Knowledge of oneself is not easy to attain. Plato was aware of this and in this paper we aim to show that he suspected then, like psychologists know now, that one’s introspective capacity to attain knowledge of oneself is very much restricted and that we must rely on the other as a source of such knowledge. We further argue that, for Plato, this knowledge is not easily achieved given not only the shortcomings of the first-person perspective but also the limitations of the third-person one.

Keywords: Knowledge of the self, introspection, self-ignorance, self-error, cognitive bias.
RESUMEN

El conocimiento de uno mismo no es fácil de obtener. Platón estaba consciente de esto; en este artículo nos proponemos mostrar que él sospechaba en aquel entonces lo que psicólogos saben ahora, que la propia capacidad introspectiva para obtener conocimiento de uno mismo es muy restringida y que debemos confiar en los demás como fuentes de tal conocimiento. Además argumentamos que, para Platón, este conocimiento no se logra fácilmente, no solo debido a las deficiencias de la perspectiva de la primera persona, sino también a las limitaciones de la perspectiva de la tercera persona.

Palabras clave: Conocimiento de sí mismo, introspección, auto-ignorancia, auto-error, sesgo cognitivo.

INTRODUCTION

When we talk about self-knowledge we might be talking about knowledge of one’s nature and character. Alternatively, we could be talking about knowledge of one’s own present mental states (say, beliefs and desires). These are clearly not the same thing as long as the self itself isn’t merely a collection of present mental states. But one might think that, given that for Plato the self is the soul (which we might understand as the mind), there isn’t much difference between one and the other. This however would be wrong, although it is hard to state that Plato has a unified or systematic conception of the self.

2 It is understandable that a Plato scholar may have reason to reject an assimilation or even a comparison of the Platonic ψυχή to contemporary “mind”: when Plato speaks of the soul he isn’t just talking about a psychological power or about the mental states related to it, but also about an entity that is supposed to can determine the moral quality of a person. See Griswold, 1996, pp. 2-3. As observed by Brickhouse & Smith, 1994, p. 101, n. 41, Plato’s Socrates does not conceive of personhood as entailing possession of the body. If there is personal survival of the death of the body, the soul should continue to be identical to the person whose body died. For the discussion of the identification of the soul with the self in Plato see Gerson, 2003, pp. 22-24.

3 Or even a unified or systematic conception of knowledge. As objected by a reviewer, Plato’s conception of knowledge not always is consistent. In some early dialogues, the model is craft (Apology, Charmides) or recollection (Meno). In some later dialogues, the model is recollection of cognitive contact with Forms (Phaedo) and a power of the soul (a capacity that can attain “what completely is”; Republic 477a-d). The objection is that if we can’t provide a consistent view of knowledge in Plato, we cannot provide a consistent view of self-knowledge either. The reviewer is right when claiming that we cannot furnish a consistent view of knowledge, but we suspect that nobody can, as there is not such a consistent epistemological view in Plato. Our discussion is mainly based on Alcibiades I (hereafter Alc. I), whose authorship is disputed (see below), and the Apology; but as rightly pointed out by the reviewer, there are important details provided by Plato on the topic of the self and self-knowledge in Gorgias and Phaedo. However, we prefer to avoid the discussion of those dialogues, so we can focus better on the main purpose of this paper (as explained below).
Given that for Plato if one wants to know oneself, one must look at one’s soul (since the self-care is the care for one’s own soul—Apology 29e-30b—, the care of virtue, which is a state of one’s soul—Crito 45d; Euthydemus 275a; Philebus 11b), interesting questions can be raised. When we look at one’s soul, are we just looking at present mental states? And in what sense, if any, do we look at our own soul? If we take seriously the perceptual metaphor⁴, are we to assume that illusions and the like are possible in these cases? If that is so, how can we correct for self-error and self-ignorance? Here we try to unearth the Platonic answers to these questions and to show that Plato suspected then, like psychologists know now, that one’s introspective capacity to attain knowledge of oneself is very much restricted and that we must rely on the other as a source of such knowledge. At any rate, we want to further argue that, for Plato, knowledge of the self isn’t easily achieved given both the limitations of the first- and third-person perspectives⁵.

The paper proceeds as follows. Firstly, we consider whether knowledge of oneself (soul) is merely knowledge of one’s own mental states. Secondly, we examine in what sense we look at our own self (soul) and, in particular, the possibility of error and ignorance given the first-person perspective. Thirdly, we argue that the third-person perspective can in principle help us correct the cognitive failures and limitations of the first-person perspective, but that we might be unlikely to take advantage of it. Finally, we offer some concluding remarks.

Before developing our argument, we would like to clarify the way in which we use both Plato’s dialogues (mainly the Apology, the Theaetetus, and the Philebus) and Alc. I, a text whose authenticity is still under dispute⁶. We do not intend

---

⁴ As presented in Alc. I 133b7-10. All our references to Plato are based on the Greek text edited by Burnet, 1976-1979.

⁵ Our talk of perspectives refers to the epistemic access of some subject to some object. The object could be one’s mental states or personal traits or someone else’s or something else. However, while the third-person perspective can be applied to all kinds of objects (including oneself –just like Shoemaker’s alleged self-blind person does– see Shoemaker 1994), the first-person perspective is restricted to the epistemic subject herself. Another and more important difference between those two perspectives is the sort of access that they involve. My access to my sensation of pain differs fundamentally from someone else’s access to this pain of mine. I know that I am in that state because I have the experience of it. Someone else, on the other hand, knows that because of public evidence based primarily on observation of behaviour to which anyone can in principle have access (i.e. public evidence).

⁶ In the 19th century, especially due to the authority of Schleiermacher, the authenticity of this dialogue has been questioned. Some contemporary scholars also forcefully challenge the authenticity of Alc. I, notably Smith 2004 and, more recently, Renaud & Tarrant 2015, who tend to support the inauthenticity (p. 38, 46, and 267), even though they declare that their purpose is “to explore the hypothesis that the dialogue may be interpreted along the same lines as any other dialogue of Plato” (p. 5). There are some other interpreters who tend to take Alc. I to be authentic (Annas 1983, pp. 114-115, quoted by Gill 2006, p. 344. See also Pradeau 2000, p. 20-29. Denyer, 2001, p. 14-26.
to engage in the discussion of whether this dialogue belongs to Plato; we will assume without argument that it may be “Platonic” in character or “a legitimate and influential part of the Platonic tradition”. Both the Theaetetus and the Alc. I are especially relevant for our project since they allow us to stress the idea that what is dealt with there might have considered the first and third-person perspectives in knowledge. Someone might think that the use we give of these (and other) “Platonic” passages takes no account of the context in which the texts are embedded. We certainly are aware of that, but this is the manner we have to put into dialogue Ancient philosophy with some contemporary views and show that the Ancients noted almost the same problems we can discuss nowadays (or so we shall argue).

KNOWLEDGE OF ONESELF AND KNOWLEDGE OF ONE’S MENTAL STATES

Mind-body dualism is usually attributed to Descartes by contemporary philosophers of mind. This view gives rise to the still famous and difficult problem of mind-body interaction: if mind and body are two so different entities as Descartes holds, the problem is, for example, to explain how it is possible for the mind to make our limbs move. But the problem goes back to Plato, who not only distinguishes the soul (or mind: ψυχή) from the body, but also attempts to prove that one’s soul is the one’s self, and that, given that the soul is able to use and rule over the body, the soul is more valuable than the body (cf. Alc. I.129e-130c, but see also Phaedo 79e8-80a6; Laws 896c1-3). Insofar as Plato holds that there

Gerson, 2003, p. 14, n.1, prefers not to take a strong position on the authenticity –both of the Theaetetus and the Alc. I,–, although he tends to accept their authenticity). It should be recalled that the view that the Alcibiades I belongs to Plato isn’t new (cf. Adam 1901, pp. 40-65).

As Renaud and Tarrant (2015, p. 5) elegantly refer to the dialogue.

“Platonic” meaning here both those dialogues that unquestionably belong to Plato and the Alc. I, whose authorship is disputed.

Cartesian dualism is well attested in the Meditations on First Philosophy II (AT VII 23-25; IX 20-21); cf. Notae VI-VII (AT VIII 343-344). See also Letter to Elisabeth, Primitive Notions (May 21 and 28, 1643; AT III 665; AT III 691-692). Descartes seems to have assumed that the brain is “the seat of our mental life”, since there is a pervasive system of correlations between mental events and brain processes (see Meditations on First Philosophy VI. See also Discourse on Method Part V et passim).

Plato’s “substance dualism” can be seen in Phaedo 78c-79d. As observed by Long (2005, p. 174), philosophy of mind is a relatively recent discipline “indelibly coloured by Descartes and the sharp distinction he introduced between the physical and the mental. Inasmuch as Plato discusses such concepts as belief, desire, pleasure, action, and emotion, he can be brought into contact with what we call philosophy of mind. But the mind/body distinction, in the way it has preoccupied modern
are two distinct domains of entities (psychological and physical items, the former being immaterial, and the latter material), one could link his view to the so-called “classical dualism” discussed in contemporary philosophy of mind (a dualism inspired by Descartes).  

Now in some Platonic dialogues (notably in *Alc. I*, but see also *Apology* 36c5-6) the main issue is what caring for oneself is (τὸ ἑαυτὸν ἐπιμελεῖσθαι); according to the *Alc. I*’s argument, people usually think they are caring for themselves when they are not, since while one can care for what one has, one actually cannot care for oneself (*Alc. I* 128a1-d7). Self-care usually is focused on “taking care of what belongs to oneself” (such as a ring that belongs to a finger). But taking care of what belongs to oneself is different from taking care of oneself, because the “oneself” isn’t what belongs to oneself. Plato’s interest concentrates on elucidating when and in what respect it can be said that a person really takes care of herself. But in order to do that, one should know what the “self itself” actually is. It is with athletics that we take care (ἐπιμελούμεθα) of our body, and we take care of what belongs to our body with weaving and other skills (128c1-4). So when one is taking care of what belongs to oneself, one isn’t taking care of oneself.

In virtue of the inductive method used by Socrates (Plato), it follows that, if this is the case in different skills or crafts (τέχνα) with regard to different objects, it also should be so in the case of what is to be a thing itself, on the one hand, and what belongs to such a thing itself, on the other. So while we take care of each thing with one skill, we take care of what belongs to it with a different skill. But it appears that, although one can take care of what belongs to oneself, one cannot take care of oneself. Once it is accepted that the “thing itself” and what belongs to the “thing itself” aren’t the same, one can proceed to generalize and apply this distinction to the self and what belongs to the self. This being so, if we know “the itself” (αὐτὸ) we ourselves are, then we might know how to care for ourselves, but if we don’t know what we ourselves are, we will never know how to do that. But the point is how could we find out what the “itself itself” is (αὐτὸ ταύτο).  

---

philosophy, has only partial bearing on the Platonic dualism of body and psyche”. In Modern philosophy the starting-point of enquiry is used to be the mind’s consciousness of itself (for a full discussion of these distinctions, see Broadie 2001, pp. 295-308).

11 We are aware that this kind of treatment of Plato can be disturbing for some Platonists. Yet, given that we believe we can bring Plato into discussion with contemporary philosophy –keeping in mind, of course, the more classical treatment of Plato’s texts by Platonic scholars– we are willing to take the risk of interpreting Plato within a contemporary philosophical framework. See also n.1 above.

12 Denyer 2001 ad locum prints αὐτὸ τὸ ταύτο; this reading stems from the indirect evidence (see Olympiodorus, *In Platonis Alcibiadem* 201, 17; 204, 3-4; 205, 3; 209, 24) and maybe from *Alc. I* 130d4. This way of reading the passage allows Denyer to understand the expression αὐτὸ τὸ ταύτο as referring to Platonic Forms (p. 211-212). But it isn’t necessary to change the version of the MSS:
what we have to examine is “the itself itself” we are, for we won't assert that there is anything of ourselves that is more “authoritative” (κυριώτερον) than our soul (ψυχή; *Alc. I* 129a8-9; b1-3. 130d3-6)\(^{13}\).

*Alc. I*’s emphasis upon the fact that the active factor is the soul, not the body, is in line with the tenet developed in a more detailed way in the *Phaedo*: I move my arm because I want to, or decide to do that. But in the Platonic ontological map wanting and deciding are psychological items, so the soul is what sets in motion one’s own body (i.e. what “rules” over one’s body)\(^{14}\). This shows that the soul is the user and the body is what is used\(^{15}\), and a reliable answer to the question “what is the man?” is “the man is what uses the body”, or insofar as the soul uses the body, one could say “the man is the soul”\(^{16}\).

one might argue that αὐτὸ τὰυτό just refers to the “thing itself” as different from what belongs to the thing (see the next note). As persuasively argued by Smith (2004, p. 105), within the *Alc.* I the construction αὐτὸ τὰυτό obviously refers to the self of Alcibiades, “a single, individual self that is not the metaphysical standard for all individual selves”. Croiset, 1980, ad locum (Tome 1) prints and translates the transmitted text (αὐτὸ τὰυτό) thus: “ce que c’est au just que soi-même”. In a footnote he paraphrases the innovative and uncommon expression αὐτὸ τὰυτό like this: “un être lui même”. See also Pradeau 2000, ad locum, who translates this difficult expression by “ce soi-meme lui-meme”. In his note 121 to the text (p. 210), he explains that his translation renders a “redondance pronominale assez inhabituelle”.

\(^{13}\) With the expression “the same in itself” (or “the thing itself”) we don’t mean a standard Platonic Form, but the thing one calls foot or finger in Plato’s examples when attempting to distinguish what something really is and what belongs to it. On this point see Pradeau, 2000, p. 210-211. “Thing itself” is the unfortunate way we have to render the Greek αὐτὸ τὸ αὐτὸ in these passages; sometimes the expression is rendered by “the self itself” (Gerson 2003, p. 22, n.14; Renaud & Tarrant 2015, p. 20 –where they recognize that both the translation and the Greek αὐτὸ τὸ αὐτὸ are clumsy—, 58, 250) surely assuming that in the analogy with the crafts “the thing itself” is the self, the soul, or the person. However, as observed by Gill, 2006, p. 347, the English term “self” is “a rather peculiar usage, representing the conversion of a reflexive suffix (as in ‘himself’) into a noun”. But the Greek αὐτὸ τὸ αὐτὸ doesn’t mean a fundamentally reflexive relationship, but “what is the core or essential features of ‘the itself’” (349). The passage from which the above argument comes is part of the final section of an account intending to show that the human being should be identified with her soul.

\(^{14}\) See Plato, *Phaedo* 98c-e. The real causes are the Athenians’ thoughts that it was better to condemn Socrates. So the real cause of Socrates remaining seated on his bed was both what the Athenians and Socrates thought. Of course, thinking is a psychological item, so Socrates’ soul “ruled” over his body.

\(^{15}\) For the “user-used” terminology see *Alc. I* 129c2-130a1; but see also *Phaedo* 79c3, where the soul is what uses the body (τὸ σώματι προσχρήσαται).

\(^{16}\) If the body cannot be the man, what rules the body cannot be the combination of soul and body, this being so because the body doesn’t take part in ruling, so no combination (of body and soul) could rule. Therefore, if man cannot be the body or the combination of soul and body, it must be the soul (*Alc. I* 130d). Gerson, 2003, pp. 22-29, provides a full discussion about the identification of the person or self with the soul. As pointed out by Gerson, “there is no other dialogue among those recognized as ‘early’ in which the identification is made so explicitly”.
So to know oneself is to know one’s soul (mind). Now if knowledge of oneself is to know one’s soul, one should look at what one really is (soul), and especially at that region where the highest good (virtue, wisdom) occurs (Alc. I 133b7-10). So the argument appears to be that in order to know yourself, you need to look into what you really are. But could this knowledge of oneself just be knowledge of one’s own present mental states? The answer seems to be “no.”

Quickly put, while we seem to enjoy some sort of privileged access to our own present mental states, we importantly don’t in the case of character traits, abilities, values and the like. Indeed, Plato believes that we enjoy such privileged access in the former case, as we shall next see, but not in the latter, as we shall see in the next section.

But first let us make clear what we mean by “privileged access.” Self-ascriptions of mental states are normally claimed to be special, in the sense of being epistemically distinctive. This claim is normally understood in terms of privileged access, but there are many different senses of the term (e.g., Alston, 1971, pp. 223-241; Gallois, 1996; Gertler, 2003). We can anyway differentiate between two main epistemic claims: on one the hand, the security of self-ascriptions and, on the other, the peculiarity of self-ascriptions. Roughly, the latter maintains that self-ascriptions can be known differently from the way we know the ascriptions to others, while the former states that self-ascriptions can be known better than the ascriptions to others.

With regard to the latter claim, some of the epistemic features that have been put forward for explaining the security that self-ascriptions are meant to enjoy are: infallibility, incorrigibility and self-intimation. A self-ascription is infallible when it is the case that, if one self-ascribes a mental state, then one has in fact the mental state. A self-ascription is incorrigible when it is the case that, if one self-ascribes a mental state, then one cannot be shown to be mistaken about it by others.

17 This region is said to be “divine” and is assimilated to knowing and understanding (ιδέεται τε και ἐφησέλαται), as that region of the soul resembles god or the divine (depending on the reading one adopts at 133c4: ὑπὸ —cod. Bodleianus— or ὑπὸ —cod. Venetus—). Certainly, we are not assuming that looking at one’s arete is simply the same as looking at our internal states. Pradeau (2000, pp. 76-78; also commented on by Gill 2006, p. 357) has suggested that self-recognition is conceived in “aspirational terms”: to see yourself is to see what you could become, that is, wisdom, knowledge, and thought, some capacities that are godlike. But in a sense you really are the best thing you are able to be.


19 Peculiarity and security are two independent claims. So, for example, Ryle, 1949, ch.VI, denies the peculiarity claim (our access to our own minds is the same in kind from the one to other minds) while accepting the security claim (we have more access to the relevant evidence in our case than in the case of others); so he rejects only one of the two claims associated with privileged access. The reverse case is also possible.
(note that incorrigibility doesn’t entail infallibility, even if it does suggest it). A self-ascription is self-intimating when it is the case that, if one has a mental state, then one self-ascribes the mental state.  

The thesis that each person has privileged access, in the security sense, to her own present mental states is suggested to some extent in a striking passage in the Theaetetus (hereafter Th.), where Plato argues thus: (i) when what acts upon me (τὸ ἐμὲ ποιοῦν) is for me and not for someone else, I, and not someone else, perceive it. (ii) Therefore my perception (ἀισθητική) is true for me – “because it is always of my own being” (ἐμὴς οὐσίας ἄει ἔστιν) – and, according to Protagoras, I am the judge of the things which are for me, that they are, and of the things which are not, that they are not (Th. 160c4-9). This passage is part of the concluding section of an argument that attempts to prove that the state in which the person is when having some perceptual experience isn’t irrelevant. We certainly aren’t assuming that in this passage of the Th. Plato explicitly suggests that we have “secure access” to our mental states. But if what we argue in what follows is sound, he seems to have considered the problem that contemporary philosophers call “secure access to our mental states”. If this passage is analysed in its context, it’s arguable to maintain that the problem Plato is regarding isn’t alien to the issue of the access to our mental states. In Th. 159a ff. he develops the following argument: (i) there is an infinite number of both active and passive things. So, (ii) when one thing mixes at one moment with one thing and at a different moment with another, it won’t produce the same effect each time but different effects. Now (iii) if this is applied to us, one could notice that, for example, “Socrates ill” and “Socrates well” are different states of Socrates, that is, “Socrates in sickness” is different from “Socrates in health”. This should mean that when anyone of the active factors finds Socrates in health, it will be dealing with “one Socrates” and when it finds him ill with “a different Socrates”. Now, if Socrates drinks wine when he is healthy, the wine appears (φαίνεται) to him pleasant and sweet (159c11-12), but when the same active factor (e.g. wine) finds Socrates ill, he is not exactly the same person. Thus both Socrates (ill) and the draft of wine generate different things: a perception of bitterness on the tongue and bitterness coming to be and moving with regard to the wine. But the wine, Plato insists, is not “bitterness”, but “bitter”, and the person is not “perception”, but “percipient” (οὐκ ἀισθητικὴν ἄλλα ἀισθητανόμενον; 159e4-5). (iv) Hence, the subject will never become thus percipient of anything else, since a perception of something else is another perception, and makes another and a changed percipient. But what acts upon the subject, when meeting with another

---

20 Of course, there are other ways of capturing the security claim, but for present purposes we don’t need to introduce them.
21 This is so, Socrates states, because what is active and what is passive, “moving simultaneously”, produce both sweetness and perception.
person,\textsuperscript{22} cannot generate the same effect since such effect depends both on the active factor (the wine) and on the state in which the person is (healthy or ill). (v) Finally, Socrates states that perception is an “intentional item”, i.e. when the subject becomes peripient, she becomes peripient of something (160a-b), and when the perceived object becomes sweet or bitter, it must become so for somebody. Then both what perceives and what is perceived “are or become for each other, for our being is by necessity tied to a partner” (160b6-7). Whatever it is that we are or what we come into being, we (as peripients) are in a reciprocal relation to the perceived object. Thus it cannot be said that something is or comes into being in itself and by itself.

Bearing all of this in mind, let’s get back to \textit{Th.} 160c4-9 (discussed above): (i) what acts upon me (a sweet object, for example; \textit{Th.} 159c-d) is for me and not for anyone else (it is I who perceives such an object, so nobody else can experience the sensations I am experiencing). (ii) My perception is true for me insofar as it is always of my own being. Thus (iii) I am the judge of the things that are for me, that they are, and of things which are not, that they are not.\textsuperscript{23} As long as the suggested infallibility (which denies the possibility of error)\textsuperscript{24} isn’t found in the ascriptions of

\textsuperscript{22}At this point the Greek is ambiguous (as long as \textit{όλλα} can be both neuter and masculine. See McDowell, 1996, ad locum). We think it should be masculine since the whole argument stresses the fact that the state of the person isn’t innocent in the account of perception. So, the way in which the subject is disposed somehow determines the kind of perception such a subject will experience.

\textsuperscript{23}We are aware that this is supposed to be Protagoras’ relativistic and subjectivist view, or rather Plato’s version of what he takes to be Protagoras’ thesis. Polansky (1992, p. 107) points out that “this establishes the truth of perception on the basis of the flux doctrine”, the flux doctrine being Heraclitus’ theory that Plato tends to conflate with Protagoras’ view. Plato proceeds to ascribe to Protagoras “a secret doctrine” (\textit{Th.} 152c) which, actually, is a version of the view that everything is changing all the time, an approach usually associated to Heraclitus by Plato (cf. \textit{Cratylus} 402a8-10; 411b-c). Burnyeat (1990, p. 39) asserts that “the real Protagoras didn’t hold the subjectivist thesis”, since “he was a relativist who maintained that every judgement is true for (or in relation to) the person whose judgement it is”. His point is that, although the \textit{homo mensura} thesis introduces a form of relativism, Protagoras’ relativism cannot be the “crude relativism” that assumes that Protagoras is defending a subjectivist rather than a relativist thesis. If the Protagorean view is a “subjectivist thesis”, Protagoras would be fully endorsing the stance that every judgment is true absolutely, not merely true for the person whose judgment it is (Burnyeat, 1976a, pp. 46-47. Cf. also Burnyeat, 1976b, pp.172-173; both papers were reprinted in Burnyeat 2012, pp. 2-47). For a still helpful discussion of the \textit{homo mensura} tenet, see Lafrance, 1981, pp. 227-242. Lafrance (p. 228, n.110), unlike Burnyeat, interestingly thinks that both Plato and Aristotle, in attributing to Protagoras the approach that assumes that all the judgments are true, interpret in the same way Protagoras’ formula. We don’t intend to engage in this discussion here, but to point out that the passage can be understood as part of Plato’s exploration of what knowledge is when it is identified with perception, a view that anyway is finally overturned by the character Socrates.

\textsuperscript{24}The self-intimation thesis, on the other hand, denies the possibility of ignorance. As long as Plato doesn’t accept this thesis, his position isn’t Cartesian: in the sense of denying both self-error and self-ignorance regarding one’s own mental states.
present mental states to others, as one would expect (after all, people can deceive us regarding their own mental economy—indeed, good actors are the sort of people who can easily do this), then Plato seems to be suggesting that we have secure access to our own present mental states. But it seems clear that, for example, we don’t seem to enjoy a secure (i.e. infallible) access to our character traits (e.g. Dunning, Meyerowitz, Holzberg, 1989, pp. 1082-1090). In fact, contrary to what is the case with regard to one’s self-ascriptions of one’s mental states, in the case of character traits there is no presumption of truth. In other words, there is no presumption that you aren’t mistaken about such things. So such self-ascriptions can and are normally challenged (say, you tell your spouse that you are affectionate and the reply you get is: “You must be kidding!”). And importantly, as we shall see in the next section, Plato is aware of the fact that we don’t enjoy a secure access to our character traits. If this is so, knowledge of oneself couldn’t simply be knowledge of one’s own present mental states.

Moreover, the limitations of the first-person perspective seem to extend beyond character traits. There also seems to be no secure access in the case of the prediction of future mental states and behaviour (Wilson & Gilbert, 2003). The same seems also to be the case with the causal sources of our actions (Nisbett, & Wilson, 1977). And of course the fact that we are often wrong about these predictions and causes might be related to the fact that we also are about our character traits. After all, if one thinks wrongly that one is friendly, one might wrongly predict certain friendly behaviours. Similarly, if one thinks wrongly that one is friendly, one might wrongly think that such alleged friendliness is the cause of a certain friendly-like behaviour. So not only one’s knowledge of oneself doesn’t seem to be subsumed under knowledge of one’s mental states, one’s secure access seems very much restricted.

LOOKING AT OUR OWN SOUL

In the *Alc.* I the topic that triggers the issue of what one is in himself presents Socrates trying to remind Alcibiades that if he wants to take charge of the political matters of the city before he has learnt to take care of himself, he will be in trouble. It is from this observation that in the last section of the dialogue the discussion concentrates on what it means to know and take care of oneself. When the character Socrates attempts to explain the meaning of the Delphic inscription, “know thyself” (129a; 132d-e), he suggests that what the inscription is saying is “see yourself” (132d6). Socrates warns Alcibiades that before he gets involved in politics, he must learn what the antidote against the terrible dangers related to political activities is. So Alcibiades says: “try to explain how we should take care of
ourselves.” Socrates replies that they have already said what they are and what they are not: they are soul, not body. If this is so, one should take care of the soul and look at it (Alc. I 132c1-2). According to Socrates, the better way to understanding the meaning of the injunction “know thyself” is only (μόνως) by means of sight (132d2-3). And this “perceptual talk” allows us to posit the possibility of error and ignorance with respect to our access to our own soul: more specifically, to that region of the soul where excellence occurs. Let us explain.

A perceptual model regarding some such access maintains both that a first-level state is conceptually and ontologically independent of the second-level state and that the former causes the latter. These are the two fundamental features of any such model (Shoemaker, 1994). Now, firstly notice that these two features of the model already allow us to accommodate the fact that successful self-ascriptions involve a genuine cognitive achievement (since they track their ontologically distinct self-ascribed states). Secondly and more importantly, given that the connection between the first and the second-level states is a contingent one, the model allows, as one would expect, for self-ignorance and self-error (since the states are distinct existences, the one could exist without the other). In fact, just as we can fail to perceive and misperceive external-world objects, we can also do so in the case of internal-world ones. The commonplace appearance/reality gap regarding external-world objects is then extended. On the one hand, self-ignorance is possible because one could lack contact with the independent reality. And on the other, self-error is possible because one could make imperfect contact with the independent reality.

Of course, one can take the perceptual analogy further and suggest that there is an actual “eye of the soul.” This “inner eye” would then detect by some sort of perception the states to be self-ascribed. But this view seems wrong. There is no inner sense-organ and our self-ascriptions don't have a distinctive phenomenology.

25 However notice that the latter feature requires the former, given that in order for one state to cause the other, they need to be, as D. Hume, 1978 (Book I, Part IV, section 2) says, “distinct existences”, i.e. ontologically distinct states.

26 Maybe Socrates might argue that when he thinks that he is both rightly able to identify his own state of mind and the state of mind of others he has “a genuine cognitive achievement”. Socrates is entirely sure that he is wise in no matter great or small and that he knows nothing (Apology 21b; 22c). And this must be so because he is aware (this is the first-person perspective) that he isn't able to give a correct account of some cognitive contents (such as the moral virtues he is asking the others about). On the contrary, the other people wrongly think that they are capable of saying what justice, piety, moderation, and so on, are. Note that the weight of Socrates' argument is on his own state of mind, that is, on the first-person perspective. Thus Socrates' genuine cognitive achievement is that he is aware of his ignorance (23b).

27 Plato uses the expression “eye of the soul” (Republic 533d2; Phaedo 83a-b) but in a very different context to make reference to the ability one's soul has to grasp Forms. The metaphor of the “eye of the soul” can be understood as an understanding (of the Forms) obtained through a “noetic act” (as suggested by Vegetti, 2003, p.425).
proponents of the perceptual model should play the analogy down and just focus on the idea that there is some tracking of the states. Still, given this model, we should take the possibility of error and ignorance seriously. There is a potential mismatch between, say, one’s self-ascriptions and one’s states. And in order to avoid error and ignorance, one has to make adequate contact with the independent reality. Indeed, the soul is there to be discovered rather than invented by our inquiry into it. And since discovery is a matter of contingent success, there is no guarantee that one will discover such reality.

Now, given this model, security (i.e. infallibility) about our access to that region of the soul where excellence occurs isn’t a live option. This also suggests that the methods by which we arrive at our knowledge of that region of the soul and our knowledge of our own mental states are different. The knowledge in those cases is achieved in different ways (since one method allows for error but not the other). So, for Plato, we cannot identify our knowledge of that region of the soul where ἀρετή occurs with our knowledge of our own mental states. The former couldn’t simply be the latter, since when we look inward at that region of the soul, our lenses might distort what we see. In the Philebus, Plato gives cases of people who falsely believe to know themselves: some people think that they are stronger or more handsome than they really are and the largest group of people believes that they are better persons regarding their virtue (ἀρετή δοξάζοντες), although they are not (48e9-10). It is regarding wisdom (σοφία) that most people are full of rivalry and of an apparent and false wisdom or intelligence\textsuperscript{28}.

This isn’t a strange phenomenon and there is a lot of evidence regarding this sort of shortcomings. For example, when asked what we are like, we report on our traits and abilities but those reports are biased in a self-aggrandizing way. Most people think that they are more beautiful, amiable, intelligent, etc., than the average person. This is statistically impossible and so some must be wrong. This bias is known as the “better-than-average effect” (Alicke, 1985), but of course it is not the only one we suffer in this domain. Another example would be the self-serving attributional bias, which is a tendency to attribute the causes of positive events to one’s own personality, while attributing the negative events to factors outside of one’s own control (Bradley, 1978. Mezulis, Abramson, Hyde, Hankin, 2004). Moreover, there is a tendency for people to non-consciously conform to those around them\textsuperscript{29}. And since people tend to deny this social influence,\textsuperscript{30} we

\textsuperscript{28} Philebus 49a2. On “apparent wisdom” (δοξάσσομαι), see also Philebus 49d, Sophist 231b, and Phaedrus 275b2, where Plato speaks of those who imagine they have come to know much while merely appear to be wise instead or really being so (δοξάσσομεν γεγονότες ἀντὶ σοφῶν). On this Phaedrus passage, see Griswold 1996, p. 206-207.

\textsuperscript{29} See, e.g., Darley, &. Latane, 1968, Charttrand, & Bargh, 1999.

\textsuperscript{30} See, e.g., Pronin, Berger, Molouki, 2007.
can easily mistake (positive) behaviours that are the result of this implicit social influence that we are anyway unaware of and attribute them, given our self-serving attributional bias, to our character.

We are susceptible to a myriad of biases and other cognitive obstacles (Dunning, 2005) that limit significantly our knowledge of ourselves. This suggests that in fact there is often a bad fit between one's own perceptions of oneself and what effectively the case is. In particular, we seem to suffer from many enhancement illusions that suggest that the correlation between self-perception and reality can be rather weak. This seems to leave us with a rather bleak picture regarding the prospects of achieving knowledge of oneself. The question then is how, if at all, we can correct for this self-error and self-ignorance.

THE THIRD-PERSON PERSPECTIVE

As seen above, according to \textit{Alc. I} Socrates, the better way we have for understanding the meaning of the injunction “know thyself” is by means of sight (\textit{Alc. I 132d2-3}).\footnote{Some scholars think, though, that the idea of knowing oneself through another isn’t the central theme in the \textit{Alc. I} (Pradeau, 2000, p. 74ff., cited and endorsed by Gill, 2006, p. 356-357). By contrast, we believe that, given the emphasis Plato makes upon the idea that one is able to know himself through another person’s sight, the point is relevant. It must be recalled that, as we mention below, the inscription “know thyself” advises our eye “as if it were a human being.” So the visual image and the vivid dialogue are related to each other.} This suggests a perceptual model that makes sense of self-error and self-ignorance within this domain. But in this section we want to argue that it can also be understood as suggesting a way to correct those limitations of the first-person perspective.

Given the connection between “know thyself” and sight, if the inscription advised our eye as if it were a human being and said to it “see yourself,” the injunction would be advising that our eye look at something in which it could see itself (\textit{Alc. I 132d7-8}). The artefacts in which one can reflect oneself are mirrors but also the other’s eyes which can work as mirrors: when one looks at someone else’s eye, one’s face appears in it. This is the pupil, which functions as a sort of image of the person who is looking (133a2-3). Then an eye –in contemplating another eye and looking at that which is the best of it, the part with which it can see (i.e. the pupil)– could see itself. That is, when one sees someone else and looks at what is the best in this someone else, one could see oneself and hence know oneself. So the interesting point made is that one can see oneself (i.e. to know oneself) through someone else’s sight. It is the person seeing herself, but thanks to the other person’s mediation\footnote{This reflection picture is here understood as presenting third-person (public) evidence.}.
Now, given that the above first-person perspective limitations are due to first-personal obstacles (e.g., better-than-average and self-serving attributional biases), such cognitive shortcomings shouldn’t be found in the third-person perspective. The others, who anyway have all the same first-personal biases and the like that one has but which won’t affect them when considering one’s case as opposed to theirs, can provide one with the knowledge of oneself that one seeks but one’s cognitive obstacles threaten. So Plato’s insight that one can see oneself through someone else’s sight should be understood as the insight that one can get better information about oneself through someone else. Indeed, if one isn’t aware of one’s biases and the like and so cannot correct them, the third-person perspective can offer one a way to overcome such cognitive shortcomings. And, in fact, although we can recognize, for example, our general susceptibility to biases, when dealing with particular cases we seem to have some sort of bias blind spot. That is to say, we seem to have a bias regarding the recognition of our own biases in concrete cases. And although we readily recognize the biases of others, we fail to see those biases in our own case (e.g., Pronin, Gilovich, Ross, 2004; Pronin, Lin, Ross, 2002).

So the third-person perspective has the potential to be very useful in our search for knowledge of ourselves. Unfortunately for us, such potential might not often be actualized. This is due to the fact that we also have a tendency to confirm our own existing self-views (Swann, 2011). In fact, we are resistant to changing both our positive and negative self-views. So one is normally resistant to whatever others might tell one about oneself as long as the information conflicts with one’s existing self-view. To make matters worse, we also have a tendency to see others as biased while still believing in our own objectivity, which is particularly prominent in cases of disagreement. In other words, there are some psychological mechanisms that might blind oneself from the input of the other. So, although in principle others are better sources of knowledge of oneself than oneself (say, one’s mother might have much deeper insight into one’s character than one does), one’s cognitive make-up is such that one will not often recognize others as such.

Of course, this is not to deny that, under some conditions, people might be open to acquiring knowledge of themselves through others. For example, one might be open to acquiring such knowledge when, among other things, the

---

33 See above (§2) for the denial of social influence. Similarly, people seem to fail to notice their prejudices about race, gender, age, etc.; see, e.g., Devine 1989; Greenwald and Banaji 1995; Dovidio et al. 2002.

34 So, for example, to know that you are egotistical you have to believe so, but it might be hard for you to do so if such character trait is at odds with your self-view.

35 As pointed out by Kennedy & Pronin, 2008.
other is recognized as a genuine expert on the particular issue considered. After all, if one does not consider oneself as an expert on the issue at hand, the other in that case would be regarded as one’s epistemic superior and so when realizing that there is a disagreement between one and the expert regarding that issue, it is pretty straightforward that one should adopt the expert’s view on the issue (or at the very least suspend judgment about the issue). This is of course what we normally do when we are confronted with the conflicting opinion of an expert who is considered as our epistemic superior. We normally change our mind when we are aware that expert astronomers, historians and physicists, who are regarded as such by us, disagree with us on issues related to their expertise. In fact, just as we consult doctors for their opinions regarding issues about our physical health and lives, many of us consult (and pay good money to) psychologists, therapists and the like for their opinions regarding issues about our mental health and lives (i.e., regarding issues related to our self-views)\(^{36}\).

Still, is Plato aware that others might not actually be exploited as sources of knowledge of oneself? Of course, he is aware that we aren’t often willing to give up our own beliefs, whatever their subject matter; after all we believe what we do because we believe it to be true, since nobody admits that her own belief (οἴηςις) is false (\textit{Tht.} 171a6-7; b1-5; 200a3).\(^{37}\) Given this Platonic view, we can envisage a general willingness to retain the beliefs we have, including importantly those about ourselves. Moreover, given that we enjoy secure access to our own mental states and that we seem to be aware of it,\(^{38}\) we seem to have prima facie reasons not to doubt that we in fact hold such beliefs. So, although one’s self-views might just be false, we have secure access to those beliefs that make up one’s self-views and are unwilling to give them up since what we believe, we believe it to be true. Plato then seems aware that we might be resistant to changing our self-views given the input of others.

In fact, he seems aware of the fact that we might be especially resistant when it comes to changing our own self-views. This is because, although throughout the Platonic dialogues characters change and give up beliefs they held before encountering Socrates (but only after arduous argumentation), not even Socrates is willing to accept (without examination) the testimony of the Oracle of Delphi regarding his wisdom. When the Oracle tells Chaerephon that Socrates is the wisest man, Socrates did cast

\(^{36}\) For a similar view, see Plato, \textit{Tht.} 178c-d, where the doctor’s opinion should be favored over the layman’s one.

\(^{37}\) On this issue see \textit{Alc. I} 113b8-11; 117b-118a. \textit{Charmides} 166d-167a. \textit{Sophist} 228c-d. For a similar approach in Davidson, see his 2001, p. 4.

\(^{38}\) At least, the presumption of truth that self-ascriptions regarding mental states enjoy in ordinary contexts (where we don’t have reasons, say due to insincerity or inconsistency, for doubting the ascription) supports the latter.
doubts on that and he didn’t accept it straightaway. In fact, he doubted the testimony of the Oracle, who happens to be the expert on the matter and whose infallibility can hardly be put into question (the Oracle is after all a spokesperson for Apollo and of course the gods are never wrong: indeed, if a god says it, it must be true), and he talked to all sorts to see if this was actually true. In the end Socrates realized he was the wisest, since he at least knew the limits of his knowledge. While he had such self-insight, others around him lacked it. Socrates was aware that he knew nothing (Apology 21b4-5; 22c9-d1), but others in fact perceived themselves as wiser than they really were. Still, even Socrates, who actually happens to be the wisest man, doubted what his epistemic superior (and the best expert you could actually find on the matter) said. This then seems to suggest that Plato is aware of the fact that we are even less open to the conflicting opinions of others when they concern our own self-views. Plato seems to think that we are more resistant to changing our own self-views than other views. All of this suggests that Plato also seems to be aware of the limitations of the third-person perspective to actually correct the shortcomings of the first-person perspective. Plato, as seen before, is aware of our own limitations.

39 Another way to put this is to argue that Socrates isn’t skeptical about what the god says to Chaerephon (i.e. “Socrates is the wisest of men”), but he is attempting to understand what the god has said (see Brickhouse & Smith, 2000, p.77). For the sake of our argument we would like to exploit the possibility that Socrates has a doubt about what the god says insofar as he is sure that he isn’t a wise person (this suggestion can gain some support if one considers that Socrates doesn’t accept straightforwardly and without investigation what the god says). This isn’t the same as saying that Socrates wants to prove that the Oracle is wrong (Nehamas, 1990, p. 306). Nehamas’ opinion is endorsed by Woolf (2008, p. 80, n.8), who suggests that, given the uses of ἐλέγχειν, “to put to the test”, in Apology “it seems hard to doubt that Socrates envisages the possibility, at least in principle, of the Oracle’s refutation”. This, we hold, cannot be the case, since Socrates is aware that surely (δὴ τοῦ) the god cannot lie (Apology 21b6). The examination Socrates performs of what the Oracle says doesn’t necessarily mean that he is trying to refute the Oracle (besides, he explicitly recognizes that the oracle is irrefutable. See Apology 22a7-8: ἀνέλεγκτος ἢ μαντεία; this line is cited by Woolf, who however thinks that Socrates envisages the possibility of the Oracle’s refutation). He can be just examining what the Oracle says in order to try to understand it. Our view rather is that, even accepting both that Socrates is trying to understand what the Oracle says about him and that he is sure that what the Oracle says must be true (after all, he recognizes that the god cannot lie), he doesn’t accept immediately the testimony of the god, the god being his “epistemic superior”. And this is because Socrates understands the testimony of Apollo as a riddle (Plato, Apology 21b4: αὐτίκεται), given that it seems obvious to him that others are wiser than he is. So he doesn’t deny the authority of what the god says about him.

40 See Plato, Apology 21b5-6; Woolf, 2008, pp. 80-81.

41 Although he seems to know that he knows nothing.

42 At Apology 21c Socrates claims that poets, because of the knowledge they have, thought themselves to be very wise people in other respects too, which they weren’t. Socrates’ advantage over them was the same he had over the politicians and others: he doesn’t know, but he doesn’t believe incorrectly to know either.
to acquire first-hand knowledge of ourselves. He is also aware that the way to correct such deficit is by relying on second-hand knowledge provided by others, since others, when looking to our souls, aren’t susceptible to the first-personal cognitive obstacles that threaten our knowledge of ourselves.\textsuperscript{43} Perhaps we could even say that, given those obstacles, others are in a position to be the primary source of knowledge of ourselves.

Unfortunately, our cognitive make-up is such that we are also unlikely to accept the testimony of others (including that of recognized experts) when they disagree with our self-views. So such source of knowledge of ourselves is often not exploited by us. This of course seems to leave us with a bleak picture regarding the extension of our knowledge of ourselves. Overcoming biases and the like isn’t easy (especially if one isn’t even aware of them), but we shouldn’t forget that Socrates’ story is one of success. And as long as Plato is trying to provide guidance on how to gain such self-knowledge then he must presuppose that it is attainable, although it is hard to attain.

It seems that Plato aims to provide, throughout his dialogues, practical advice as to how to, metaphorically speaking, put one’s own epistemic house in order. Although he doesn’t furnish a clear set of “practical guidelines” to improve both our self-knowledge and knowledge in general, there are some insights that can be taken as part of an implicit orientation to do so. Firstly, there is no philosophy without a “dialogical conversation”. Secondly, if the willingness to maintain a collaborative attitude toward the dialogue is removed, there cannot be progress in the philosophical conversation and in the educative process of which that conversation is a crucial part (Plato \textit{Gorgias} 495a2-9, 500b7; \textit{Republic}, 349a7-8). Thirdly, taking the conversation seriously means not only considering seriously the thesis under discussion but, more importantly, being aware of the fact that one should be responsible for what follows from his own assertions (this explains why one should say what one really thinks. This is the “coherence”–\textit{όμολογία}–urged by Plato everywhere in his dialogues)\textsuperscript{44}. Fourthly, from early Platonic dialogues onwards, Plato declares that to be refuted shouldn’t be understood as an insult; on the contrary, refutation is what guarantees that one can check his own beliefs and, if necessary, to correct them (\textit{Gorgias}, 458a2-b1; \textit{Euthydemus} 295a; \textit{Sophist} 230b-e).

\textsuperscript{43} Of course, as seen, others have all the same first-personal biases and the like that one has but they won’t affect them when considering the other’s case.

\textsuperscript{44} \textit{Gorgias} 461a1-5; 461b8-9; 468e2; 470b1-7. \textit{Th.} 154d-e; 164c8, where it is emphasised that the agreement cannot be merely nominal, like the agreements the contenders are used to entering. For the relation between \textit{όμολογία} and \textit{διαλέγματα} cf. Plato, \textit{Crito} 48b-c, \textit{Protagoras} 322d-e, with the remarks by Brancacci, 2005, p.137-138.
So it seems that for Plato the limitations can be overcome but certainly the path to knowledge of oneself isn’t an easy one\textsuperscript{45}. In this respect too, our knowledge of ourselves is different from our knowledge of our own present mental states. While the former is hard to come by, the latter is normally thought to be easy to come by. Anyhow Plato was right to think that we need the other to know ourselves, given that we are subject to illusions (which are the product of first-personal biases and the like) that preclude such knowledge. But he also realized that some of those first-personal obstacles in fact will stop us from considering the other as a legitimate source of knowledge of ourselves. So achieving such knowledge of ourselves will normally prove difficult, though not impossible, to achieve.

CONCLUDING REMARKS

The path to reach the Temple of Apollo and so to gain the self-knowledge involved in the injunction “Know Thyself” (γνῶθι σεαυτόν) was certainly not easy. Maybe, one might think, we can achieve such knowledge simply by some armchair reflection, as we can in the case of our own present mental states. But the Ancient Greeks weren’t just concerned with knowledge of one’s own mental states. The sort of self-knowledge they were after was knowledge of the self: of one’s nature and character. And looking inward to achieve this self-knowledge isn’t enough. At least Plato’s lesson seems to be that it isn’t that easy. In fact, it isn’t so because one cannot often achieve that knowledge on one’s own. Perhaps one needn’t travel all the way to Delphi to achieve it, but one still needs the other (some external source) to learn about oneself. This is because when we look inward, our lenses might distort what we see (due to the illusions our first-personal biases and the like cause). But, as Plato teaches us, the perspectives of others have much to offer us when searching self-knowledge: others are a source of self-knowledge, or can contribute significantly to that. However, although others can be such sources, they might be sources to

\textsuperscript{45} To put it in slogan form, Plato thinks that the limits of the first- and third-person perspectives can be overcome by the second-person perspective. The distinction between the first- and the third-person perspective is fairly well-established, especially in the philosophy of mind, but the second-person perspective isn’t. So let us just say a few words about it. As suggested above (n.4), perspectival differences are primarily differences in epistemic access, not differences in epistemic objects. The main feature of the second-person perspective, as we understand it, is that involves an interactive process between (at least) two subjects (as in the dialogical conversation). Neither first- nor third-person access share this feature. Indeed, this difference between perspectives can be introduced in a rough-and-ready way by taking the first-person perspective presenting a “subjective” perspective, the third-person-perspective an “objective” one, and the second-person perspective an “intersubjective” one.
which we are blind. Our psychological make-up (including biases and prejudices) might blind us to them and so put such knowledge almost beyond our reach, even in the case of those who are wise, as in Socrates’ case.

To sum up, although Plato doesn’t speak of first- and third-person perspectives in knowledge, he surely considered both the advantages and disadvantages of those viewpoints in examining self-knowledge.

REFERENCES


Burnyeat, M. “Protagoras and self-refutation in later Greek Philosophy”, 

Burnyeat, M. “Protagoras and self-refutation in Plato’s *Theaetetus*”, 

Burnyeat, M. *The Theaetetus of Plato* (Translated by M. J. Levertt. Revision by Myles Burnyeat; Introduction to the *Theaetetus* by Myles Burnyeat), Hackett: Indianapolis. 1990.


Darley, J. Latane, B. “Bystander intervention in emergencies: Diffusion of responsibility”, 


Dunning, D., Meyerowitz, J., Holzberg, A. “Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of ability”, 


Gallois, A. *The World without, the Mind within*. CUP: Cambridge, 1996.


Nisbett, R., Wilson, T. “Telling more than we can know: verbal reports on mental processes”, *Psychological Review*, 84 (1977): 231-259.


