

SELF, SIGN, BELIEF, AND AGENCY

Waldomiro J. Silva Filho

Universidade Federal da Bahia, CNPq (Brazil)
waldomiro.silva@pq.cnpq.br

Maria Virgínia Dazzani

Universidade Federal da Bahia (Brazil)
dazzani@superig.com.br
dazzani@ufba.br

Abstract

This article presents a discussion of notions of self and belief from a semiotic approach that integrates theories of action. In spite of the semiotic nature of subjective experience and likewise of the incompleteness of our (self-) understanding, the weaknesses of our reason, the events of our lives, we reflect and act in the light of our beliefs and reflections. Even if the person can assume multiple positions in the same flow of communication, reflection is present at moments in which that person faces situations wherein inquiring him/herself about his/her belief becomes unavoidable for the conduct of the person's life. To reflect is nothing but to expose one's own attitudes under a certain angle, considering one's own contingencies, changes, precariousness and uncertainty. From a semiotic point of view, to reflect is to interpret one's own flow of signs in (but not in the idealistic sense, since it also involves emotions and feelings) in order to act.

Keywords: Agency; Belief; Self; Self-knowledge; Semiotic.

Resumo

Este artigo apresenta uma discussão acerca das noções de self e crença a partir de uma abordagem que integra teorias semióticas e teorias da ação. O argumento central é que, no cenário da vida comum, o self-awareness e a agency deveriam ser concebidos como um processo semiótico e prático. Considerando o caráter semiótico da experiência subjetiva e, do mesmo modo, a incompletude do nosso



(auto)entendimento, as fraquezas da nossa razão, as vicissitudes da nossa vida, mesmo assim, refletimos e agimos à luz das nossas crenças e reflexões. Mesmo que a pessoa possa assumir múltiplas posições no mesmo fluxo comunicativo, a reflexão está presente no momento em que essa pessoa se encontra em situações em que a pergunta a si mesmo sobre sua crença torna-se inevitável para a condução da sua vida. Refletir não é mais do que expor as próprias atitudes sob um certo ângulo, considerando as próprias contingências, mudanças, precariedades e incertezas. De um ponto de vista semiótico, refletir é interpretar o próprio fluxo sógnico do pensamento (mas não no sentido idealista, envolvendo também as emoções sentimentos) com o objetivo de agir.

Palavras-chave: Agência; Crença; *Self*; Autoconhecimento; Semiótica.

Self and Subjectivity

Salgado & Hermans (2005) describe how in the recent history of ideas in Philosophy and Psychology, the concept of self has undergone a significant shift on the influence of dialogism and the socio-cognitive perspective. Indeed, in the twentieth century, different traditions of Psychology, Psychoanalysis, Philosophy, Social Sciences and Cognitive Sciences criticized the modern conception of subjectivity and self. According to the modern perspective (post-Cartesian), the "Notion of self is related to, one might say, and constituted by a certain sense (or perhaps a family of senses) of inwardness" (Taylor, 1989, p. 111). Moreover, according this perspective, it was conceivable that "we have selves the way we have heads or arms, and inner depths the way we have hearts or livers, as a matter of hard, interpretation-free fact" (Taylor, 1989, p. 112). Much of the research on the nature of the human mind, in a way, depended on this model, as for example the idea of mental representation of the external world (in epistemology), self-awareness, personal identity and deliberation (in moral philosophy) and of course, the very idea that there is a special category of properties in humans that makes them persons and not just natural beings (in Metaphysics).

The core of Salgado & Hermans's (2005, pp. 6-7) argument is that some criticism

on the modern notion of self, when defending argue that the self is not an entity or metaphysical substance and by relying on tools of the language theory, of communication and socio-historical constructionist, they eventually sacrifice their own idea of subjectivity. This is the case, for instance, of the theory of multiplicity of selves (Gergen, 1991): the meanings that identify the self become completely dependent of the socio-linguistic negotiation between individuals and these relationships occur in many different contexts which allow meaning games between individuals. Self (or selfhood) is no longer understood as a (necessary) individual's instance or property, becoming just a particular case of intelligibility, that is to say, what an individual is must be exclusively thought of as what he/she symbolically and pragmatically negotiates in communicative contexts. Moreover, in these games individuals can take multiple positions. The diversity of communicative contexts and positions that individuals occupy in communicative exchanges creates a "multiphrenic self", which is "a self that needs to deal with different ways of meaning-making." (Salgado & Hermans, 2005, p. 7).

According to Salgado & Hermans (2005), the theoretical issue here is that this point of view actually expands the sense of self, but has as a consequence the fact that the notion of subjectivity becomes meaningless, because there is a radical shift from the subject (from speech) to the web of language and discourse. If the subject is just a diluted position in a web of communication, then it makes no sense to speak of something like subjectivity or constituting traits the subject. Notwithstanding, Salgado & Hermans (2005) believe that subjectivity is still a topic that should be considered and theorized and, therefore, the dialogical self theory presents itself as an alternative to both the traditional conception (Cartesian) of mind and self and to the perspective of multiple selves (Salgado & Hermans, 2005, pp. 9-10). In other studies, Salgado & Ferreira (2005) and Salgado & Gonçalves (2006) proposed a triadic dialogical model which aims to describe subjectivity as a communicational process which integrates three elements: a) I (as the center of the experience), b) *another* to whom the I addresses and c) a third diffuse and invisible element, which may operate as a mediator, denominated by Salgado as "*potential audiences*" (Salgado, 2006, p. 150). The formulations elaborated by Salgado (2006) and other dialogical self theorists are very interesting, stimulating and complex, but what we want to highlight here is precisely the fact that there is a perspective about the self that preserves the notion of subjectivity as "the personal side of a communicational process", that is distinct from the modern tradition which conceived the self as "the foundation of knowledge"



(Salgado & Hermans, 2005, p. 10).

We have, then, a double-sided question: in what sense can we speak of a personal side of the communicational process and what consequences can we draw from this theoretical idea to interpret the notion of subjectivity?

To answer this question, we follow two paths: on the one hand, we discuss the semiotic aspect of subjective experience and on the other, the relationship between belief and agency. With this, we will seek, in the final section of this paper, this convergence between semiotic sense of subjectivity and the notion of practical action.

Sign, Self, and Subjectivity

One of the central aspects of Charles S. Peirce (1931-35; 1958) criticism to the spirit of cartesianism is related to the tendency to think the human mind as absolutely private, self-centered and independent of social and natural processes. Peirce questions the image of a self that has a pure intuition of objects and of itself (something like a faculty of intuition of intuition). According to the broad intellectual tradition criticized by Peirce, we must establish a description of psychic life from a direct understanding of psychic or spiritual facts and assume a primary and privileged aspect self-consciousness. An introspective psychology would be the epistemological foundation and intuitive awareness of the self (the ability to intuitively distinguish between different activities of the spirit, such as perceive, think and dream) would be the primary source of evidence and truth. Thus, self-awareness would be separated from general consciousness, preceding it logically and ontologically; thus, self-awareness – the intuitive awareness of consciousness - would be the first in time and reason.

According to Peirce (1868), the problem is that in order for the awareness of consciousness to fulfill its function, it should be unrelated to any concept and belief and could not be influenced by any other cognition:

Self-consciousness, as the term is here used, has to be distinguished both from general consciousness, from the internal sense, and from pure apperception. Any cognition is a consciousness of the object as represented; by self-consciousness we mean a knowledge of ourselves. Not a mere feeling of subjective conditions of consciousness, but of our personal selves. Pure apperception is the self-assertion of THE ego; the self-consciousness here meant is the recognition

of my private self. I know that I (not merely the I) exist. The question is, how do I know it; by a special intuitive faculty, or is it determined by previous cognitions? [Emphasis added by the author] (CP 5.225)ⁱ

Embarrassing is the fact that the consciousness we have of ourselves as individuals is not of an intuitive nature. For him, it would be impossible to establish a rationale (about objects and about oneself) that departs from a principle which has not been formed in one's real life, made concrete by education and habits. Self-consciousness is a discourse that integrates passed data (what we were in the past according to the testimony of memory, documents, photographs and reports of others) and anticipations ("I'll have lunch tomorrow with Mary and John") and is actually, a discursive construction deriving from various sources. Using a common expedient to Wittgenstein - child pseudo psychology - Peirce says that is not observed in children traits such intuitive self-awareness as a requirement for knowing things (CP 5228). The child expresses capabilities to think well before being aware of him/ her. This acquisition of capacities is mainly due to the discovery of the central relationship of his/her body, and especially the experience of ignorance and error in the interaction with the world and with other men and women.

It is through the mediation of language taught to the child, that he/she becomes aware that others' words classify and categorize things; in this same movement, the understanding of his/her ignorance and his/her mistake, which takes place when the child realizes that the warnings of adults are confirmed by experience and their spontaneous anticipations are invalidated, the child's self-image starts to be built.

The other's language and act of witnessing provides the "first draft of self-consciousness" (CP 5.233). The child has the experience of error, and error can only be explained only assuming that one I is the subject, and the place of error and accuracy, the real and the unreal, truth and falsehood, is something that applies not only that subject, but to all subjects: "Ignorance and error are all that distinguish our private selves from the absolute *ego* of pure apperception" (CP 5.235). Or, what is worse, the infallible certainty, immediately resting on the authority of the first person of singular, when interdicting the possibility of error, forbids correctness and truth: "Supposing that a man really could shut himself up in such a faith, he would be, of course, impervious to the truth" (CP 5.214). The awareness we ourselves have is inexorably an inferential feature (even if we adopt inferences that lead us to consider such consciousness is immediate). How do we resort to intuitive self-awareness as the



first knowledge if our consciousness is the result of conjectures and hypotheses acquired by education and coexistence with others? It would be comfortable if we had the ability to separate intuitively what we conceive and believe, what we learned and what comes to our experience:

...we believe and what we merely conceive, we never, it would seem, could in any way distinguish them; since if we did so by reasoning, the question would arise whether the argument itself was believed or conceived, and this must be answered before the conclusion could have any force. And thus there would be a regressus ad infinitum. (CP 5.239)

How to end this vicious circle?

However, one cannot deny that we have some knowledge of our mental life, but knowledge of the internal world (imagination, emotions etc) passes inexorably through the knowledge of what was outsourced and meant, i.e., of what represents something to the interpreter - the inner facts can only be learned through concatenation of external facts:

If we seek the light of external facts, the only cases of thought which we can find are of thought in signs. Plainly, no other thought can be evidenced by external facts. But we have seen that only by external facts can thought be known at all. The only thought, then, which can possibly be cognized, is thought in signs. But thought which cannot be cognized does not exist. All thought, therefore, must necessarily be in signs. (CP 5.251)

Here, sign is understood as anything that under certain aspects and qualities represents something for someone, creating in the interpreter's mind (CP 4536) representation of a representation, or rather, a new sign equivalent to the first sign, or more developed than it. It is a serious mistake to imagine that, for Peirce, thought is a spiritual and self-sufficient reality (which, uses language voluntarily and arbitrarily to express or communicate); and sign a mental representation. If we consider the case of a conventional sign, such as words, it may seem that it requires a previous inchoative thought in order to assure the meaning. Peirce explicitly rejects the idea that thought antecedes signs. Every thought is in itself symbolic in nature, consisting of the manipulation of signs. To Peirce, indeed, the sign needs something previous to attach its meaning, even though the preceding thing is also a sign. The conventional sign

interprets earlier signs and is interpreted by other signs following it: there is no sign that does not require some form of interpretation and does not refer itself to a sign. Therefore, the sign is not a representation in the sense of the mentalist conception of empiricism, but a relationship that involves language, the world, the body, the rule, what is general.

The great distance between Peirce and linguistic theories of the sign is mainly due to the fact that he is busy with a theory of knowledge and a theory of mind to solve the mind-world dichotomy, diluting the corresponding concept of representation and at the same time, preserving the positive sense of the real knowledge. The lesson he drew from nominalist Ockham and Scotus, Peirce knows that thought is addressed to the particular and singular in the real experience of external events, but does so resorting to relationships, connections, series, signs and inferences. Thought can only be symbolic because it cannot be but a thought of laws and general terms. Like Aristotle said, based on the pure particular and singular, one concludes nothing, nothing can be deduced, and nothing is known. There is in Peirce a fallibilist conception of knowledge: he frankly acknowledges that in all our claims to knowledge remains an element of chance unchecked and therefore the order that directs the formation of our discourse about things depends on the order of conviction, notions and concepts from which we depart. That's why every case of apparently straightforward and intuitive awareness (e.g., awareness of the most basic data present in spirit) has the nature of a hypothesis because every pretense of knowledge involves an assumption that apply in relation to an object or a series of objects from a certain method to classify, systematize and articulate (Gallie, 1970, pp. 63, 66).

At the pure immediate present, there is no pure thought and, likewise, no operations on intuition of consciousness that resemble normal use of the concept of thought. By thought or knowledge, Peirce believes an understanding, an intellection, a work of establishing relationships and chains that takes time, and not a mere presence. There can be no actual knowledge of the immediate present, or of the absolutely immediate:

From the proposition that every thought is a sign, it follows that every thought must address itself to some other, must determine some other, since that is the essence of a sign. This, after all, is but another form of the familiar axiom, that in intuition, i.e., in the immediate present, there is no thought, or, that all which is reflected upon has past. Hinc loquor inde est. That, since any thought, there must



have been a thought, has its analogue in the fact that, since any past time, there must have been an infinite series of times. To say, therefore, that thought cannot happen in an instant, but requires a time, is but another way of saying that every thought must be interpreted in another, or that all thought is in signs. (CP 5.253)

Compliance with the external facts, far from metaphysical realism, refers to what we call provisionally semiotic externalism or pragmatic realism. Now here is a hermeneutic imperative which states that every thought determines another thought which interprets it, and conversely, every thought assumes another of which it is an interpretation. This is the basis of Semiotics and of Peirce's realism, namely that every thought is a sign given by another thought and, in turn, determines another.

Peirce also asks whether a sign can be meaningful if, by definition, is a sign of something absolutely unknowable (CP 5.254). To Semiotics, then, to establish the theoretical impossibility of any conception of the object epistemologically unknowable and non-signic escapes all our forms classifying and thinking. All our conceptions are obtained through abstractions and combinations of cognitions first occurring in judgments of experience. Therefore, there can be no conception of the absolutely unknowable, since none of this occurs in the experience. But the meaning of a term is the conception it conveys. Therefore, a term may not have such a meaning (CP 5255).

The term unknowable comes from an unseen contradiction because any idea that defines something apart from a set of possible knowledge (based on the current stage of knowledge) is the result of an abuse of language. Every thinkable is conceivable, and what is conceivable in reality must be theoretically knowable and signifiable: if something is part of the nature of signs and meaning, it is therefore a knowable concept. But, intuition would operate with what is absolutely unknowable, because it would deal with what is not part of language and previous knowledge. But the unknowable (if there is something like that) just does not have any relation with thought, much less with the formation of a solid foundation for building an image of what is real and true.

Because, in general, we do not have the power to intuitively distinguish a way of intuitive knowledge of the non-intuitive, we cannot, likewise distinguish this knowledge which is present to our consciousness, what is given and what is built, but by understanding with the help of elements extracted from inference. Each event known – for example, the outbreak of a new experience – is not instantaneous, but an ongoing process which is uninterrupted interpretation of thoughts by other thoughts and fusion

of impressions of representations that give unity and direction. To indicate the cognition by which a knowledge or sign was determined is to explain and give meaning to the determinations of that cognition (CP 5260): this is the only way to explain the knowledge of oneself, of things and of the mind of others.

Thought and Sign

The idea that there is no thought without sign, and that sign is interpretation and inference, undermines the foundational value of pure intuition and intuitive awareness of self and shatters the strong sense of subjectivity. Thus, the self can be understood in the flow of language (or semiosis), not as an instance or entity, but as an individual's position in this semiotic stream in the interaction with the natural world, with culture (including the narrated past and uses) and with other people.

Thus, the criticism of the classical conception establishes another status for subjectivity and self. The nerve of this semiotic solution of subjectivity was determined by the third inability, presented in *Some Consequences of Four Incapacities* (Peirce, 1868): the inability to think without signs. In other words, it means that thought is composed of a network of logical-psychological states, which are signs or have the function of a sign: thought-sign. Peirce insists on saying that these signs are phenomenal manifestations of ourselves, so that when we think, "ourselves, as we are right now," i.e., in the modification that affects us, "we are like a sign" (CP 5283). The individual in the environment of the natural and social world joins the process of semiosis as subject of thought, emotions, and actions:

It is that the word or sign which man uses is the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thought, proves that man is a sign; so, that every thought is an external sign, proves that man is an external sign. That is to say, the man and the external sign are identical, in the same sense in which the words homo and man are identical. Thus my language is the sum total of myself; for the man is the thought. (CP 5.314)

First of all, the man-sign and thought-sign are terms that hold the doctrine of categories: no immediaticity as such is element of thought, nor pure quality, nor the pure existing singularity. The spirit-mind is a sign developing according to the laws of inference, because in Science and Moral, and also in Aesthetics, when we adopt



beliefs, our words publicly learned with other men and women justify our relationship with our objects (those who in a Cartesian way would be within us and would only be known by us through an introspective self-examination). All that belongs to I (all inner ontology) is first signified in the public fabric objectivity.

The infancy of subjectivity (if there is something like that) can only be understood in this fabric: according to the rule, to the belief, to the objectivity, to the sign. How did it appear in *What Is Pragmatism?* We read:

Two things here are all-important to assure oneself of and to remember. The first is that a person is not absolutely an individual. His thoughts are what he is "saying to himself," that is, is saying to that other self that is just coming into life in the flow of time. When one reasons, it is that critical self that one is trying to persuade; and all thought whatsoever is a sign, and is mostly of the nature of language. The second thing to remember is that the man's circle of society (...), is a sort of loosely compacted person, in some respects of higher rank than the person of an individual organism. It is these two things alone that render it possible for you (...) to distinguish between absolute truth and what you do not doubt. (CP 5.241)

It is not the purpose of this chapter a careful analysis of the self semiotic theory and its current developments in psychology. What we want is to highlight, on the one hand is the place of belief and the sign in the economy of the self and, on the other hand, the nature of the relationship between individuals and these beliefs in the course of their working lives, agency, as actions oriented by beliefs and based on a type of responsibility and implication which makes this belief or this sign the other person's belief or sign.

Belief and Content

Consider a situation in which a person can take different I-positions, as in the classic studies of Dostoyevsky on Bakhtin (Bakhtin, 1929, p. 5-46) which is the recurring source of dialogical self theory (e.g. Hermans, 2004) "... the dialogical self is based on the assumption that there are many I-positions that can be occupied by the same person" (Hermans, 2001, p. 249). However, in these circumstances, one can only be identified – even if temporarily – if one says something with the strength of the first person (first person as in grammar).

In order for a person to join a communicational process, he/ she must be able to articulate sentences in the first person that convey content, it is necessary that we give this person on one hand, the ability of self-identification and on the other, the mastering of the language. First, if a person does not have these skills, if he/she cannot distinguish him/ herself from others in the continuum space-time (people, animals and objects), we cannot ensure that he/she is properly using the pronoun I. Second, because if it does not dominate the broad aspects of language, we cannot say that he/she is saying anything or is stating a meaningful sentence. The absence of these two skills disables the person to participate in the dialogical language game.

From the conceptual point of view and with regard to the very idea of self that makes all the difference, because in our common experience we almost always refer to ourselves as persons under this award of a special ability that enables us to participate in communicational process. We have no reasons to suspect that the fellow next to us, our former teacher, or the waitress are not people, that they are zombies or robots – we also do not think this absurdity about ourselves. Anyone who has watched science fiction movies knows that the difference between humans and zombies is not in the way these two kinds of beings behave, speak, walk, serve coffee or explain a theorem. Unlike in the movies is possible that the good guy is fooled until the last minute without realizing that the person next to him is not a person. The difference is very simple: people think, believe, wish; zombies are guided by an external force and have no desire, do not think, do not believe; and of course, we assume that our colleague, teacher and waitress think, believe, desire.

The notion of a person requires the idea that we can give the individual the ability to properly use the pronoun I in phrases that indicate the individual's own events and states, such as feelings, thoughts. Events and mental states such as beliefs, desires and thoughts are called, according to the philosophical jargon, intentional mental states that have content or meaning. In a daily basis, in our ordinary psychology, we refer to these mental states using phrasal expressions such as "John believes that water quenches thirst," "I wish my team were champion." Philosophers call such mental states propositional attitude because they express a specific attitude of a subject in relation to a specific proposition (Richard, 1990). Therefore, we have: a belief that [the water quenches thirst], the thought that [things are not going well], the desire that [our team wins], the fear that [the economic crisis worsens]. Thus we can speak of the semantic content or the meaning of belief, desire etc. Moreover, when we believe and wish, we almost always believe in something or want something. Other mental states



such as sensations, pain, tickling are not intentional mental states and have no semantic content.

In any case, the proposition is that (and only that) of what you can say it is true or false. Moreover, the proposition has an informational character: "If I know something and communicate it to you (in English, French or through a gesture or drawing) what you get is the proposition that I know." (Dennett, 1982, p. 6). The terms that make a proposition have, on one hand, an intension or a semantic brand or simply meaning (chair has the semantics brands of, or means "that which serves to sit," "be a household utensil" etc.) and, on the other hand, an extension, namely the objects that can be referred to by the word (as in the case of many possible objects that can be referred to when we use the word "chair"). A person who is in the mental state of believing and uttering a declarative sentence that expresses this belief must master these semantic aspects of the sentence pronounced.

The strongest trait of the proposition is that it can be grasped by more than one mind. Different beliefs – like different thoughts – denote different mental states. This may seem a truism, but, as we shall see when dealing with externism or anti-individualism, it is crucial to evaluate the attribution of rationality through attribution of intentional mental states. This is because if we accept that the object of a belief-thought is a proposition and that the proposition, in turn, is compositionally understood, that is, its semantics depends on the semantics of its constituent terms (signs-words that appear in the sentence), then we also have to accept that having a belief involves mastering the meaning of words that express that belief.

Likewise, if we want to understand what role played by belief in the psychological life of a rational agent, we should investigate how the person understands, masters and uses the content of his/her beliefs in relation to how the agent understands and masters the meaning of his/her words and phrases. It is worth mentioning that when we are dealing with what a person believes we are identifying (or specifying) the content of his/her belief. Similarly, when we ask about what a declarative sentence means we are asking about how to identify and specify the semantic content of the sentence. The question about the content of the belief and the meaning of the phrase expresses a dual relationship between the person and the proposition and between the phrase and the meaning (McGinn, 1982, p. 207). In the context of the attribution of mental states, these two paths intersect.

To conceive that a person may not have beliefs or thoughts, or that this person may unreasonably act, without motivation of beliefs and thoughts, is usually the same as to say that this person does not act rationally. Therefore, to speak of a person means, in regular situations, to consider someone as a rational agent, because in addition to feeling naturally inclined to assign thoughts and beliefs to that person, believing that these thoughts and beliefs influence or even explain the person's actions and behavior, we also feel naturally inclined to believe that this person can understand, know the content, critically and reflexively evaluate and consider his/her own thoughts and beliefs and guide his/her actions in light of that understanding. In this sense, if there is no transparency of attitudes, or if this transparency fails, we cannot say that a person knows what he/she is talking about, or knows the reasons for his/her actions and, therefore, we lose an element that is considered crucial to establish the characteristic of human action, namely rationality and sense of responsibility and choice.

The relationship between attribution of beliefs to one person and the understanding that this person has of the content of his/her beliefs and rational action are in the center of the debate on self-awareness and rationality. Self-knowledge is that knowledge "one has of one's own states – e.g. knowledge of one's own current experiences, thoughts, beliefs or desires" (Gertler, 2011, p. 2). This knowledge would enable the person to discriminate his/her own thoughts and excerpt a unique authority in relation to them. Some authors still think that if there is a consistent argument evincing that there is no self-knowledge in this sense, and accordingly no first-person authority, our personal image becomes deeply shaken.

Transparency Condition

Often the discussion about the notion of the self is treated exclusively as a theoretical and epistemological point of view. However, following the line of reasoning adopted in this article, we consider that the theme of self is especially relevant when we are dealing with situations in which the person is the subject of a lawsuit. We understand that action as intentional behavior is guided by reasons of beliefs and will (Sobel & Wall, 2009; Stout, 2005). This is the context of "epistemic agency". A key issue of epistemic agency is: what is the specific feature of the relationship we have with our own mental states as opposed to the relationship we have with the mental states of others?



For Moran (2012, p. 212), some questions are crucial to understanding the relationship between self and agency:

We may ask, first of all, how self-knowledge matters to agency itself, that is, whether the specifically human forms of rational agency can be understood apart from the capacity for self-knowledge of the mental life that is expressed in that agency. Is our capacity to act for reasons, to be self-guided in that sense, dependent on our ability to know our mental life 'immediately'? Can the ordinary ability to respond to reasons in one's thinking, to consider reasons for and against some belief and respond accordingly, be understood apart from our capacity for immediate self-knowledge? And would the absence of the ordinary capacity for self-knowledge make no essential difference to our rational agency? And from the other direction we can ask how rational agency itself may matter to the understanding of self-knowledge; that is, whether the ordinary capacity to know what one thinks about something is part of the same capacity to determine one's thought about that thing. Is our ability to know what we believe 'immediately', and with a kind of authority not shared by what we say about the beliefs of others, tied to the fact that our beliefs and other attitudes are expressions of our rational agency, and is there a notion of responsibility applying to a person's relation to her attitudes that is related to the capacity for first-personal knowledge of them?

We can say that an important aspect of our mental life is the fact that people have the legitimate right (entitlement) to consider and reflect on their own beliefs. And that possibility to consider their own beliefs often plays a role in the conduct of their lives and changes their mental states.

The idea of mental transparency often appears in the discussion of self-knowledge and, almost always, is associated with two different conceptions, which we will call Russell's principle and Evans' principle. Gareth Evans (1982, p. 89) coined the expression Russell's Principle. According to the former, the principle states, "a subject S cannot form a judgment on something if he does not know what that of which the judgment is about." The subject must have a discriminative ability to distinguish the object of his/her judgment from other different objects. In addition, to Bertrand Russell (1910-11), the understanding of simple facts requires first that we have a direct knowledge of our mental states. What makes this principle essential is the intuitive fact that it seems paradoxical that we can have a thought without knowing what we have

what it is all about. Russell's principle implies what Timothy Williamson (2000) called luminosity, that is to say that a condition for the subject to have a belief is the fact that he also has clear, distinct and infallible knowledge (second order knowledge) on the content of this belief. From this point of view, something like a method of acquiring this type of knowledge is at stake, for a person would know the content of his/her mental states without the benefit of an empirical investigation, introspectively. To talk about our beliefs, desires and thoughts would require a look inside ourselves.

Evans' principle offers a perspective on transparency which does not involve introspection or brightness. For Evans (1982), in the position of first person of the present (in grammatical terms), understand the meaning of having beliefs involves a reference to the same reasons that might be required to understand what this belief is all about. That is because my belief in *p* is equivalent, for me and for common conditions, to my belief in the truth of *p*. If someone asks "Do you believe that the global crisis will affect the Brazilian economy?", when replying, the subject does not investigate a phenomenon that occurs in the inner space of his/her mind, looking inward. Instead, he/she directs his/her attention to the facts and economic conditions in Brazil and to the information which can be accessed in the newspapers. Therefore, he/she will be treating the same phenomenon the same way if he/she would have to answer the question "Will global crisis affect Brazilian economy?" For Evans, if the subject is in a position to assert that "p...", then he/she is in a position to say "I believe that p..." (Evans, 1982, p. 225-226).

Considering this, Richard Moran proposes what he calls the condition of transparency (Moran, 2001, p. 62), which could be presented as follows:

(CT) The question of the first person present tense about his/her own belief ("Do I believe that p?") is answered in relation to the same reasons that justify an answer to the corresponding question about the world (about the truth of p).

For Moran (2001), if we accept the condition of transparency, we can understand the most striking feature of the self, namely, the prospect of asymmetry between first and third person.

But we can ask ourselves in which situation it is relevant to mention that we have (or not) self-knowledge? If we were to use a Wittgensteinian parlance, we would question: in what language game makes sense to ask if we know our own thoughts and beliefs? In fact, our question is even more basic: if we consider the scenario of



ordinary life, does it make sense to ask about self-knowledge? Let us recall that the transparency condition deals with a specific situation, when the person in the present tense, inquires him/herself about his/her own belief: "Do I believe that p?" Well, if it is related to ordinary life we should still ask another question: why would a person inquire about his/her own beliefs? In these situations, the end result is almost never an explanation or knowledge, but a decision the person will take, affecting the conduct of his/her life, often changing his/her thoughts. This distinction between a purely theoretical attitude and a practical or deliberative attitude introduces a special dimension to the issue of substantiality of self-knowledge: often we reflect on our thoughts and beliefs because we need to make decisions and act. This is the point that concerns the transparency condition.

What is relevant here is not to offer a good argument or explanation to clarify how a person attributes to him/herself an attitude that would put him/her in a particularly safe situation, but to understand how the expression of a belief is the expression of an attitude that matters to the person. When we say "I believe that p", we are expressing our belief in p; differently, when we say "Joseph believes that p", we are giving Joseph a certain mental state without committing ourselves to the truth of p. Therefore, in certain paradigmatic circumstances, when using the words "knowing" and "believing" it makes sense to doubt Joseph's current belief, but it does not make sense to doubt our current belief. This commitment to the belief accompanies first-person reports. From this angle, self-knowledge would involve first the ability we have to express or manifest our mental states and not just the capacity that we have to assign these states to ourselves, just as we do when we attributing mental states to others. Second, our ability to meet our current beliefs can, at least in certain situations, be understood as reflecting an ability to act on ourselves. In this second sense, self-knowledge is directly related to the universe of moral problems.

Belief and Answerability

To we say we have a belief - that we claim to believe – is not to assign ourselves a mental state, but to express a belief about something. Wittgenstein had already explored the strangeness of this (philosophical) position of the first person in relation to his/her own beliefs. In the second part of *Investigations*, we read "This is what I think: believe is a mental state" (Wittgenstein, 1953, p. 191). The strangeness is due to the fact that if we accept this, we must devise a symmetry of the occurrence of the mental state in cases of first and third person. So, in order to have access to our mental state,

it would be necessary to pay attention to ourselves as others do, hear us speaking and to draw conclusions of what we say (Wittgenstein, 1953). This image of the person in face of his/her beliefs is somehow justified by a certain skepticism about self-knowledge. The same way we often ascribe mental states to others and we are usually mistaken regarding their actual states, it can happen to us when we assign mental states to ourselves. Strictly speaking, this is not so serious since it is actually possible for a person to state (honestly) that he/she has a certain mental state, and still be wrong about that. The romantic comedies are rich in examples of people who say they hate someone else, express it in words and actions, but deep down they are in love (but only the loved person does not know it). The problem is that when we mistakenly attribute a mental state to another person and we are mistaken in our assignment, it does not affect our rationality or the other person's rationality. But when we deceive ourselves about our own mental state, our rationality is threatened. Moreover, when we say we have a belief, this statement necessarily adheres to the truth of the proposition and has a role in all our attitudes and actions (Moran, 2001, p. 83-94).

The Paradox Moore (REF) orbits around this. It is not based on linguistic and logical impossibility, since the two phrases ("It's Raining" and "I believe it is not raining") can both be true simultaneously. But here is the knot of the very idea of this belief, since, despite the fact that both sentences may be true, the act of saying "It's raining and I do not believe it's raining and this (the fact that it is raining and I do not believe it's raining) is something in which I believe" does not make sense (Shoemaker, 1996, p. 76).

There is a normative commitment in this belief. In most common uses, to speak of believing simply means believing that p , and at the same time to consider that the "belief that p " is something reasonable and of course believable. In certain contexts, one can even say that the *belief that p* is justified or defensible. For Moran (2012, p. 214), the fact that a belief is reasonable or intelligible to a person is something relevant to establish why that person has that attitude, "as it became part of his/her mental life." Likewise, the subject of belief can also put him/herself the question: "Why believe that?" (Moran, 2012, p. 218-210). The fact that we do not separate the fact of having a belief and the fact of having a reason to believe it enables this form of transparency, that is, a person can normally tell us what he/she thinks about something reflecting on that belief. Wittgenstein writes: "And then it would also be possible for someone to say, 'It's raining and I do not believe it' or 'it seems to me that my ego believes this, but that's not true.' It would be necessary to form an image of a behavior in which two



people were speaking through my mouth "(Wittgenstein, 1953, p. 192). We cannot put the question to ourselves whether we believe p without thereby raising the question about the truth of p , after all there is only one person speaking when talking about p .

What would be really distinctive is that in relation to our own beliefs we have practical authority because our belief that there is water in the cup causes us to hold it, bring it to the mouth and quench our thirst, even if we do not have the epistemic authority to discriminate the water. This authority involves commitment to the truth of the belief and responsibility. As Stuart Hampshire wrote, a rational subject is an "author responsible for his/her beliefs" (Hampshire, 1965, p. 80). When someone says "I believe *that X*", "I think *that X*", "I wish *that X*", this belief, thought, desire belongs to that person; that person is agent, is author of the propositional attitude and he/she has a special responsibility that no one else has. This belief, thought, desire is not a "mere succession of representations (which, for some reason, the person is the only witness)" (Moran, 2001, p. 32).

Part of that aspect of being a rational agent is to be able to submit one's own attitudes to a review in such a way that this review makes a difference to what the attitude is. Someone is an agent with respect to one's own attitudes to the extent that he/she reflects on what is true, what he/she desires, pretends, intends, and believes (Moran, 2001, p. 64). There is an ethical aspect inseparable from the agent. Self-knowledge in this sense is not a matter of having access to information about ourselves (information that, for some reason, we could not have.) Moran (2001, p. 89) says: "when I express a belief I am not dealing with an empirical psychological fact about myself." In ordinary life, when an individual deliberately reflects about his/her beliefs and thoughts, he/she is not doing an investigation or enlisting evidence and proof. Meanwhile, even without epistemic guarantees, when an individual takes into account his/her own beliefs, he/she can (and often does) change his/her mind (his/her thoughts, beliefs, and of course actions).

Self and Reflection

All this is based on an expectation of rationality attributed to people and on the assumption that we are able to reflect on our own attitudes, meanwhile, this is a modest expectation. But, in fact, what the notion of transparency sets is that, in the context of ordinary practical life, even if our beliefs are usually not the object of reflection, or even if we are never instigated to discriminate its reasons, the fact is that

it is possible for a person to reach a conclusion, determine his/her beliefs on something simply by relying on his/her access to the reasons supporting this conclusion. The assumption of transparency does not claim that all beliefs are formed through deliberation or reasoning or that we need to apply or use a method to make our beliefs clearer. The condition of transparency simply assumes that this is an ability that people have when their actions are governed by reasons.

Beliefs and other attitudes are transparent in this sense: one does not need to practice something like a method of transparency for his/her attitudes to be transparent – to reflect on one's own reasons (reflexivity) is a person's ability, it is not a prerequisite or requirement. Here, the assumption of rationality is purely normative, it is a minimal, modest, basic assumption (Moran, 2003, p. 413), that is to say people can ask and answer questions like "What do I think about X?". People, as rational agents, have the ability to ask questions about their own mental states. The common man reflects and can ask about his own reasons.

As much as this assumption of self-knowledge is essential to understanding people, it refers to a human capacity that is partial, imperfect and fragile (Moran, 2004, p. 456). We must not forget that self-awareness is one of the means we have in order to know what happens to us, whether we have such a belief or not. But neither this is the only mean, nor it is always a safe. The testimony of others and countless other ways can be even more reliable than consciousness itself. Our daily life shows us that many times other people are able to better describe our interior, our moods than ourselves. Such possibility "has nothing to do with the unquestionable fact that only I feel [think, believe in] what I feel [think, believe in]." (Marques, 2007, p. 23). Therefore, in relation to our own beliefs, we do not maintain a relationship of knowledge or ignorance: "I might know what the other thinks, and not what I think. It is correct to say: 'I know what you think', and incorrect to say, 'I know what I think' "(Wittgenstein, 1953, p. 222). The fact that we have a belief does not impose light to the phenomenon.

Reflexivity and Agency in the Frail Transparency of What is Mental

In spite of the semiotic nature of subjective experience and likewise of the incompleteness of our (self-) understanding, the weaknesses of our reason, the events of our lives, we reflect and act in the light of our beliefs and reflections. Even if the person can assume multiple positions in the same flow of communication, reflection is present at moments in which that person faces situations wherein inquiring him/herself



about his/her belief becomes unavoidable for the conduct of the person's life. To reflect is nothing but to expose one's own attitudes under a certain angle, considering one's own contingencies, changes, precariousness and uncertainty. From a semiotic point of view, to reflect is to interpret one's own flow of signs (but not in the idealistic sense, since it also involves emotions and feelings) in order to act. Indeed, this is the sense of the maximum of pragmatism, since for Peirce, "[a]ccording to the maxim of pragmatism, then, we must ask what practical difference it can make Whether the value is p or something else" (CP 5.20).

In this sense, this is a partial capacity since transparency does not mean a complete understanding of the contents of thought, this is a fragile transparency because the psychopathology of everyday life is full of trivial examples in which transparency fails and we do not know of our own reasons.

It is perfectly reasonable to assume that a person can ignore or mistake aspects of their subjective mental states. It is also conceivable that he/she decides to do something against their better reason (eg: knowing that smoking is bad for health, yet continue the habit) or to construct a false self-image when, for example, a friend, relative or therapist might be in a better position to assess whether he/she is faithful, reliable, loving, etc. The transparency is still imperfect, since there are no epistemic and metaphysical guarantees.

Because of the multifaceted and ambiguous nature of language in our day-to-day common experience, one can perfectly master a concept sufficiently without a thorough understanding of the specific concepts and words. In a perfectly intuitive sense, we have thoughts and beliefs about the life of early Portuguese settlers in Brazil or about the lives of women in Afghanistan or the Taliban, or on stem cell research or global warming. If we are asked about the precise definition of these terms and concepts used, that originate from our thoughts and beliefs, we may have some difficulty in recognizing that we will use some words imprecisely. We will realize that we do not have a clear theory and that a historian, an anthropologist, or a biologist could do it better. In any case, we already recognize that our musings are crossed by an undeniably incomplete understanding, We can only guarantee that we were thinking and believing these things because often the source of our justification and our epistemic duty is not that we know and understand everything thoroughly but that we can explain its nature and all its truth conditions. We can think and believe things.

but we lack the authority to explain the meaning of words and other terms on which we express ourselves.

And finally, the fact that someone examines one's own beliefs, investigates one's reasons, does not mean that, in the end, his/her effort is successful, that everything becomes clear and distinct, that the reasons become clear and the light becomes present. Things may ultimately remain obscure and, likewise, rational people can remain confused and make misleading choices, even because, "it is there – like our life" (Wittgenstein, 1969, § 559). But there is a "self" (and self-knowledge) when a person is the author of his actions in a way that is irremediably personal.

References

- Bakhtin, M. (1929). *Problems of Dostoevsky's Poetics*. Edited and translated by Caryl Emerson. Minneapolis, London : University of Minneapolis Press, 1984.
- Dennett, D. (1982). Beyond Belief. In: Woodfield, A. (ed.). *Thought and Object: Essays on Intentionality* (pp. 1-95). Oxford : Clarendon Press.
- Evans, G. (1982). *Varieties of Reference*. Edited by John McDowell. Oxford, New York : Clarendon Press, Oxford University Press.
- Gallie, W. B. (1970) *Introduzione a Charles S. Peirce e il pragmatismo*. Translated by Roberto Tettucci. Firenze: C/E Giunti.
- Gergen, K. (1991). *The Saturated Self: Dilemmas of Identity in Contemporary Life*. New York : Basic Books.
- Gertler, B. (2011). *Self-Knowledge*. London : Routledge.
- Hampshire, S. (1965). *Freedom of the Individual*. New York : Harper & Row.
- Hermans, H. J. M. (2001). The Dialogical Self: Toward a Theory of Personal and Cultural Positioning. *Culture & Psychology*, Vol. 7(3), 243-281.
- Hermans, H. J. M. (2004). The Dialogical Self Between Exchange and Power. In: Hermans, H. J. M. & Dimaggio, G. (eds.). *The Dialogical Self in Psychotherapy* (pp. 13-28). Hove and New York: Brunner-Routledge.
- Marques, A. (2007). Vivência e Significado. In: Wittgenstein, Ludwig. *Últimos Escritos sobre a Filosofia da Psicologia* (pp. 5-26). Lisboa : Calouste Gulbenkian.
- McGinn, C. (1982). The Structure of Content. In: In: WOODFIELD, Andrew (ed.). *Thought and Object: Essays on Intentionality* (pp. 207-258). Oxford : Clarendon Press.
- Moran, R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*.



- Princeton, Oxford : Princeton University Press.
- Moran, R. (2003). Responses to O'Brien and Shoemaker. *European Journal of Philosophy*, 11:3, pp. 402-419.
- Moran, R. (2004). Replies to Heal, Reginster, Wilson, and Lear. *Philosophy and Phenomenological Research*, vl. LXIX, n. 2, pp. 455-472.
- Moran, R. (2012). Self-knowledge, 'Transparency', and the Forms of Activity. In: Smithies, D. & Stoljar, D. (eds.) *Introspection and Consciousness* (pp. 211-236). Oxford : Oxford University Press.
- Peirce, C. S. (1868). Some Consequences of Four Incapacities. In: *The Essential Peirce: Selected Philosophical Writings*. Volume 1 (1867-93). Bloomington/Indianapolis: Indiana University Press, 1998.
- Peirce, C. S. (1931-35). *Collected Papers of Charles S. Peirce*. Vols. 1-6. Ed. Hartshorne and P. Weiss. Cambridge: Harvard University Press.
- Peirce, C. S. (1958). *Collected Papers of Charles S. Peirce*. Vols. 7-8. Ed. A. W. Burks. Cambridge: Harvard University Press.
- Richard, M. (1990). *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge : Cambridge University Press.
- Russell, B. (1910-11). Knowledge by Acquaintance and Knowledge by Description. In: Salmon, N. e Soames, S. (eds.). *Propositions and Attitudes*. (pp.16-32) Oxford : Oxford University Press, 1988.
- Salgado, J. & Gonçalves, M. (2006). The dialogical self: Social, personal, and (un)conscious. In: Rosa, A. & Valsiner, J. (eds.), *The Cambridge Handbook of sociocultural Psychology* (pp. 608-621). Cambridge: Cambridge University Press.
- Salgado, J. & Hermans, H. J. M. (2005). The Return of Subjectivity: From a Multiplicity of Selves to the Dialogical Self. *E-Journal of Applied Psychology: Clinical Section*, N. 1 (1), 3-13.
- Salgado, J. (2006). Thought as Dialogue: a comment on Marková. *International Journal for Dialogical Science*. Vol. 1, No. 1, 149-154.
- Salgado, J., & Ferreira, T. (2005). Dialogical relationships as triads: Implications for the dialogical self theory. In: Oles, P. K. & Hermans, H. J. M. (eds.), *The Dialogical Self: Theory and Research* (pp. 78-99). Lublin, Poland: Wydawnictwo KUL.
- Shoemaker, S. (1996). Moore's Paradox and Self-knowledge. In: *The First-person Perspective and Other Essays* (pp. 74-93). Cambridge : Cambridge University Press, 1996.
- Sobel, D. & Wall, S. (eds.) (2009). *Reasons for Action*. Cambridge: Cambridge

University Press.

Stout, R. (2005). *Action*. Chesham : Acumen.

Taylor, C. (1989). *Source of the Self: The Making of the Modern Identity*. Cambridge, Mass. : Harvard University Press.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford : Oxford University Press.

Wittgenstein, L. (1953). Part II. In: *Philosophical Investigations* (pp. 173-232). Oxford, Cambridge : Basil Blackwell, 2000.

Wittgenstein, L. (1969). *On Certainty*. Edited by G. E. M. Anscombe and G. H. von Wroght. New York : Harper Torchbooks.

ⁱ To refer to the texts of Peirce, we will use the abbreviation CP in reference to the Collected Papers of Charles S. Peirce (Peirce, 1931-35, 1958) followed by the volume number and paragraph quoted.