

Being Self-Involved Without Thinking About It

Confusions, Virtues and Challenges for Higher-order Theories

Miguel Ángel Sebastián

msebastian@gmail.com

1 Introduction

Theories of consciousness aim to understand the nature of our conscious experiences. One of the main challenges that they face is to explain what it takes for a mental state to be conscious rather than unconscious. Higher-order theories attempt to answer this question. This family of theories can be characterized by their endorsement of what Rosenthal (1997) calls ‘the transitivity principle’ (Weisberg 2011b, Rosenthal and Weisberg 2008), the idea that a state’s being conscious consists in one’s being conscious of being in that state; where ‘being conscious of’ is unpacked as some sort or other of higher-order awareness.¹

TP A mental state M is conscious if and only if the subject is aware of herself as being in M.

According to higher-order theories, conscious states are the targets of a higher-order awareness or representation, where the higher-order awareness concerns both the conscious state and the subject of experience. Objections to higher-order theories typically focus on the relation between the higher-order representation and the conscious state. I will argue that higher-order theories have resources to dispel these kinds of worries. On the other hand, the required self-awareness is often ignored or set aside in discussion and theorizing of the higher-order approach. I will argue that the real challenge for higher-order theories —but also for competing theories— derives precisely from this component.

¹ Higher-Order theorists disagree with regard to the best way of characterizing such higher-order awareness. The main dispute is whether such an awareness is perception-like or thought-like. Theories that maintain that it is the latter are called Higher-Order Thought (HOT) theories (See e.g. Brown 2015, Gennaro 2012, Rosenthal 1997, 2005, Weisberg 2011b). These theories hold that when I have a phenomenally conscious experience as of red I am in a mental state with certain content, call this content ‘red*’—. For this mental state to be phenomenally conscious, there has to be, additionally, a higher-order thought targeting it, whose content is something like ‘I see red*’ or ‘I am in a state that represents red*’. On the other hand, Higher-Order Perception (HOP) or ‘inner-sense’ theories maintain that what is required is a (quasi-) perceptual state directed on to the first-order one (see e.g. Armstrong 1968, Carruthers 2000, Lycan 1996).

For this purpose this paper is organized as follows. In §2 I deal with conscious states. I focus on two challenges arising from their theory of which states are conscious (§2.1): offering a justification of the transitivity principle in the first place and accommodating cases where the higher-order awareness misrepresents the conscious state. I show that both challenges can be met if we distinguish the two senses in which the term ‘conscious state’ can be used (§2.2). In §3 I deal with self-awareness in higher-order theories. I first present the different views with regard to first-person or *de se* representation (§3.1) and then present Rosenthal’s theory of self-awareness (§3.2). Against his view, I argue (§3.3) that conscious beings lack the cognitive resources that the theory predicts they should have to entertain a conscious experience, and that Rosenthal’s views on self-representation do not make room for the intuition that different subjects can have experiences with the same phenomenal character. Finally, in §4 I sketch the view that results from considering all the challenges. The problem of explaining self-awareness remains as the main challenge but the analysis reveals the constraints and tools that are available to face it.

2 Conscious State

2.1 Two Challenges

The Transitivity Principle (TP) is key to support higher-order theories, at least according to Rosenthal. However, TP is also a source of problems.

In the first place, TP seems to lead to a regress: if a state’s being conscious requires one’s being conscious of that state and this in turn another conscious state we fall into a vicious regress.² Defenders of higher-order theories reply that the higher-order state that makes us conscious of the mental state is unconscious, thereby avoiding the regress. However, this leads us straightforwardly into the problem of offering a justification of TP.

Although many find the Transitivity Principle (TP) plausible, it requires a justification. Rosenthal (2005) thinks that it is part of our folk-psychology that a state we are not aware of in any sense cannot be regarded as a conscious. This might be right, but insufficient if the theory attempts to be a theory of our conscious experience and explain what it takes for a state to be conscious in the phenomenal sense —what Block (2011) calls ‘*ambitious* higher-order theories’. In such a case, we need a justification that links the folk-psychological notion of conscious state with that of a phenomenally conscious state —a state such that there is something it is like to have it. Kriegel (2009a) argues that the required evidence that supports TP cannot come from empirical evidence nor conceptual analysis, and that it should come from what is phenomenologically manifest; i.e., from the information the conscious state carries:

² For detailed discussion see Caston 2002, Kriegel 2009b, Williford 2006.

its content. If Kriegel is right, then the higher-order theorist seems to be in trouble. The higher-order representation carries the required information, and it represents that I am in the conscious state. However, defenders of higher-order theories hold—in reply to the apparent circularity of TP—that this is an unconscious state, and hence it cannot provide the required phenomenological justification. The conscious state, on the other hand, cannot do the job because it conveys something about the world—the presence of certain property, red*, in the example of an experience as of red—, but not the state the subject is in.

A different challenge derives from the combination of TP and the possibility of misrepresentation (Block 2011, Byrne 1997, Neander 1998). If the relation that holds between the higher-order awareness and the conscious state is that of representation then there has to be room for misrepresentation. But in case of misrepresentation, the higher-order awareness represents oneself as being in a conscious state that one is not actually in. In reply, defenders of higher-order theories explain that one would still enjoy the corresponding experience because one would still be aware of oneself as being in a conscious state. The conscious state would be a mere intentional object—as the pink elephant is when one hallucinates it (Rosenthal 2011b, Weisberg 2011a).

This analysis of cases of misrepresentation is problematic nonetheless. Consider the following uncontroversial claim:

EXP Having a conscious experience E just is being in conscious state M.

When I have a conscious experience E I am in certain state M, and by definition there is something it is like for me to be in M. Contrast this situation with one in which I were under the effect of general anesthesia. I would be in a different kind of state, M*. There is nothing it is like to be M* and hence I would not have a conscious experience. A theory of phenomenal consciousness aims to explain what distinguishes states that are like M (conscious states) from those that are like M* (unconscious states). So, I take it that any theory of phenomenal consciousness has to accept EXP. And *ambitious* higher-order theories are theories of phenomenal consciousness.

This seems to lead to a contradiction. In case of misrepresentation I am aware of myself as being in a state M, therefore M is conscious (by TP) and defenders of higher-order representation hold that in such a case I have an experience. But this is a case of misrepresentation and hence I am not in a conscious state. So, by EXP I do not have a conscious experience. As a consequence, in case of misrepresentation I have and do not have a conscious experience.³

One alternative is to deny that a conscious state can be misrepresented by holding that

³ See Farrell (2018) for a related argument.

a conscious state is one that represents itself.⁴ So, in case of misrepresentations M fails to represent itself, the subject is not in a state that represents itself and hence she is not in a conscious state. Therefore, according to the theory she has no conscious experience. The challenge for this same-order view is to explain how a state can represent itself in a way that is compatible with naturalistic theories of mental content.⁵

I think that the higher-order theorist can —and probably should— avoid the problems derived from this particular form of self-awareness where the state represents itself. In the next subsection, I will argue that the tension derived from cases of misrepresentation is only apparent and is due to a failure in distinguishing two senses in which the term ‘conscious state’ is used in TP and EXP. This allows a response to the problem of offering a justification for TP on behalf of higher-order theories.

2.2 Reply: Two senses of Conscious State

The term conscious state is used in EXP to refer to the state in virtue of which one undergoes an experience. We can use the term ‘phenomenally conscious state’ to refer to this kind of states. When one is in a phenomenally conscious state one has a conscious experience: there is something it is like to be in a phenomenally conscious state and there is nothing it is like to be in a state that is not phenomenally conscious. The state that is responsible for the experience is the higher-order one: a phenomenally conscious state is one that makes me aware that I am in such-and-such state.

On the other hand, the term ‘conscious state’ is used in TP to refer to the state that I am aware of myself as being. We can call this kind of states ‘transitively conscious states’. A state is transitively conscious if and only if I am (adequately) aware of myself as being in it. I propose, contrary to same-order representationalism, that the phenomenally conscious state and the transitively conscious state are two different states. So, when I have an experience, I am in a phenomenally conscious state, one that represents that I am in a transitively conscious state. We should rewrite EXP and TP to remark the different senses of conscious state that they deploy:

TP* A mental state M is *transitively conscious* if and only if the subject is aware of herself as being in M.

EXP* Having a conscious experience E just is being in a *phenomenally conscious* state M.

⁴ See for example Brentano (1874/1973), Burge (2007), Caston (2002), Gennaro (2012), Metzinger (2003), Kriegel (2009b), Williford (2006), Sartre (1936/1960), Williford (2015).

⁵ For an suggestive proposal see Kriegel 2009b. Cf. Sebastián 2012b

EXP* is the uncontroversial statement that having an experience just is being in a phenomenally conscious state. According to TP*, a state is transitively conscious if and only if I am aware of myself as being in it. A higher-order theorist can hold that a phenomenally conscious state is one that makes the subject aware of herself as being in a certain state (the transitively conscious one). This resolves the apparent contradiction derived from cases of misrepresentation. In such cases, I am having the experience because I am indeed in a phenomenally conscious state—a state represents that I am in a certain state, namely, the transitively conscious one. Given that this is a case of misrepresentation, I am not in the state the phenomenally conscious state represents myself as being, I am not in the transitively conscious state. But this is not a problem because this is not in tension with EXP*.

At this point, some readers might be inclined to resist the idea that the transitively conscious states and the phenomenally conscious state are different states. It is often claimed, especially within the phenomenological tradition, that phenomenally conscious states reveal themselves or make us aware of themselves. I think that phenomenology is silent on this and that there is no sound argument in favour of such a claim. Conscious experience makes us aware that we are in certain states and relations to the environment, they do inform the subject that she is in a certain state.⁶ Representational theories—whithin which higher-order theories are included—attempt to characterize what it is like to be in a phenomenally conscious state by focusing on the content of the phenomenally conscious state, on what the experience conveys to its subject. Higher-order theorists can agree that my experiences convey that I am in a certain state. But is there any good reason to think that the state I am aware of myself as being is the phenomenally conscious one?

As we previously saw, Kriegel argues in favour of same-order representationalism that unless the higher-order state is conscious, we cannot justify TP, because TP has to be justified by phenomenology. With the help of the analysis that I have offered, we can see that his argument fails. The phenomenally conscious state is unconscious only in the transitive sense, in the sense that it is not a state I am aware of myself as being—unless there is a third-order state that targets it. In the framework we are considering, phenomenology is determined by awareness in experience, that is, by the content of the phenomenally conscious state. In our case, the higher-order state, which content is ‘I am in such-and-such state’. According to the reading of the theory that I am offering, in having an experience, we are indeed aware of ourselves as being in a certain state, and this justifies TP. This is compatible with our not being aware of the phenomenally conscious state in the transitive sense.

The justification of TP aside, there are very few arguments in the literature in favor of the

⁶ Whether such a state is an internal state, as Rosenthal would be inclined to think, or a relation to the world is irrelevant for current purposes.

claim that phenomenally conscious states represent themselves. One of the few is offered by Sartre (1956, pp. 19-20).⁷ He notes that we can immediately and non-inferentially report the state we are in and what we are doing. If I ask you what you are doing, you would immediately reply—in a seemingly non-inferential way and without any need to appeal to introspection—that you are reading a paper. Sartre thinks that we have to be pre-reflectively aware of the experience to explain this fact: the experience has to inform its subject of its own presence. The argument is intended to be an abductive one (Zahavi and Kriegel 2016),⁸ where the best explanation for what one is in a position to believe immediately and non-inferentially just by virtue of having the experience is given by what the experience conveys. But in such a case there is no need to hold that phenomenally conscious states have to convey their own presence to explain one’s epistemological position when having an experience. If what is to be explained is the fact that you immediately report that you are reading the paper when you are asked in terms of what the experience conveys, then your experience has to convey that you are reading a paper. This is fully compatible with the view that I am advocating here. What the experience conveys corresponds to the content of the phenomenally conscious state. In having such an experience you are indeed aware of being in a certain state, the transitively conscious state, that in this case would have to correspond to the state of reading a paper, if Sartre is right. And this explains your epistemological situation in terms of the content of the phenomenally conscious state. There is no need to endorse the claim that the experience represents itself.

In his criticism of higher-order theories, Block (2011) claims that explaining the transition between (phenomenally) conscious and unconscious states, requires an account of the difference between having and lacking self-representational properties. He thinks that this issue is of a piece with the explanatory gap and that for this reason higher-order theories “makes consciousness out less puzzling than it really is.” I agree with Block to some extent. I agree that self-representation is key to understanding the problem of consciousness, but not because the conscious state represents itself but rather because it represents oneself.

The claim that a mental state M is self-representational is ambiguous, and theorists agreeing that a form of self-awareness is required for consciousness disagree with regard to how such self-awareness is to be understood. It can mean that M represents itself or that M represents oneself (Sebastián 2012a). According to the first reading—that we might call ‘Mental-State-

⁷ Chadha 2017 offers a similar argument from the Buddhist-Abhidharma tradition based on memory and Gurwitsch based on vehicles of representation (Gurwitsch 1985, Williford 2011). Cf. Howell and Thompson (2017).

⁸ The intended argument cannot be deductive one because, as Schear (2009) notes, we cannot logically infer from an epistemological relation (knowing that I am doing what I am doing) a phenomenological conclusion (that my conscious state is conveyed by my experience).

Involving’— the (phenomenally) conscious state also represents itself. The same-order theories considered above are an example thereof. According to Mental-State-Involving theories, the content of the state M one entertains when having, say, an experience as of red is something along the lines of *M represents red** or *This-state represents red**, where the expression ‘this-state’ corresponds to a mental demonstrative that points to the state itself (M in this case). The phenomenally conscious state and the transitively conscious state are identical. On the other hand, *Self-Involving* theories maintain that the conscious state represents that the subject herself is in a certain state (the transitively conscious state). The content of a phenomenally conscious state is something along the lines of *I am in a state that represents a red apple* or *I am seeing red*.⁹ Finally there are theorists that endorse a combination of these two, holding that the experience concerns both itself and the subject. Something like *I am in this-state*.¹⁰

Higher-order theorists can reject the view that a state represents itself (as in mental-state-involving theories). But this, *pace* Block, does not mean that there is no self-representational aspect as TP (and TP*) clearly endorses a self-involving view. However, I agree with Block that higher-order theorists make consciousness seem less puzzling than it is. They have underestimated the problems derived from reference to oneself.

3 Self-Representation

According to TP, when we have an experience we are aware of ourselves as being in a certain state (the transitively conscious one). So, having an experience requires in addition to awareness of the transitively conscious state, awareness of ourselves in some sense to be determined.

The first challenge is the apparent tension between TP and our incapacity to find anything like the self in introspection. Hume seems to be right when he points out that:

[W]hen I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception. (Hume 1739, p. 252)

Rosenthal (2011) acknowledges that we do not perceive ourselves in, say, consciously perceiving the world. However, he replies that this doesn’t show that we are not aware of ourselves in having an experience because perceiving is not the only way to become aware of things. We can also become aware of things by virtue of having thoughts about them. Rosenthal thinks

⁹ See e.g. Brown (2015), Sebastián (2012a, 2019). Block himself suggest in several places his preference for a self-involving view, for example when he points out that there is a *me-ishness* about conscious states.

¹⁰ E.g. Gallagher (2000), Zahavi (2005).

that if the phenomenally conscious state —the higher-order awareness— is a thought, then we can accommodate Hume’s observation. Moreover, Rosenthal (2002) notes that there are distinctive mental qualities associated with each perceptual modality. And he argues that if the higher-order awareness were a perceptual one, then we would expect some kind of additional quality associated with it. However, there is no room for such an extra quality, and there appears to be none. Rosenthal’s Higher-Order Thought (HOT) theory seems to be in the best position to explain the data. I will come back to these arguments later to argue that Rosenthal’s arguments in favour of higher-order awareness being a form of thought or belief are not compelling. Let’s accept for the moment that the higher-order representation is a thought.

According to HOT theories, I have a conscious experience when I have the right kind —occurrent, assertoric and seemingly non-inferential— of higher-order thought to the effect that I am in certain state (the transitively conscious one). In the previous section I dealt with the relation between the experience and the state it represents myself as being. Let’s focus now on the fact that the experience conveys something *about myself*.

3.1 *De Se* Representation: the Skeptic, the Globalist and the Portabilist

There are several ways in which I can think of myself. I can think of myself appealing to some description, for example when I think of the author of this paper. If I happen to be identical with a certain entity —maybe a certain body—, I can think of myself by thinking of that entity. I can also think of myself demonstratively (that person) or using my name (Sebastián). In all those cases, it seems that I can wonder whether I am really thinking of myself, as I might ignore that I am Sebastián or a certain entity, that I am the person I am demonstrating or that I am the author of this paper. On the contrary, I can think of myself in a first-personal or *de se* way, in which case there *seems* to be no room for such a wonder. This seems to be the kind of thought that we naturally express using the first-person pronoun —for example *that I am in such-and-such state*.

It is widely accepted that *de se* representation is special in some way, that it plays a fundamental role in cognition and that this requires us to change the way in which we think of propositional attitudes (Castañeda 1966, Chisholm 1981, Lewis 1979, Perry 1979). For example, Perry (1979) argues that first-person representation is fundamental for the rationalization of action. In his famous example, Perry is in a supermarket and unbeknown to him a package of sugar in his trolley is spilling sugar. He sees the trail of sugar and starts to look for the messy shopper. Only when he realizes ‘I am the messy shopper’ he stops following the trail and sorts the package in his trolley. However, several authors in recent years have cast doubt on the idea that first-person representation is especial (Burge 2019, Cappelen and Dever

2013, Magidor 2015). Skeptics about *de se* representation argue that there is no distinctive problem with first-person thoughts beyond well-known problems like the change in cognitive significance of correferential terms (Frege’s puzzle). *De se* skeptics think, for example, that Perry’s messy shopper case is analogous to one in which the Joker attacks the CEO of Wayne Enterprises only when he realizes that he is the Batman.

De se skeptics and *de se fans* (Morgan and Salje 2020) disagree on whether we should change our classical understanding propositional attitudes due to the phenomenon of first-person representation, where the classical understanding is defined by the following two features (Perry, 1979; Cf. Ninan 2016, Torre 2018, García-Carpintero 2017):

GLOBAL: thoughts have global truth values —as opposed to having truth values only relative to subjects or times. Given a complete description of the world we can assign them a truth value.

PORTABLE: thoughts are portable. Different individuals can entertain the same thought.

De se fans disagree with regard to which feature should we give up to offer an understanding of *de se* attitudes —and on whether there is any substantial difference between giving up GLOBAL or PORTABLE (Lewis 1979; cf. Feit, 2008). Perry agrees that first-person representations are particular cases of coreferential opacity. However, in line with Frege’s suggestion that “[e]veryone is presented to himself in a special and primitive way, in which he is presented to no one else” (Frege, 1984, p. 359), he abandons PORTABLE, distinguishing the content of the thought from the thought-state through which such a content is accessed. Thoughts have absolute truth value (he respects GLOBAL), but a complete rationalization of action requires not only contents, but also the state through which they are accessed. Thought-states are particulars that cannot be shared by different individuals, thereby abandoning PORTABLE. When David and Miguel have the thought they would express with the sentence ‘I am in M’, David thinks that David is in M and Miguel thinks that Miguel is in M: they have thoughts with different content. The behavioral dispositions that they share under these conditions are explained by the fact that they are in the same kind of thought-state. Moreover, when Miguel thinks that David is in M, he has a thought with the same content that David’s first-person thought. However, this is a different kind of thought because David’s first-person belief includes a form of access to such content that only he can have.

Lewis (1979), on the other hand, gives up on GLOBAL, taking the content to be true or false only relative to a subject and a time. He proposes contents to be properties (Chisholm 1981, Sosa 1981) that the subject self-attributes, selecting not classes of worlds but classes of centered worlds (pairs of worlds together with a designated subject and time). According to Lewis, David and Miguel have beliefs with the same content when they believe ‘I am in

such-and-such state' —they self-attribute the same property, that of *being in such-and-such state*, or select the same set of centered worlds—, thereby respecting PORTABLE.

I will argue that higher-order theorists should be a *de se* fans and endorse a portabilist view —they should respect PORTABLE thereby rejecting GLOBAL. But let me first discuss the details of Rosenthal's theory of self-awareness.

3.2 Self-Representation in Rosenthal's Theory

Rosenthal (2011a, p. 28) stresses that for a mental state to be transitively conscious —and to thereby have a conscious experience—, it will not suffice that one is aware of someone that happens to be oneself is in that state, one has to be aware of *oneself, as such*, as being in the state. If we express the content of our higher-order thoughts we would naturally deploy the first-person pronoun. According to Rosenthal, if I am aware that the person (Sebastián) or the entity I am identical with (X) is in such-and-such a state, that would not result in the such-and-such a state being (transitively) conscious unless I am aware that I am Sebastián or that I am X, for otherwise the thought would not make me aware in any relevant way of myself as being in such-and-such a state. Rosenthal thinks that we can overcome this problem without giving up on GLOBAL if the subject were aware that the subject who has the HOT is the subject the first-person concept deployed in the HOT refers to. However, if the HOT were to make explicit that the individual who has the HOT is in such-and-such a state, then one would be aware not only of such-and-such a state but also of the HOT. And in this case the HOT would become transitively conscious. But, as we have seen, this is not what typically happens, and the HOT is not transitively conscious unless there is a third-order thought that targets it. To deal with this problem, Rosenthal proposes that the HOT only identifies tacitly the transitively conscious state as belonging to the individual that has the HOT, without the HOT having such explicit content. The remaining challenge for the Higher-Order Thought theorists is to explain how the first-person concept refers in a way that is compatible with the theory. Rosenthal meets this challenge by endorsing what I take to be a variety of the globalist strategy.

Rosenthal (2011a) argues that this is possible if i) the HOT represents the transitively conscious state belonging to some particular individual S and ii) S is disposed to identify the individual the HOT refers to as the individual that has the HOT. This way, although the HOT does not represent the individual as the thinker of the HOT, the individual is disposed to make such an identification *if the question arises*. We can capture the content of the HOT with a proposition whose value depends only on the way the world is —in my case, that Sebastián is in such-and-such a state—, so we do not have to give up GLOBAL. But on top of the content we need to also appeal to the way in which such a content is accessed, we need to appeal

to some other distinctive feature of the state that has this content. In Rosenthal's view the thought-state is characterized by the disposition to identify the individual the HOT refers to as the individual that has the HOT. This is compatible with a skeptic and with a globalist *de se* fan view. It seems reasonable to hold that the disposition to make the identification between the subject that has the HOT and the subject the HOT refers to is one that only the subject that has the HOT can have. And in such a case, only the person who has the HOT can have the required kind of thought—a thought that is individuated by a content and a thought-state—. So, Rosenthal gives up PORTABLE while holding GLOBAL. If this is correct, he is a *de se* fan of the globalist kind.¹¹

3.3 Objections to Rosenthal's proposal

I see three problems with Rosenthal's approach. The first two are related to his particular theory of self-representation. The third one is a general objection to *de se* skeptics and globalist *de se* fans.

First, many have doubted that human babies and animals have the conceptual capacities that the Higher-Order Thought theory demands, but setting that aside, they definitely do not have the disposition to make the identification required to possess the first-person concept deployed in the HOT. Rosenthal is aware of this problem and he claims in reply that:

[Animals and human babies] have no irrelevant, inessential ways of referring to themselves in thought. They do distinguish themselves from everything else, and can thereby refer to themselves in thought. But their HOTs do not require the essential indexical, since distinguishing themselves from everything else provides the only way they have to refer to themselves (2011a, p. 34).

I think that we have good reasons to think that this is wrong. The capacity to attribute mental states to others have been demonstrated in, for example, corvids and canids (Bugnyar and Heinrich 2006, Hare and Tomasello 2005, Stulp et al. 2009, Udell et al. 2008). Some of these animals might not recognize themselves when they see their image, for example, reflected in a river, but they might recognize the presence of an animal in these circumstances—we all have seen dogs barking at mirrors and reflections in windows. So, one would expect that by

¹¹ Several authors have thought that a globalist approach in general has problems accommodating the phenomenon of Immunity to Error through Misidentification (IEM) (cf. e.g. García-Carpintero 2018) a conclusion that Rosenthal is happy to embrace. In his view, the reference of the first-person pronoun is indeed mediated by an identification, so there should be room for the possibility of error in this process. Rosenthal (2011a) argues that cases of dissociative identity disorder might illustrate such a possibility.

seeing their own image reflected in a river they can come to attribute a certain mental state to the animal they are seeing.¹² Therefore, it is false that they lack inessential ways to refer to themselves and hence the problem of fixing the reference of the *I* concept in their HOTs persists.

Second, as I have argued elsewhere (Sebastián 2018), Rosenthal's theory of self-reference is inadequate for humans either, because it is implausible that even we, human adults, have the cognitive capacities that it demands. Abstracting from the details, the reason is that if one is to identify the individual the HOT refers to as the individual that has the HOT, then it has to be possible that the HOT becomes transitively conscious: under certain circumstances, the HOT can become transitively conscious so that the subject knows what is the thought whose subject she has to identify. But according to the theory, this would require that it is possible that the subject has a transitively unconscious third-order thought which also deploys the first-person concept. For this concept to refer in the required way the subject requires the disposition to identify the individual the third-order thought refers to as the individual that has the third-order thought. But if the subject is to have such a disposition then it has to be possible that the third-order thought becomes transitively conscious.... and so on, *ad infinitum*. So, the reference fixing mechanism of the self concept deployed in the higher-order thought depends upon an arbitrary tall hierarchy of dispositions to have higher-order thoughts. And it is implausible that cognitively limited systems like us have such a capacity.

The final, and I think fatal, objection is that, if conscious experiences entail a form of self-awareness, then a globalist approach —and the *de se* skeptic *a fortiori*— cannot make room for the possibility of different individuals having experiences with the same phenomenal character. Here is the argument.

The phenomenally conscious state —the state such that there is something it is like to be in it— is a state that makes me aware that I am in a certain state (the transitively conscious state). This allows us to naturally dispel the apparent tension derived from cases of misrepresentation and to provide a phenomenological justification of TP*. Phenomenology (what it is like to be in a phenomenally conscious state) is determined by the content of the awareness that the phenomenally conscious state provides. And it conveys that I am in a certain state.

Not everyone would agree that phenomenology is exhausted by what the phenomenally conscious state represents, but everyone that accepts that experiences have content should agree with the idea that two experiences have the same phenomenal character only if there is no difference in content —only if they do not convey different things. This is a fairly weak

¹² See Perez and Gomila (forthcoming) for reasons to doubt that perception —even when one has the capacity— suffices for mental attribution.

condition, compatible even with the idea that there is more to phenomenology than what the experience represents. Crucially, it is one that the defender of representationalism in general—and Rosenthal, who thinks that consciousness is a matter of explaining how things appear to us, in particular—should endorse. So:

Let E_1 and E_2 be two numerically different experiences of two different subjects. Let PCS_1 and PCS_2 the corresponding phenomenally conscious states:

1. If E_1 and E_2 have the same phenomenal character, then PCS_1 and PCS_2 have the same content.
2. A phenomenally conscious state is one that represents (*de se*) that I am in certain state (TP* and EXP*)
3. According to the *globalist de se fan* (and to the *de se skeptic*), two different individuals do not have representations with the same content when they have a first-person thought.
4. PCS_1 and PCS_2 cannot have the same content.
5. E_1 and E_2 cannot have the same phenomenal character.

∴ Two different subjects cannot have experiences with the same phenomenal character.

Pre-theoretically, we want to leave open the possibility that David and Miguel have the same experience, for example with regard to the color—maybe under different circumstances. But those who, like Rosenthal, think that the phenomenal character is determined holistically might have independent reasons to resist this claim. In Rosenthal's theory, the phenomenal character of the experience depends on the state one is represented as being. And those states are individuated holistically, depending on the relations they bear to other (qualitative) states (Rosenthal 2010, 2015). Therefore, we have reasons to deny that, in general, any two individuals can have experiences with the same phenomenal character. However, it is possible for two different individuals to share the cognitive structure, function or quality space that holistically individuate qualitative states—in the extreme case, just think of two physical duplicates with the same history of interactions with the environment. Thus, independently of issues regarding holistic versus atomistic individuation of states, we do want to leave open the possibility that, in some cases, two different individuals can have experiences with the same phenomenal character—let then Miguel and David be two such individuals. This would require that there is no difference in the content of their experiences. However, this is not what happens if we accept GLOBAL, as globalist *de se fans* and *de se skeptics* do. According

to Rosenthal, the content of the experience that David has is something like that David sees red* or that David represents red*, whereas the content of Miguel's experience is that Miguel sees red* or that Miguel represents red*. Their respective experiences represent themselves as being respectively in the same kind of state. Moreover, Miguel and David share the thought-state determined by a shared disposition to identify the individual the PCS refers to as the individual that has the PCS —and arguably other dispositions because, *ceteris paribus*, they will have similar behavioural dispositions when they are in their respective states. We can even leave open the possibility that the thought-state contributes to phenomenology, so that there might be something more in common in David's and Miguel's experience. But, according to the globalist *de se fan*, David and Miguel cannot have experiences with the same phenomenal character, given the difference in the content of their respective states. However, this is a possibility that the portabilist *de se fan* can accommodate.

4 Meeting the Challenges: Self-Involving, Portabilist and Non-Conceptualist.

Many authors in the history of philosophy —from Aristotle to Kant and from Brentano to Sartre— have argued that self-awareness is key to the problem of consciousness. Although some have argued to the contrary, I think that the transitivity principle defended by Rosenthal is heir to this tradition.

As we have seen, self-awareness can be unpacked in terms of awareness of the experience itself (mental-state involving) or as awareness of oneself (self-involving) —leaving room for view that endorses both forms of self-awareness. In the first part of the paper (§2), I dealt with some objections to higher-order theories that concern their understanding of what states should be regarded as conscious. I argued that there is no good reason to endorse a mental-state-involving view and that they should favour a reading of TP that dissociates the phenomenally conscious state from the state we are aware of ourselves as being in having the experience —the transitively conscious one.

TP (and TP*) demands some sort of awareness of oneself in experience: a self-involving self-awareness. Although most of the debate surrounding higher-order theories focuses on the notion of conscious state, the required first-person representation is the real source of problems, as I have argued in the second part of the paper (§3). It is in virtue of underestimating the problem of the first-person perspective that I agree with Block that higher-order theories make consciousness out less puzzling than it really is. Rosenthal (2011a) takes the problem seriously.¹³ Although he flirts with skeptical views —according to which there is nothing distinctively problematic in *de se* representation—, his proposal should be understood as a

¹³ For complementary discussion regarding the *de se* phenomenon and higher-order theories see Weisberg (2019).

de se fan of the globalist kind. I have argued that his proposal is inadequate because it demands cognitive capacities that conscious beings seem to lack and because it does not allow for different people to have experiences with the same phenomenal character—at least not in a way that is compatible with TP*.

Contrary to *de se skeptics* and globalist *de se fans*, *portabilist de se fans* give up GLOBAL and endorses PORTABLE. According to the portabilist, the correctness conditions—the content—of an indexical representation depends not only on the way the world is but also on another parameter that we call ‘*the pole* of the representation’. In the case of first-person representation the pole is an individual. This opens the possibility for two individuals to have first-person representations with the same content—respecting PORTABLE—and that in turn allows to them to have experiences with the same phenomenal character. There is no difference in the content when David and Miguel represent *I am seeing red*. David and Miguel can self-attribute the same property, that of *seeing red* or select the same set of centered worlds, and hence they can have experiences with the same phenomenal character. Those who endorse a self-involving view have to be portabilist if they want to leave open this possibility.

Finally, I want to deal with the need to appeal to thought to characterize awareness in experience. Rosenthal appeals to thought in the first place because, as he notes, there is no distinctive quality associated with the higher-order awareness and he thinks that we would expect there to be one if the form of awareness were not a thought. I strongly disagree and I think that there is no reason to expect so. Independently of theory laden views on what mental qualities are, they play the role of explaining differences in the phenomenal character of experience.¹⁴ So, pre-theoretically we should postulate a distinctive mental quality whenever there is a difference in phenomenal character, and none when there is any. And therefore, we should not attribute a distinctive mental quality to higher-order awareness if higher-order awareness is necessary for there to be phenomenal character in the first place—because it would be ubiquitous in experience and hence unable to make a difference. In particular, according to the view I am defending, the phenomenal character of the experience is determined by the content of the phenomenally conscious state; that is, the higher-order representation. All phenomenally conscious states represent oneself as being in a certain state. Differences in the phenomenal character of conscious experiences are given, precisely, by the differences in the state they represents oneself as being—the transitively conscious state—and not by the fact that one is represented as being in it, given that this is something common to all phenomenally conscious states. The theory predicts that there are only as many kinds of qualities as kinds of transitively conscious states: there is no additional quality to be expected from higher-order awareness.

¹⁴ See e.g. Berger (2018), Rosenthal (2005, 2010, 2015).

Rosenthal also thinks that Hume's observation provides reasons in favour of Higher-Order *Thought* theory. But the portabilist is in a position to respond to Hume without endorsing the claim that the phenomenally conscious states have to be thoughts. As we have seen, *de se* awareness in experience cannot be a matter of representing a particular entity, the self. Otherwise the entity that Miguel represents and the one that David represents would be different, and then there would be a difference in content, which entails a difference in phenomenology. Hume's mistake, and Rosenthal's with him, derives from a failure in appreciation in the difference between the way the world and the subject —the pole of indexical representation— enters the adequacy conditions of *de se* representation.

This can be straightforwardly appreciated by considering other forms of indexical representation in experience. Visual experiences uncontroversially represent the world from a certain location. The one I have right now, represents a rubick's cube to the left of the monitor from here (*de hinc*). Hume —and I for what it is worth— is incapable of introspecting the self in his experiences. I predict that he would also be unable to introspect the *hereness* in the case of a visual experience. But we do not conclude from this failure in introspection that the visual experience does not concern a location. Neither do we conclude from this that the location is not part of the adequacy conditions of the perception or that it has to be represented in thought. So, why should one conclude that we are not self-aware or that self-awareness has the form of a thought?

It seems reasonable to conclude that we should do so only if we accept GLOBAL and we think of the location (in *de hinc* representation) or the subject (in *de se* representation) as another object of experience. But, as we have seen, there are good reasons to reject this view. If, as I blink, I were teleported to New York to a room with an identical set up I would have an experience with the very same phenomenal character, and both would be correct or adequate. Those who endorse GLOBAL —the skeptic and the globalist— can explain the adequacy by a change in the content: the first experience represent a rubick's cube to the left of the monitor from such a location in Mexico and the second from another location in New York. But as a consequence they would predict a difference in phenomenal character, against the assumption. On the contrary, the portabilist, rejects GLOBAL. For her, the adequacy of a state depends not only on the way the world is but also on a pole —a location in this case. This way she can explain that they are both correct or adequate without postulating a change in content, and thereby accommodate the absence of difference in phenomenal character. We do not conclude that the experience does not concern the pole. The lesson learned is that we fail to introspect the pole of representation, not that the experience does not concern it. We fail to introspect the location in *de hinc* representation and likewise we would expect to fail to introspect the subject in *de se* representation. There is nothing to fear from Hume's observation as it is

exactly what the portabilist would predict.

If this is correct, we are left with no reason to appeal to thoughts to explain phenomenal consciousness. We do not need to attribute conscious beings the conceptual capacities required to entertain a thought with content attributed to the phenomenally conscious state—embracing state non-conceptualism (Heck 2002). More generally, we do not need to attribute conceptual capacities to a being in order to attribute to it a conscious experience, avoiding overintellectualizing the requisites to have the experience. The real challenge that remains is to explain what it takes to be aware of oneself in the required portabilist and non-conceptual manner. A challenge that I think goes hand in hand with the problem of explaining phenomenal consciousness.

Rosenthal would disagree with many of the details of the self-involving view that I have sketched in this paper (portabilist and non-conceptualist), starting from the amount of things that I concede to his opponents. Despite the differences, the view is very much indebted to his work.¹⁵

References

- Armstrong, D.: 1968. *A Materialist Theory of the Mind*. London: Routledge.
- Berger, J. 2018. A defense of holistic representationalism. *Mind & Language*.
- Block, N. 2011. The higher order approach to consciousness is defunct. *Analysis* **71**(3), 419–431.
- Brentano, F.: 1874/1973. *Psychology from an Empirical Standpoint*. International Library of Philosophy.
- Brown, R. 2015. The HOROR theory of phenomenal consciousness. *Philosophical Studies* **172**, 1783–1794.
- Bugnyar, T. and Heinrich, B. 2006. Pilfering ravens, corvus corax, adjust their behavior to social context and identity of competitors. *Animal Cognition* **9**, 369–376.
- Burge, T.: 2007. Reflections on two kinds of consciousness. *Philosophical Essays, Vol. II: Foundations of Mind*. Oxford: Clarendon Press. pp. 392–419.

¹⁵ I would like to thank Josh Weisberg and the members of *Ego-Ipsum* for their comments on a previous version of the paper. Especially to Santiago Echeverri (also for suggesting to me the title), Ángeles Eraña, Carlos Montemayor, Francisco Pereira, Diana Pérez and Pablo Quintanilla. I am grateful to Dan López de Sa for discussion on the *de se* phenomenon and to Ken Williford on its relation to consciousness. Financial support for this research was provided by DGAPA projects IA400218 and IG400219, and the Spanish Ministry of Science, Innovation and Universities via research Grant PGC2018-095909-B-100.

- Burge, T.: 2019. Psychological content and egocentric indexes.. in A. Pautz and D. Stoljar (eds), *Blockheads!: Essays on Ned Block's philosophy of mind and consciousness*. MIT Press.
- Byrne, A. 1997. Some like it hot: consciousness and higher-order thoughts. *Philosophical Studies* **86**, 103–129.
- Cappelen, H. and Dever, J.: 2013. *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.
- Carruthers, P.: 2000. *Phenomenal Consciousness: a naturalistic theory*. Cambridge: Cambridge University Press.
- Castañeda, H.-N. 1966. 'he': A study in the logic of self-consciousness. *Ratio* **8**, 130–157.
- Caston, V. 2002. Aristotle on consciousness. *Mind* **111**(444), 751–815.
- Chadha, M. 2017. Inner awareness is essential to consciousness: A buddhist-abhidharma perspective. *Review of Philosophy and Psychology* **8**, 83–101.
- Chisholm, R. M.: 1981. *The First Person: An Essay on Reference and Intentionality*. Minneapolis: University of Minnesota Press.
- Farrell, J. 2018. Higher-order theories of consciousness and what-it-is-like-ness. *Philosophical Studies* **175**, 2743–2761.
- Feit, N.: 2008. *Belief about the Self: A defense of the Property Theory of Content*. Oxford University Press.
- Frege, G.: 1984. *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell.
- Gallagher, S. 2000. Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences* **4**(1), 14–21.
- García-Carpintero, M. 2017. The philosophical significance of the de se. *Inquiry*.
- García-Carpintero, M. 2018. De se thoughts and immunity to error through misidentification. *Synthese* **195**, 3311–3333.
- Gennaro, R.: 2012. *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. MIT Press.
- Gurwitsch, A.: 1985. *Marginal Consciousness*. Athens, Ohio: Ohio University Press.

- Hare, B. and Tomasello, M. 2005. Human-like social skills in dogs?. *Trends in Cognitive Sciences* **9**, 438–444.
- Heck, R. G. 2002. Nonconceptual content and the space of reasons.. *Philosophical Review* **109**, 483–523.
- Howell, R. J. and Thompson, B. 2017. Phenomenally mine: In search of the subjective character of consciousness. *Review of Philosophy and Psychology*.
- Hume, D.: 1739. *A Treatise of Human Nature*. Oxford University Press, USA.
- Kriegel, U. 2009a. Self-representationalism and phenomenology. *Philosophical Studies* **143**, 357–381.
- Kriegel, U.: 2009b. *Subjective Consciousness: A Self-Representational Theory*. Oxford University Press, USA.
- Lewis, D. 1979. Attitudes de dicto and de se. *Philosophical Review* **88**(4), 513–543.
- Lycan, W. G.: 1996. *Consciousness and Experience*. The MIT Press.
- Magidor, O. 2015. The myth of the de se. *Philosophical Perspectives* **29**(1), 249–283.
- Metzinger, T.: 2003. *Being No One: The Self-Model Theory of Subjectivity*. illustrated edition edn. The MIT Press.
- Morgan, D. and Salje, L. 2020. First-person thought. *Analysis* **80**(1), 148–163.
- Neander, K. 1998. The division of phenomenal labor: A problem for representationalist theories of consciousness. *Philosophical Perspectives* **12**(S12), 411–34.
- Ninan, D.: 2016. What is the problem of de se attitudes?. in M. García-Carpintero and S. Torre (eds), *About Oneself*. Oxford: Oxford University Press.
- Perez, D. and Gomila, A.: forthcoming. *Social Cognition and the Second Person in Human Interaction*. Routledge.
- Perry, J. 1979. The problem of the essential indexical. *Noûs* **13**, 3–21.
- Rosenthal, D.: 1997. A theory of consciousness. in N. Block, O. J. Flanagan and G. Guzeldere (eds), *The Nature of Consciousness*. Mit Press.
- Rosenthal, D. 2002. How many kinds of consciousness?. *Consciousness and Cognition* **11**, 654–657.

- Rosenthal, D.: 2005. *Consciousness and mind*. Oxford University Press.
- Rosenthal, D. 2010. How to think about mental qualities. *Philosophical Issues*.
- Rosenthal, D.: 2011a. Awareness and identification of self. in J. Liu and J. Perry (eds), *Consciousness and the Self: New Essays*. Cambridge: Cambridge University Press.
- Rosenthal, D. 2011b. Exaggerated reports: reply to block. *Analysis* **71**(3), 431–437.
- Rosenthal, D.: 2015. Qualities spaces and sensory modalities.. in P. Coates and S. Coleman (eds), *Phenomenal Qualities: Sense, Perception and Consciousness*. Oxford: Oxford University Press.
- Rosenthal, D. and Weisberg, J. 2008. Higher-order theories of consciousness. *Scholarpedia* **3**(5), 4407.
- Sartre, J.-P.: 1936/1960. *The Transcendence of the Ego: An Existentialist theory of Consciousness*. New York: Hill and Wang.
- Sartre, J.-P.: 1956. *Being and Nothingness*. trans. Hazel E. Barnes New York: Philosophical Library.
- Schear, J. K. 2009. Experience and self-consciousness. *Philosophical Studies*.
- Sebastián, M. Á. 2012a. Experiential awareness: Do you prefer "it" to "me"? *Philosophical Topics* **40**(2), 155–177.
- Sebastián, M. Á. 2012b. Review of u. kriegel "subjective consciousness: a self-representational theory". *Disputatio* **6**, 413–417.
- Sebastián, M. Á. 2018. Drop it like it's hot: a vicious regress for higher-order thought theories. *Philosophical Studies*.
- Sebastián, M. Á.: 2019. Subjective character, the ego and de se representation: Phenomenological, metaphysical and representational considerations on pre-reflective self-awareness. in M. F. Marc Borner and K. Williford (eds), *Senses of Self: approaches to pre-reflective self-awareness*. ProtoSociology.
- Sosa, E.: 1981. Propositions and indexical attitudes. in H. Parret (ed.), *On Believing*. Berlin: Walter de Gruyter. pp. 316–332.
- Stulp, G., Emery, N., Verhulst, S. and Clayton, N. 2009. Western scrub-jays conceal auditory information when competitors can hear but cannot see.. *Biology Letters* **5**, 583–585.

- Torre, S. 2018. In defense of de se content. *Philosophy and Phenomenological Research*.
- Udell, M., Dorey, N. and Wynne, C. 2008. Wolves outperform dogs in following human social cues. *Animal Behavior* **76**, 1767–1773.
- Weisberg, J. 2011a. Abusing the notion of what-it-is-like-ness: A response to block. *Analysis* **71**, 438–443.
- Weisberg, J. 2011b. Misrepresenting consciousness. *Philosophical studies* **154**(3), 409–433.
- Weisberg, J.: 2019. Higher-order theories of consciousness and the heidelberg problem. in M. F. Marc Borner and K. Williford (eds), *Senses of Self: Approaches to Pre-Reflective Consciousness*. Vol. 36. ProtoSociology. pp. 340–357.
- Williford, K.: 2006. The self-representational structure of consciousness. in U. Kriegel and K. Williford (eds), *Self-Representational Approaches to Consciousness*. The MIT Press.
- Williford, K.: 2011. Pre-reflective self-consciousness and the autobiographical ego. in J. Webber (ed.), *Reading Sartre: On phenomenology and existentialism*. New york: Routledge.
- Williford, K.: 2015. Representationalisms, subjective character and self-acquaintance. in T. Metzinger and J. Windt (eds), *Open MIND*.. Frankfurt: Mind Group.
- Zahavi, D.: 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. MIT Press.
- Zahavi, D. and Kriegel, U.: 2016. For-me-ness: what it is and what it is not. in A. E. . W. H. D. Dahlstrom (ed.), *Philosophy of mind and phenomenology: Conceptual and Empirical Approaches*. London: Routledge.