if one has a given mental state M, then one is aware of it, i.e., the mental state M is phenomenologically salient to the subject. Coliva gives us reasons to prefer strong transparency over weak transparency. This preferred variant of transparency states that, given C-conditions, which include concepts' possession, cognitive well-functioning, alertness and attentiveness, and to the exclusion of unconscious and purely dispositional mental states, if one has a given mental state M, then one will *be in a position to* judge or believe (or both) that one has it. I highlighted "be in a position to" passage because, as I think, it is subject to different interpretations, as I will explore next.

Finally, weak authority states the idea that, 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. On the other hand, strong authority is the idea that, given C-conditions, 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.

I begin with the passage, in the description of strong transparency, according to which, given C-conditions, if one has a given mental state M, one will *be in a position to* judge that one has it. As I think, this passage allows three possible interpretations, which I will call the *Wittgensteinian* interpretation; the *metaphysical* interpretation; and the *epistemic* interpretation. The Wittgensteinian one is the interpretation according to which "to be in a position to" judge that

one has a mental state M is a feature of the grammar we have, the option that has inspired expressivism. The metaphysical interpretation is another option, which has its roots in constitutivism, according to which “to be in a position to” judge that one has a mental state M is a feature of the kind of *metaphysical* relation we have with M. Finally, the epistemic interpretation, which can be identified in some epistemic accounts of first-personal self-knowledge, states that transparency consists in an *epistemic* relation between oneself and M.

Here, I will explore the third option, i.e., the idea that transparency is a feature of the kind of epistemic relation we have with some of our own mental states, such as sensations, intentions-as-commitments, and beliefs-as-commitments, in an attempt to save weak epistemic accounts of first-personal self-knowledge. What follows is the idea that one knows that one is feeling ψ , intending to φ or believing that p *based on* the transparency of ψ -sensation, φ -intention or the belief that p is the case. In this view, transparency is seen as playing the epistemic role in our self-ascriptions of some of our own mental states—precisely, those that are transparent to us. I am not sure, however, if the epistemic relation we have with our own mental states involves the kind of normativity that accompanies characteristic instances of knowledge.

The present view is not inconsistent with Wittgensteinian criteria for knowledge self-ascription. To see why, we need to consider the pragmatics of self-ascription. Consider questions such as “how do you know that you feel ψ / believe that p ?” In ordinary conversation, we assume, in cases of first-personal self-knowledge, the authority of the first person. My bet is that such questions—and the answers that would be appropriate to them—would hurt two elements of cooperative conversational practices, namely, informativeness and relevance. They hurt the element of informativeness that is expected in a conversation because the most immediate answer to such questions no longer tells you what is already assumed in the question: “I know that I feel ψ because I am feeling!”, or “I know that I believe that p because I believe!”. Therefore, it is also not relevant to ask someone with questions such as these. The strangeness with which we would react to such questions is explained by factors of the order of pragmatics. What follows is the idea that the *presumption of the first-person authority* is explained by cooperative conversational practices, i.e., the normal operation of a cooperative conversation. I highlighted “presumption of the first-person authority” because this is not an explanation about first-person authority itself, but about its recognitional conditions. This is because, in my view, first-person authority itself depends ontologically on the transparency of

the mental states, which, as I maintain, is a feature of the kind of epistemic relation we have with some of our own mental states.

Before ending my review of this incredible monograph, I need to comment a question that remains unanswered: cases of first-personal self-knowledge are accompanied by the characteristic normativity present in cases of knowledge? If “no”, the consequence of this view is that epistemic normativity is not a necessary condition for knowledge, because there are cases of knowledge, namely, first-personal self-knowledge, without epistemic normativity. If “yes”, then this normativity needs to be explained. Maybe the person’s conceptual mastery and its cognitive state of alertness, attentiveness, etc. can be the explanation of the epistemic normativity present in cases of first-personal self-knowledge. And if this is so, then the present weak epistemic account is consistent only with weak authority, because only weak authority allows the *possibility of error* in judging that we have a mental state M—and normativity in general, as I think it is plausible to assume, presupposes the possibility of error.