DOI: https://doi.org/10.26512/rfmc.v9i1.41032

Realizing Who I Am

[Darme Cuenta de Quién Soy]

Antoni Gomila *

Abstract: In this paper, the old view of self-knowledge as a practical achievement is vindicated. Constitutivism, the view that connects self-knowledge to rational agency, thus taking a step towards this practical dimension, is discussed first. But their assumption of an epistemic asymmetry that privileges self-knowledge is found mistaken. The practical dimension of self-knowledge, its potential transformative power, is accounted in terms of the interiorization of the concepts acquired in intersubjective interaction.
Keywords: Self-Knowledge. Second Person. Practical Knowledge.

Resumen: En este trabajo se reivindica la concepción tradicional del autoconocimiento como una forma de saber práctico. Para ello, se discute primero el constitutivismo, la posición que dentro de la filosofía anglosajona busca en la dimensión práctica la clave de las propiedades distintivas del autoconocimiento, que es directo e inmediato. Sin embargo, el constitutivismo asume aún la tesis de la asimetría epistémica, que privilegia el autoconocimiento, un supuesto que se muestra incorrecto. Reconocer la dimensión práctica del autoconocimiento, los efectos potencialmente transformadores de la reflexión sobre uno mismo, implica verlo como una capacidad resultado de la interiorización de las capacidades conceptuales desarrolladas en la interacción intersubjetiva, aplicadas a la propia experiencia subjetiva.
Palabras Clave: Autoconocimiento. Segunda Persona. Conocimiento Práctico.

*Full Professor at the Department of Psychology of the University of the Balearic Islands (Spain). E-mail: toni.gomila@uib.cat. ORCID: https://orcid.org/ 0000-0001-7325-961X.

(cc) BY-NC-ND

1.

"Know yourself", said the Delphic oracle. As influentially interpreted by Socrates in the Alcibiades I, it turned out to be a deep thought: a principle of wisdom, a guide for responsible action. It was seen as a criptic statement of the idea that a good life is an examined life, a life lived with a reflexive attitude towards oneself. What's remarkable about it, is that this attitude requires effort, "askesis": it is something to be reached and exercised, a cognitive achievement; something that most people generally do not attain most of the time. Knowing oneself is not easy, nor without consequences; it is not something natural, spontaneous, or that suggests itself out of the blue. On the contrary, self-knowledge is effortful and has a transformative potential. This Socratic approach underlies different Western traditions and practices of "knowing oneself", such as "vitae examinatae", confessions, jesuitic espiritual exercises (of which Descartes adopted their structure in the Meditations), or essays (Montaigne, Bacon). It is also what gives psychotherapy its "healing" bite: through psychotherapy one first understands oneself, so that one can change oneself. This link between self-knowledge and practical reason reached its zenith with Kant's and Fichte's notion of autonomy: a good life is one carried out according to one's understanding of Reason's requirements, and this involves a reflexive attitude towards one's own mental states. Self-knowledge became a basic condition for autonomy, and autonomy, a condition for rational action.

This practical view of self-knowledge, though, receded into the background once self-knowledge, at least some forms of self-knowledge, were distinguished as the most certain forms of knowledge, and so taken as the cornerstone of a foundationalist epistemological project of Cartesian inspiration. While the practical stance towards self-knowledge was recovered in the 20th century by some existentialist writers (most notably Sartre), it also made an impact in angloamerican philosophy in recent years (notably: Moran, 2001; Bilgrami, 2006). It is not my goal in this paper to review the details of this question in the history of philosophy. My goal is rather to contribute to this recovery in order to try to overcome the vexing perplexities of self-knowledge.

Naturally enough, the shift of perspective, though, is not neutral: the "marks" of self-knowledge very much depend upon which epistemic stance is adopted. Ever since Descartes, self-knowledge has been viewed, not as the effortful outcome of a reflection on oneself as an agent, but as the most certain knowledge, one to which we have a privileged access, by means of a process of self-observation. Later criticisms of this model of self-knowledge, which I take here for granted, have led to a de-

flationary account of its marks: instead of infallibility and transparency, it has been attributed the properties of incorregibility and self-intimation (Shoemaker, 1996), traits that are also contentious on empirical grounds, or so I will claim. What is not generally questioned is that self-knowledge is direct (non-evidential) and immediate (non-inferential) in a way that grants it epistemic authority –although it is barely noted that these are phenomenological properties, which, by themselves, say nothing as to how self-knowledge is produced; but lacking an alternative model to the self-observation one as regards its production, it is unclear what the grounds are for the claim of epistemic authority of the first person. My contention is that when a proper understanding of the practical dimension of self-knowledge is achieved, its claimed status as a special kind of knowledge loses its grip. In my view, it is the psychological role of self-knowledge what makes it remarkable, not its epistemological authority. Not its epistemic quality –a sort of certainty-, but the way we get it and its role in agents like us.[1]

To show this, I will first discuss "constitutivism", as a contemporary attempt to go beyond the self-observation model of production of self-knowledge, and to move towards a practical view of self-knowledge, but an approach still concerned with accounting for its directness and immediacy. Next, I will argue that self-knowledge lacks those epistemic marks in general; it is rather a form of reflexive knowledge, grounded in the conceptual abilities acquired in interpersonal interaction.

2.

One trend in this recent shift to a (more) practical standpoint for self-knowledge has been called "constitutivism". Proposed by several writers, notably Wright and Bilgrami (Wright, 1989; 1990; Bilgrami, 1992; 1998; 2006), its main idea is to account for the epistemic authority of self-knowledge in terms of the conditions for rational agency. To my lights, constitutivism is not a satisfactory approach, though. First, because instead of adopting the practical stance to clarify the marks of self-knowledge, these authors reason the other way around; that is, accepting as their starting point the claimed marks of self-knowledge, as characterized from the epistemological stance, its aim is to try to account for these

[1] Cf. Cavell, M. (2006), pp.144: "We have resolved the seeming contradictions between the subjectivity demanded by **self**-knowledge and the publicity demanded by self-**knowledge**, in distinguishing **what** I know from **how** I know it." However, Cavell still assumes an account of the differences in the "know" in terms of immediate vs inferential (see sec. 5). I'll be trying to point out to the reflexivity of self-knowledge, instead of appealing to this taken-for-granted phenomenal properties, heir to the epistemological project that sees also a difference in the "what". Finkelstein, D. H. (2008) also assumes the asymmetry in his defense of expressivism, thus missing the practical dimension of self-knowledge.

(cc) BY-NC-ND

epistemic marks in terms of the practical dimension of the subject that produces them. Second, because in so doing, although they reject the self-observation model of self-knowledge (in any of its guises), they fall short of producing an alternative model of its production from this practical perspective.

Both causes of concern are particularly clear in Wright's case. Wright develops from the later Wittgenstein the thesis that self-knowledge is not a cognitive, epistemological, achievement; he then proposes that it is constitutive of the notion of intentionality and agency that it involves self-knowledge: both go together, because both are central to our successful institutional practices of intentional interpretation. This amounts, for Wright, to place the reports of self-knowledge out of the scope of public assessment, and within the domain of the "subjective", along with the so called "secondary qualities": as judgement-dependent ones. To put it boldly, it is my judging that I belief that p that constitutes my believing that p, just as for something to be green is just to be experienced as green by someone.

There is a sense, to my view, in which this proposal, despite its Wittgensteinian inspiration, seems to be up to its ears involved with the subjective/objective dichotomy which Wittgenstein himself tried to overhaul, and which can (not surprisingly) be tracked back to Descartes. Worse still, making the claims of self-knowledge subject-dependent, as opposed to the objective, factual, ones, not only renders them void of any epistemic value, but also of causal import. Thus, they belong with a scheme of interpretation of what happens from the point of view of the agent, even though what really occurs may in fact be completely subject-less or a-subjective, as you like (a move reminiscent of the Kantian aporia). In other words, Wright's proposal entails an interpretive view of intentionality, as extrinsic to the objective realm of the facts.

Of course, Wright does not embrace such a crude form of Kantianism. He sets some ideal conditions that are required to hold for such a constitutive thesis, and judgement-dependency, to go through. Namely, it is required that the judgement be produced with proper attention, that the subject have the concept involved in the state, and that self-deception be excluded. His idea is that the judgements of self-knowledge are constitutive of the corresponding intentional states by default, that is, unless there is some evidence to cast doubt on them. This is surprising because if there can be grounds -objective, third-personal ones- to doubt a self-ascription (that is, grounds available as well to the subject of the self-ascription herself), then it seems that such grounds should also make such ascriptions right, over and above the "grammar" of such concepts

or the "primitivity" of the constitutive thesis. In other words, if there were no facts of the matter that make our judgements of self-knowledge true, there should not be such facts to make self-deception possible either.

There is something else even more revealing in this hedging self-deception from the ideal conditions. Because the evidence that may discredit a report of self-knowledge as one of self-deception must in principle be non-accessible from the first-person point of view (since if it were, it would not be a case of self-deception, but of cynicism or bad faith, what is not to say that it must be non-accessible to the subject, since it could be from an objective, third-personal, stance, as said before), it follows that judgements of self-knowledge cannot be viewed as authoritative *tout court*: they must comply with intersubjective constraints; they must be acceptable, credible, likely, convincing. Otherwise, they could be challenged, as rationalizations or as lies. To put it another way: what the possibility of self-deception really shows is that the subject has not the last word on the issue of which intentional state he is really in.[2] Therefore, it is difficult to accept that self-reports are judgement-dependent, subjective, as opposed to objective. There is a sense in which they are intersubjective, a sense which needs exploring. More on this later on.

In summary, Wright is concerned with the issue of the epistemic authority of self-knowledge ascriptions, not with the issue of the psychological role such ascriptions may have. But his interpretativist account seems to preempt that such an issue can even arise.

3.

Bilgrami, on his turn, while sharing the constitutive thesis with Wright, tries to go beyond the judgement-dependency account of the latter. So to speak, Bilgrami is not content with simply stating the constitutive thesis as a primitive fact about intentionality, and sets forth the task to explain such a fact. For this goal he recruits Strawson's classic strategy on the issue of freedom vs. determinism (Strawson, 1962): just as freedom must be presupposed for our practices of moral evaluation to make sense, because they amount to holding each other responsible, Bilgrami adds that authoritative self-knowledge must equally be presupposed since to be res-

[2]Wright could point out that Wittgenstein already distinguished between two ways in which a report can be untrue: being false and being deceitful or insincere, while self-reports can only be untrue in the second sense. Two comments are due on this. First, after demolishing the self-observation model of self-knowledge, Wittgenstein suggested an expressive view of self-reports, which makes some sense for qualitative states like pain, but none at all for intentional states, as Wright's proposal reflects (the expressive view is absent in his proposal). Second, and more important, Wittgenstein paid attention zero to self-deception, a kind of "contradiction in terminis" for him, given his expressive view. Accepting self-deception amounts to accepting both senses of being untrue for self-reports.

(cc) BY-NC-ND

ponsible for something we do, we have to "know what we are doing", and to know what we are doing, "we must, in general, know our beliefs and desires; and our intentions, because it is these states which bring about and explain our doings." (Bilgrami, 1992, p. 250).

Although Bilgrami's design seems to face the point about the risk of causal irrelevance I raised as to Wright's account, there are two aspects of this proposal that require commentary. First, that its wording is in the present tense, as if the scope of self-knowledge reduced to currently held states. This is surprising because, on the one hand, were to be so understood, the scope of self-knowledge would conflict with the appeal to the structure of agent responsibility, which clearly extends beyond what one is currently doing. Agent responsibility makes sense for temporally extended subjects, for whom even counterfactual situations matter ("I wish I had not tried to convince her..."). It is not clear to me that Bilgrami's version of the constitutive thesis could be spread out so as to include past mental states, given that nobody claims that memory is infallible, self-intimating, incorrigible or authoritative, and given the well-known fact that autobiographical memory works like a narrator trying to make sense of the past, thus ignoring some events, adding others, embellishing the whole[3]. Moreover, it is worth noting that such self-reports about one's past mental states show the same directness and immediacy as the ones about current states. One is led to wonder whether Bilgrami, as Wright, in trying to account for the epistemic authority of self-knowledge, have failed to address the full domain of self-knowledge judgements, and have restricted to the sub-part that is claimed to exhibit a privileged epistemic condition.

On the other hand, the present tense wording not only suggests that the constitutive thesis applies only to current mental states, but it also further entails that the proper way to conceive of them is as occurrent states, not as dispositional ones. In other words, it comes close to the idea –defended by those who conceive the mind as self-intimating– that such states are intrinsically conscious. This is what the appeal to knowing what we are doing, and for which reason, seems to entail. Again, this is problematic. To articulate the problem in Bilgrami's own way to make his proposal, it is not just current, conscious mental states that are presupposed to be accessible in our practices of moral evaluation, social accountability and personal responsibility, but also habits, tendencies, character features, virtues, personal motives, deeply

[3]See Cohen (1996) as a survey of evidence that shows the lack of reliability of eyewitness testimony, flashbulb memory, and autobiographical memory.

entrenched beliefs, values, life-long intentions, vital goals...: many kinds of in fact dispositional traits, including, of course, intentions. For many of these states, the claim of epistemic authority makes no sense.

Finally, although Bilgrami –in contrast with Wright– clearly defuses an interpretive stance on the intentional, explicitly assuming it to have an explanatory and causal dimension, his argument also misses the point about self-deception, but in a complementary way. For his point about the necessary connection between self-knowledge and responsibility would equally hold for agents whose self-knowledge judgements were always self-deceptive, whose intentional attributions were no more than rationalizations; that is, false to the facts. Indeed, he tries to avoid this consequence by assuming that in self-deception the subject holds both beliefs, that p and that not p, so at least she is right on one of them. But this view of self-deception does not fare well with what is known about our confabulatory capacities (Hirstein, 2006; 2009; Gomila, 2007), our making up of rationalizations on the spot, a phenomenon that doesn't square well with the dispositional nature of the "false" beliefs; it rather suggests that they are "made up" right away, given the context and the dialogical demands. His version of the constitutive thesis, hence, ignores the possibility that a subject may be mostly blind to her real intentions and motives (as it may happen in delusions –importantly for our purposes, it would be shallow, to say the least, to understand them as beliefs that happen to be false: they exhibit the sort of practical involvement characteristic of self-knowledge, its psychological mark).

To close these two sections, then, let me sum up what we have got from our examination of Wright's and Bilgrami's initial efforts at placing self-knowledge in a practical perspective. First, both adopt such a stance while trying to account for the marks of self-knowledge attributed from the epistemic standpoint. But, as I have tried to argue, in so doing they introduce some grounds to doubt that self-knowledge even retains the deflationary marks that they take for granted (namely, first-person authority, a sort of privileged access). The possibility of self-deception and of self-knowledge about one's past mental states and dispositional mental states cast doubt on the traditionally assumed epistemic authority of the first person –it rather seems to be a dialogical authority, an "speaker" authority that can be challenged at any moment. Second, neither of them has explored deeply enough the practical stance on self-knowledge to look for a model of self-knowledge that pays proper attention to the psychological role of self-knowledge: its potential to change the subject's attitudes or identity. The moral of the first point is that we should try to understand self-knowledge for its

(cc) BY-NC-ND

own sake, not from the point of view of its (contentious) privileged epistemic status, especially when we are supposed to have given up on the Cartesian foundationalist project, while looking for an account of its phenomenological attributes, immediacy and directness. The moral of the second point is that a practical approach to self-knowledge may provide the key to make progress in these respects, in shedding light on the role of self-knowledge in the structure of intentional agents like us. At bottom, it may well turn out that self-knowledge is to be seen along with, rather in opposition to, to our knowledge of other minds.[4] I will just suggest at the end a reading of the Greek dictum "Gnothi sauthon" as entailing that the way to find out about the dispositional states that conform what we do is through our (intentional) interaction with others. We become able at self-knowledge as we become able to understand others.[5]

4.

The first moral confronts a very long tradition of thought, one that is easily confused with common sense, or even worse, obviousness: that we are authoritative about our own mental states. The hints I've provided for doubting that we are, are barely convincing in themselves, for sure, so it is pressing to make this claim credible. Thus, I would like to explore in more detail the issue in this section.

The first thing to realize, as already anticipated, is that views of self-knowledge such as the ones discussed above (as well as the ones that, while still within the epistemic approach, share the Wittgensteinian spirit, such as Burge's and Davidson's), gain their initial plausibility from an inadvertent framing of the scope of self-knowledge to occurrent mental states, or more precisely, to avowals, their self-attribution. It could be claimed that it is just these states that we are authoritative about (because of how second-order states are produced, or because of the constitutive thesis or whatever), but such a claim would be insufficient on two counts: first, because this authority is contentious (given not just the possibility of self-deception, but also the psychological literature on the (lack of) reliability of self-reports or avowals); second, because it frames the scope of self-knowledge too restrictively, leaving aside other judgements, equally direct and immediate, about one's past

[4]Notice that, historically, those approaches that have privileged the first-person point of view have encountered a deep "problem of other minds" (rather than "of other's minds"), while those that adopt a third-personal point of view are led to struggle with the problem of self-knowledge. We are advocating a symmetrical epistemological standpoint and contend that the asymmetry is functional.

[5]In fact, the psychological evidence points out that the so-called "theory of mind" is a unitary competence, that the same concepts are deployed both in self-attribution and other-attribution. See, for instance, Gopnik, A. Meltzoff, A.N. (1994).

or future-directed mental states, whose epistemic marks are clearly not equally privileged.[6]

But it is when we turn to non-ocurrent, dispositional, mental states, that the standard view of first-person authority is clearly called into question. Let me remind, first of all, that it was precisely the acceptance of such unconscious states, due to empirical reasons, that cast into trouble the Cartesian transparency thesis about the mind, as well as the infallibility thesis, and forced the retreat to the incorrigibility and self-intimation theses. Unfortunately, the debate about self-knowledge instigated by the challenge set from the externalist view of content has tended to ignore this fact. Of course, it is a very difficult issue how to conceive of mental states such that they can be both occurrent and dispositional. But it is clearly wrong to say that they cannot be both, and to set them apart in terms of conscious awareness, the occurrent ones being self-conscious by definition.

Thus, for instance, it has been contended that we have to distinguish between self-consciousness and self-knowledge (Tugendhat, 1990). But doing so is a way of accepting the different remarks I have been making, thus implying that the current debate on self-knowledge went wrong just terminologically, being self-conscious reports, which are epistemically privileged, but conceding the point on dispositional states and the lack of any epistemological relevance of the self-knowledge of these states[7]. The cost of so doing, however, is to establish a gap between occurrent and dispositional intentional states, so that the latter fall beyond the scope of self-consciousness. I take it that the cost is to high to pay: I can only be self-conscious of a desire to travel to the moon as long as it is an occurrent, active, one –just as I can only be self-conscious of a pain as long as the pain is felt–.

Another way to make the same point is by considering how to make sense of third-personal self-consciousness from this restrictive standpoint (Eilan, 1995). This is something usual among politicians and football-players. Instead of saying "I think...", "I want...", ...they rather say "Obama thinks..." (said by Obama), "Messi wants..." (said by Messi). Given that the epistemic credentials of such statements are assumed to differ from those of the "I-" ones, are we to conclude that they are not

[6]Or, even, for that, matter, about one's bodily states, s Evans (1982) pointed out.

[7]Moran (2001) does something similar, in restricting epistemic authority just to what he calls "avowals", a small subset of self-attributions that constitute self-knowledge.

[8]It is a different question that the "I"-judgement exhibit an epistemic trait called, after Wittgenstein, "immunity to error through misidentification" (See Evans, op. cit.). As it has become clear, this trait is a by-product of the semantic value of "I" as a minimally identificatory indexical. It has nothing to do with the certainty or evidence of the content of the judgment. In fact, Wittgenstein pointed to this trait as the cause of the illusion of a metaphysical self. See Gomila, A. (1999).

(cc) BY-NC-ND

self-conscious? Are they cases in which the speaker lacks privileged access, or even authority, to his own sayings?[8]

The distinction between self-consciousness and self-knowledge, once we give up on the foundationalist project in epistemology and its search for certainty in consciousness, turns out to be an ad hoc dichotomy, a move set to stick to the epistemic privilege thesis, a kind of prejudice when taken apart of the package to which it belonged.

5.

Once self-knowledge is discharged of its traditional epistemic, foundationalist, burden, we still need to account for its phenomenological features: directness and immediacy, its non-inferential, non-evidential, character, the asymmetry with mental attribution to others. We need to find an alternative model to the self-observation one, as well as an account of the practical consequences that "realizing who I am" may have. It is tempting, though, to interpret these marks in still epistemic terms, as signs of its special mode of production and justification, and henceforth, as establishing a gap between self-knowledge and knowledge of others' minds. Thus, for instance, Boghossian says:

> [there appears to be] a profound asymmetry between the way in which I know my own thoughts and the way in which I may know the thoughts of others. The difference turns not on the epistemic status of the respective beliefs, but on the manner in which they are arrived at, or justified. In the case of others, I have no choice but to **infer** what they think from observations about what they do or say. In my own case, by contrast, inference is neither required nor relevant." (Boghossian, 1989)

Boghossian contends that the problem of self-knowledge –in contrast to the problem of other's knowledge- is that only the first is direct and immediate, while the second is inferential. If we are looking for an account of the psychological role of self-knowledge, this difference may appear as an attractive idea. Maybe self-knowledge is not epistemically privileged, but its produced differently, in such a way that this "direct and immediate" mode of production has to do with its practical import. However, I'm going to argue that this way to formulate the problem is misleading, since at the personal level the attribution to others is as non-inferential as self-attribution.

To begin with, I completely agree that we are not consciously aware of any inference to find out what our mental states are, but for that matter, we

ISSN: 2317-9570

are not aware of any inference as regards to memory, vision, or to language comprehension, for example: we do not hear noises and then infer what they mean; the experience is that of understanding them directly. Of course, Cognitive Science has made plausible, theoretically plausible, that all these mental capacities – memory, vision, language comprehension, recognition of emotions... – are sustained by unconscious computational processes which involve transformation of informational structures (or patterns of activations in connectionist networks, or trajectories in the space-state of dynamic fields, make your pick). The debate on Gibson's "direct perception" theory solved satisfactorily this question, to my lights: we see (perceptual) patterns in a robust, immediate, way (what was latter termed "modular" by Fodor), that could be called "direct" even though it involves some kind of processing. This view of perception makes possible that we can "directly and immediately" perceive aspects of our environment, through the mediation of subpersonal processes we do not experience. This view departs from an empiricism view of the mind, though, because empiricism conceives of perceptions as the outcome of a constructivist process; they are built out of sensations or impressions, which are the units which enjoy such directness and immediacy, and play the role of intermediate steps in an inferential process at the personal level. This view is to be resisted: at the personal level, we directly and immediately perceive the world. "Directness and immediacy" are not the marks of subjective states, but of our objective experience of the world.

On the other hand, immediacy and directness are also marks of our everyday attributions of mental states to other people with whom we interact face to face. We do not work, as behaviourists imagined, by getting first all the movements in detail, to infer later what's in the black box (through a theory or through analogical projection of our own inferences) (Gomila, 2001). Our experience is that of directly perceiving the emotions, intentions, and sometimes epistemic states (not just objects, movement, or spatial relations) of others. In other papers, I have called this the standpoint of the second person, which we adopt in face-to-face interaction, and which works in an implicit, practical, way (Gomila, 2002; 2007).[9] Again, this is not to deny that this experience is mediated at a subpersonal level.

Therefore, Boghossian is doubly wrong: it is not true that the thoughts of others are always known through an inferential process at the personal level,

[9]The notion of a "second person" of intentional attribution has become popular in recent years, in a variety of versions. See Gallagher (2001), De Jaegher & di Paolo (2008), Butterfill (2012), Schilbach et al. (2013), Roessler (2014), Gomila y Pérez (2017), Gomila y Pérez (2018), Pérez y Gomila (2018), and Pérez & Gomila (2021).

(cc) BY-NC-ND

and it is not true that self-knowledge proceeds without any grounds, just as a "given". Once the empiricist view of perception is given up, the epistemic asymmetry also disappears. The only difference is functional: the information drawn upon concerns oneself in the case of self-knowledge, because it is about oneself; but this does not mean that there is no perceptual channel is involved. Besides, since it is about oneself, its role in the production of behavior is obviously different from knowledge about another, but in both cases the attribution is of practical interest, and in both cases, it has a "dialogical" authority: it can be challenged in the interaction.

Of course, I am not denying that an objective, third-personal, detached, theoretically mediated, wholly inferential, stance can be adopted towards other people. This is what a cognitive scientist, or a psychologist, tries to do, to get a theoretically grounded account of the minds of other people. But such an approach can also be taken towards oneself. What I am saying is that, as regards face to face, everyday, interactions, this is not the attitude we adopt by default. What we adopt is what I have called "a second-person" point of view, not a third-personal one. I take it that something similar is already implicit in Hume's Treatise. Hume raises the question of the justification of emotional attributions, on the grounds of analogical or inductive inferences, in the Third Part of the Treatise, while in the Second he simply presents the natural phenomena of emotional interaction. "The minds of men are mirror to one another", says Hume, in what it seems to me a graphical way to push this "second person" stance.

The point of this section, then, is to oppose the last resort to keep attributing epistemic relevance to self-knowledge through retaining a phenomenological asymmetry between self-knowledge and "knowledge of other minds". The branches of Descartes' tree have spreaded a lot, and it is easy to get entangled in them, even for people who claim to refuse the Cartesian roots and trunk.

6.

Our challenge, then, is to account for how self-knowledge is produced, as a cognitive achievement, in order to explain its practical import, without relying on the standard idea that this "mode of production" is somewhat mysterious because of its phenomenology. I'm not going to be able to develop such an account in full here, but I'll try to provide the basics. It seems to me that the right strategy is to consider self-knowledge as any other area of cognition. In this regard, it seems to me that self-knowledge structurally depends upon two basic components: a phenomenological self-model (Met-

ISSN: 2317-9570

zinger, 2003; Gallagher, 2005), and the conceptual capacity to adopt a reflexive point of view towards oneself. Let's consider them in turn.

This second notion, that self-knowledge involves a reflexive attitude, which involves the deployment of a set of mental concepts, makes clear the role of the second person point of view in providing the opportunities for mastering these concepts. Mental concepts are acquired in and through intersubjective interaction (Pérez, 2013), which provides broader opportunities for grasping the complexities of mental life by making information contingent upon oneself (Butterfill, 2012). In other words, the second-person point of view is the starting point of mental development, and it provides the conceptual network to be deployed towards oneself. In self-knowledge, one confronts oneself as another, by mastering how other's reactions to one's expressive and intentional behavior help one to understand oneself. The concepts gained in the context of social interaction can be later on deployed in making sense of oneself, by reflecting on the meaning of one's dispositions and reactions.

But, of course, self-knowledge does not depend upon looking to anybody, nor in looking to one's face in a mirror. It is based on our subjective experience, our "phenomenological self" (Metzinger, 2003; Gallagher, 2005), on the particular kind of bodily system we happen to be, with inner senses that make one aware of one's inner workings. This self-awareness is organized at least at two levels: at an ecological level and at a relational level, that are captured by "the ecological self" and the "interpersonal self" (Neisser, 1997). The ecological self-concerns one's bodily awareness, and it shows up in our experiences of initiating action, of being the author of change, as well as the owner of the body that goes through such a change, and as involved in self-regulation. The interpersonal self-concerns the experiences of intersubjective interaction. Both are preconceptual experiences, that reflection may articulate in conceptual terms, once the required mental concepts are acquired through intersubjective interaction. But their deployment requires an effort to make sense of one's experience, just as in interaction one tries to make sense of another's behavior.

Let me illustrate with one example how this double structure works in producing self-knowledge, in the particular case of the reactive attitudes –such as resentment, or guilt, blame, embarrassment.... These emotions involve an assessment of the relationship between me and somebody else, in regard to something I did to them, or the other way around. This assessment is implicitly grounded in a norm or value, but doesn't require an explicit judgement, the valuing is implicit in the emotional reaction. By internalizing the point of view of the other (through a dialogical

(cc) BY-NC-ND

structure, I would say), I may come to realize that my feeling guilty for what I did –that is, making conceptually articulated and explicit an emotion that involves the relational self-, I come to understand those values that move me, I realize why I did what I did. In so doing, I put myself in the position to try to change my attitudes –maybe by being less prone to feeling shame, or by questioning the correction of such a reaction in some circumstances. I apply my conceptual abilities to the experiential, pre-reflexive, level of my mental life. Through such a reflexive attitude I also display the narrative schemas that allow me to make sense, not just of a situation/reaction, but of robust recursive patterns of action and reaction, providing a narrative of temporal continuity.

The moral of this approach is that we acquire the capacity for self-knowledge as we acquire the capacity to understand others and apply this conceptual capacities to our proprioceptive and relational experiences. This "phenomenal self" provides the preconceptual background for conceptual reflexion. In this way, we can account for the idea of intersubjective constraints on "subjective" reports: of rationality, consistency, identity (along time), etc. in so far as they are internalized with the development of self-knowledge. Self-knowledge is therefore not foundational, but revisable as any other form of knowledge. First-person authority turns out to be a kind of discursive authority. Moreover, it implies the possibility that sometimes, others may know better than oneself what one really wants, feels, intends –this discursive authority may be challenged. As a matter of fact, some subjects may barely exercise this capacity for self-knowledge; there is also room for discrepancies among subjects: some may be more reliable than others in their self-reflective practices.

It might be thought that the suggested structure is in fact a new version of the internal self-observation model, in that self-knowledge depends upon deploying concepts acquired in interaction to this bodily phenomenology which appears to be independently given. But this is not the thrust of the proposal. Self-knowledge of self-conscious emotions is not a matter of just looking inside (inside what? to what?) and reading or looking (in which code?). It is rather a matter of making explicit what is implicit in the reactive attitudes themselves, through the adoption of the reflexive attitude: taking oneself as an agent, in contrast to others, to make sense of what one experiences: dispositional states, not occurrent, ones, that just show up, expressively. The "marks" of self-knowledge, therefore, directness and immediacy, would result from this effort at self-understanding, this aiming at controlling the forces that determine our actions, in virtue that it is our own behavior that we are

dealing with. It is my action that I try to shed light on, so that what I judge about myself directly and immediately influences it. Life experience (the common sense notion of experience, not the empiricist one), interaction with others, etc. turn out to be preconditions to acquire this capacity, this ability to deal with oneself.[10]

This approach also makes clear that self-knowledge is a cognitive achievement: one may very well get oneself wrong, this is a common experience. Moreover, it is just agents with complex cognitive skills that are able to exhibit this competence. But I would rather insist that it is a cognitive achievement of practical, not theoretical, import. In other words, the "knowledge" in self-knowledge, is to be understood as wisdom, not as science. Self-knowledge is for a better life, not for quenching intellectual curiosity (Moran, 2001). This practical dimension makes it the case that what we grasp about ourselves adds up to the process of deliberation and decision. The idea here is that the self-knowledge reached by reflection also becomes part of considerations the agent has to consider in her action, besides her other motivational states. Thus, in contrast to constitutivism, this account of self-knowledge accounts for the fact that it can be revelatory, a sort of discovery, of realization, that may bring about changes in how one behaves, which goals one pursues. This is especially relevant when we, as agents, face a conflict, and have mixed, ambiguous, feelings about a situation, etc. Realizing that this is so can settle the stage for overcoming the conflict –or for living it with despair.

7.

The Greek dictum "Know yourself", to sum up, points out to this role of reflection in the structure of action in agents like us. It is not, as I understand Bilgrami, that self-knowledge is a condition of possibility of being an agent; there can be different ways of being an agent, variable in their degree of functional integration and coherence, of akrasia and self-deception, of higher-level control. Thus, I would prefer to say that different degrees of responsibility for oneself are possible, depending on the degree of self-knowledge. Self-knowledge can't be taken for granted as a condition of possibility of agency; agency is possible without it –just as massive self-deception is a possibility, or delusions. Deep motives and objective needs and interests, the values and norms implicitly involved in the "reactive attitudes", the ambiguous emotions they can give rise to when they

[10]Thus, the approach is not to be confused either with purely theoretical, third-personal, view of the mind, such as Carruthers (2011), which overlooks the practical dimension of self-knowledge.

(cc) BY-NC-ND

enter into conflict; life-long intentions and goals, virtues and character, etc. are all cases where self-knowledge can make a difference in practical terms, but it cannot be taken for granted. They are dispositional states of which we are implicitly self-aware through the "interpersonal self", that we can make explicit through reflection. Realizing that they drive our life is the key to modulate them, to take responsibility of them (but not easily self-control them –that's again difficult, an achievement, not a condition of possibility of agency). But it is this complex web of intentional, dispositional, states and relations, causally active in the production of action, that constitutes our identity as agents in the long run, in opposition to transient and cursory volitional and epistemic states, so that one has a say in what kind of person one wants to be, through the reflexive attitude that ponders reasons, solves motivational conflicts, assesses courses of action and ways of life.[11]

References

BILGRAMI, A. (1992): "Appendix: Self-knowledge and Intentionality", in *Belief and Meaning*. Blackwell.

BILGRAMI, A. (1998): "Self-knowledge and resentment", in WRIGHT, C.; SMITH, B.C. MACDONALD, C. eds.: *Knowing our own minds*, Clarendon Press.

BILGRAMI, A. (2006): *Self-knowledge and Resentment*, Harvard University Press.

BOGHOSSIAN, P. (1989): "Content and Self-knowledge", *Philosophical Topics* 17, 5-26.

BUTTERFILL, S. A. (2012): "Interacting mindreaders", *Philosophical Studies* 165, pp. 841-863.

CARRUTHERS, P.F. (2011): *The Opacity of Mind: an integrative theory of self-knowledge*. Oxford University Press.

CAVELL, M. (2006): *Becoming a subject. Reflections in Philosophy and Psychoanalysis*. Clarendon Press.

COHEN, G. (1996): *Memory in the real world*, Psychology Press.

DE JAEGHER, H. and DI PAOLO, E. 2007. Participatory sense-making: an enactive approach to social cognition. *Phenomenology and the Cognitive Sciences* 6: 485-507.

EILAN, N. (1995): "The first person perspective", *Proceedings of the Aristotelian Society* XCV, 51-66.

EVANS, G. (1982), "Self-identification", in *The Varieties of Reference*, Oxford U.P., pp. 205-233.

FINKELSTEIN, D. H. (2008): *Expression and the Inner*. Harvard University Press.

GALLAGHER, S. (2001: "The practice of the mind: theory, simulation, or primary interaction", *Journal of Consciousness Studies* 8, pp. 83-108.

GALLAGHER, S. (2005): *How the body shapes the mind*. Oxford University Press.

GOMILA, A. (1999): "Jo, L.W.", *Taula* XIX.

GOMILA, A. (2001): "La perspectiva de segunda persona: mecanismos mentales de la intersubjetividad", *Contrastes*, supl. vol. 6, pp. 65-86.

GOMILA, A. (2002): "La perspectiva de segunda persona de la atribución mental", *Azafea* NS, 1, pp. 123-138.

GOMILA, A. (2007): "El retorno de la represión", *Teorema* 26, pp. 97-112.

GOMILA, A. PÉREZ, D. (2017): "Lo que la segunda persona no es". In En PÉREZ, D. y LAWLER, D. (eds.) *La segunda persona y las emociones*. Buenos Aires: Editorial SADAF, pp. 275-297.

GOMILA, A. PÉREZ, D. (2018): "Mental attribution in interaction: how the second person perspective dissolves the problem of other minds", *Daimon*, 75, pp. 75-86.

GOPNIK, A. MELTZOFF, A.N. (1994): "Minds, Bodies and Persons: young children's understanding of the self and others as reflected in imitation and theory of mind research". In TAYLOR, S.T.; MITCHELL, R.W. BOCCIA, M.L. (eds.), *Self-Awareness in Animals and Humans: Developmental Perspectives*. Cambridge U.P.

[11] Without this reflexive attitude, phenomena like religious conversions, "falling off the horse", radical changes of life, cultural migration or self-irony would appear as mysteries.

HIRSTEIN, W. (2006): *Brain fiction: self-deception and the riddle of confabulation*. MIT Press.
HIRSTEIN, W. (2009): *Confabulation: views from neuroscience, psychiatry, psychology and philosophy*. Oxford University Press.
MORAN, R. (2001): *Authority and Strangement: an essay on self-knowledge*. Princeton University Press.
NEISSER, U. (1997): "The roots of self-knowledge: perceiving Self, It, and Thou". In SNODGRASS, J.G. THOMPSON, R.L. (eds.) *The Self across Psychology: self-recognition, self-awareness, and the self concept, Annals of the New York Academy of Sciences*, vol. 818, pp. 19-33.
METZINGER, Th. (2003): *Being No One: the Self-Model Theory of Subjectivity*. MIT Press.
PÉREZ, D. (2013): *Sentir, desear, creer: una aproximación filosófica a los conceptos psicológicos*. Prometeo Libros.
PÉREZ, D. GOMILA, A. (2018): "La atribución mental y la segunda persona". En BALMACEDA, T. PEDACE, K. (eds.) *Temas de Filosofía de la Mente: Atribución Mental*. Buenos Aires: Ed. SADAF, pp. 69-98.
PEREZ, D. GOMILA, A. (2021): *Social Cognition and the Second Person in Human Interaction*. London: Routledge.
ROESSLER, J. (2014): "Reason explanation and the second-person perspective", *Philosophical Explorations* 17, pp. 346-357.
SCHILLBACH, L., TIMMERMANS, B., REDDY, V., COSTALL, A., BENTE, G., SCHLICHT, T., VOGELEI, K. (2013): "Towards a second-person neuroscience", *Behavioral and Brain Sciences*, 36, pp. 393-462.
SHOEMAKER, S. (1996): *The First-Person Perspective and Other Essays*, Cambridge University Press.
TUGENDHAT, E. (1990): *Self-consciousness and self-determination*, MIT Press.
STRAWSON, P.F. (1962): "Freedom and resentment", *Proceedings of the British Academy*, XLVIII, 1-25.
WRIGHT, C. (1990): "Wittgenstein's rule-following considerations and the central project of theoretical Linguistics", in GEORGE, A., ed.: *Reflections on Chomsky*, Blackwell.
WRIGHT, C. "Wittgenstein's later philosophy of mind: sensation, privacy and intention", *Journal of Philosophy* 86.

Received: 02/02/2021
Approved: 12/03/2021
Published: 30/04/2021