



Research Repository

The Phenomenal Concept of Self and First-Person Epistemology

Accepted for publication in Manuel García-Carpintero, Marie Guillot (eds.) 2023. Self-Experience: Essays on Inner Awareness. Oxford University Press. Oxford.

Research Repository link: <https://repository.essex.ac.uk/35388/>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the published version if you wish to cite this paper.

<http://doi.org/10.1093/oso/9780198805397.003.0010>

The phenomenal concept of self and first-person epistemology

Marie Guillot

Abstract

I defend the view that the concept of self is a kind of phenomenal concept, anchored in self-experience. This approach sheds new light on the epistemic privileges traditionally attached to I-thoughts. Coliva (2003) argues that we should distinguish between immunity to error through misidentification and two other properties it is often confused with: the impossibility of a split between semantic and speaker's reference, and the "real guarantee". The phenomenal-concept model explains why each of these three properties arises in connection with I-thoughts, and with what frequency. It also makes sense of the fact that they are regularly conflated, given their common source in the nature of the concept of self.

Keywords

The concept of self; phenomenal concepts; self-experience; first-person epistemology; immunity to error through misidentification; the impossibility of a split between semantic and speaker's reference; the "real guarantee"

0 Introduction

In experiencing the world, we also often experience ourselves doing so. When the cat brushes my leg, part of what is phenomenally given, beside the warm, silky sensation, is that it's *me* getting the privilege of it. Let us call "self-experience" this phenomenal awareness of oneself, *as oneself*. What follows assumes that at least *some* experiences of the world come with self-experience. For reasons of space, I won't defend this assumption at length (although I will provide in Section 2.3 a brief argument supporting its plausibility)¹; the claims below are to be understood as conditionals based on this hypothesis.

The aim of this essay is to explore the significance of self-experience for the concept of self and the epistemology of the thoughts that include it, or *I-thoughts*. By "the concept of self", I mean the concept that each subject uses to think about herself *as herself* (Castañeda 1966), as she may manifest by using first-person expressions.

I-thoughts are widely believed to be protected against error in a special way, but there is disagreement on which way that is. The most discussed feature is "immunity to

¹ For a defence, see Zahavi and Kriegel (2015), Guillot (2017), and the chapters in the second part of this volume.

error through misidentification” or IEM (Shoemaker 1968). However, in her important 2003 article on the issue, Annalisa Coliva shows that IEM is regularly conflated with two other properties, which she calls respectively the impossibility of “a split between speaker’s reference and semantic reference” (hereafter “No-Split”), and the “Real Guarantee”. Coliva argues that IEM and No-Split characterise only *some* I-thoughts. What does attach to *all* I-thought, and helps define it, is the “Real Guarantee”.

Coliva’s compelling essay leaves open some further questions, which this chapter tries to answer²: (i) what is the connection between the three epistemic properties? (ii) Why are they so often conflated? (iii) What explains each of them in the first place?

Section 1 introduces Coliva’s three epistemic properties. Section 2 points out that they are observed not just in I-thoughts, but also in thoughts containing phenomenal concepts, like “red” or “painful”. My central claim is that this is more than an analogy. I-thoughts are epistemologically special for the same reason that phenomenally-grounded judgments are: because the self-concept, too, is a phenomenal concept, of a special kind. This hypothesis explains (Section 3) why each of Coliva’s three properties arises in connection with I-thoughts, and with the frequency that they do. It also entails that there is a common source for all three properties, accounting for the widespread conflation. Finally, it provides a unified account of first-person epistemology – heeding Coliva’s lesson that its three facets need disentangling, while uncovering their deeper connection.

1 Three epistemological properties of I-thoughts

Many authors believe that, when using the first-person pronoun “I” or the underlying self-concept, one possesses a sure-fire guarantee against error. However, the exact nature of that guarantee has proven elusive.

1.1 Immunity to error through misidentification (IEM)

The most discussed candidate is IEM (Shoemaker 1968). IEM protects a judgment of the form “*a* is *F*” against the following type of error through misidentification (EM):

EM [T]hat a statement “*a* is φ ” is subject to error through misidentification relative to the term “*a*” means that the following is possible: the speaker knows *some particular thing* to be φ , but makes the mistake of asserting “*a* is φ ” because, and only because, he mistakenly thinks that the thing he knows to be φ is what “*a*” refers to. (Shoemaker 1968: 557, my emphasis.)

² Coliva (2017) addresses some of these questions in a different way; but there are potential points of convergence with this chapter. As the below will make clear, I agree with her passing suggestion (2017: 241) that “there are non-conceptual yet first-personal representations, which ground our possession of the first-person concept.”

Some I-thoughts are immune to this type of mistake when formed on certain grounds. Suppose that I judge, based on inner perception, that I am hungry. Assuming that I'm right in judging (on this basis) that *someone* is hungry, I must also be right in judging (on the same basis) that *I* am hungry.

IEM_{de se} One way I can't be wrong in judging "I am F" based on grounds G is by knowing that property F is instantiated, where the bearer of F is a person *s* who is distinct from me.

But not all I-thoughts are IEM. Suppose I read in the newspaper that Marie Guillot has won the lottery, and judge, based on those grounds: "I am the lottery winner". Unfortunately, the winner is a homonym. This is a case of EM: my grounds (the newspaper) do give me knowledge that *someone* won the lottery; but I am mistaken in thinking that *I* have that property.

Here is the standard reconstruction of the general case of EM_{de se}. An erroneous judgment that "I am F" is justified³ by an inference of this form:

<i>Predicative premise</i>	<i>a</i> is F ⁴	("Marie Guillot won the lottery")
<i>Identification premise</i>	<u>I am a</u>	("I am Marie Guillot")
<i>Conclusion</i>	I am F	("I won the lottery")

An error through misidentification arises here because the identification premise is *false*: I am right that the person named "Marie Guillot" in the newspaper won the lottery (*predicative premise*), but wrong that *I* am that person (*identification premise*).

IEM is not a sure-fire guarantee characterising the self-concept as such, since some I-judgments are vulnerable to EM. According to Coliva (2003), the reason why many think otherwise is that they are confusing IEM with either of two other properties. The first one is the (alleged) impossibility of a "split between speaker's reference and semantic reference" in the case of "I", or what I will call for brevity's sake "No-Split".

1.2 The (im)possibility of a split between speaker's and semantic reference

An early appearance of No-Split, Coliva argues, can be found in the second part of G. E. M. Anscombe's famous claim that "The object an "I"-user means by it must exist so long as he is using "I", *nor can he take the wrong object to be the object he means by "I"*. (1975:145-6; my emphasis.) Leaving aside the details of Coliva's exegesis, I propose to define "No-Split" as follows:

³ In the sense that the subject could appeal to the inference in question if asked about her reasons to judge as she does. This does not entail that she has to go through the inference, consciously or unconsciously, to form the judgement.

⁴ To simplify, I am ignoring the distinction made by Pryor (1999) between this form of EM, and what he calls "which-object" EM, a variant where the predicative premise has an existential form.

No-Split Let *o* be an occurrence of a term or concept *a*. The intended reference of *o* necessarily coincides with the semantic reference of *o*.

No-Split is a guarantee against splits in reference of the kind made familiar by Donnellan (1966) and Kripke (1977) for definite descriptions in their referential uses (e.g. “Smith’s murderer”), and ordinary proper names. Here is an example of the latter kind. Suppose I only know Daniel Dennett from blurry photos. At a conference, I observe in the distance a man with a white beard and an extraordinary shirt. I wrongly take him to be Daniel Dennett (it is, in fact, Saul Kripke). I say, “Dennett is wearing an extraordinary shirt”. Now *Kripke* is arguably the person I have a primary intention to refer to here, given that the point of my remark is to single out the wearer of that special shirt. Hence, not all uses of “Dennett” are guaranteed to speaker-refer to their actual semantic referent.

According to Anscombe (as Coliva reads her), this kind of split is excluded with “I”, whether in language or in thought, because, she believes, one cannot take someone else to be oneself, as one would have to do if one were to use “I” to speaker-refer (or “thinker-refer”) to another (Coliva 2003: 426).

Coliva has two comments on this. First, it is a mistake to identify No-Split with IEM.⁵ I would put it like this: No-Split attaches to some acts of *reference* (“*a*”), while IEM attaches to some acts of *predication* (“*a* is F”). Confusing them⁶ thus obscures a requirement on the shape a representation must have if the question of IEM is to be raised meaningfully about it. In the case of I-thoughts, it is to conflate a property of *self-reference* (No-Split) with a property of *self-ascription* (IEM).

Coliva’s second point is that in any case, it is false that No-Split characterizes all uses of the word ‘I’ or the self-concept. While this may be contingently rare, the possibility exists of wrongly taking someone else to be me, and using “I” to speaker-refer or thinker-refer to that person.

To show this, Coliva uses an example from Rovane (1987: 153-4), which I will tweak a bit. Suppose Jane is visiting an art museum with her friends John and Rose. Jane, who is admiring an ancient mirror from a distance, glimpses in the mirror the silhouette of Rose, whom she takes to be herself, and an awful painting behind the silhouette. John exclaims: “There isn’t a single ugly painting in this room!”. Jane replies, based on what she sees in the mirror: “Actually, there is an awful painting behind *me*”.

Here, there is a sense in which Jane’s use of “I” speaker-refers to Rose. Suppose John points out the mistake to Jane. She might then say, “Well, *what I meant* is that there

⁵ Coliva diagnoses this confusion in Rovane (1987: 153-4) and Christofidou (1995: 227). I would add McGinn (1983: 45, 49 and footnote 11). These are merely representative examples of a widespread conflation.

⁶ It is unclear whether Anscombe herself does, since she uses her own label, *immunity to mistaken identification*; this leaves some room for questioning whether she took herself to refer to the same phenomenon as Shoemaker’s IEM.

is an awful painting behind *Rose*". This makes clear that Jane's primary referential intention in using "I" was directed towards the person in the mirror, namely Rose, whom Jane mistakenly took to be herself. The upshot is that No-Split doesn't cover all uses of "I" in language (or of the self-concept in thought).⁷

What then, is the sure-fire guarantee against error (if any) that attaches to all uses of "I", if not IEM, nor No-Split? Coliva believes there is one, which she calls the "Real Guarantee".

1.3 The Real Guarantee

To isolate the relevant property, Coliva (2003: 427 sq.) contrasts cases of splits involving "I" with cases of splits involving other singular expressions, like proper names and definite descriptions.

When a referential split occurs, what typically happens is the following: the subject *doesn't know* which person is the semantic referent of the name "*a*"; that is, she is unable to answer the question "which person is *a*?". That may be because she isn't acquainted with that person, or lacks sufficiently individuating descriptive information about them. This ignorance explains why the speaker may take another person distinct from the semantic referent of "*a*" to be the referent of that expression.

However, in the case of "I", even when there is a referential split, the presumption that *the subject knows which person she is* is never defeated. Where the referring term *a* is "I", the subject can always answer the question, "Which person is *a*?" (Coliva 2003: 428.) In other words, ignorance of the identity of the semantic reference of "I" cannot be appealed to in order to explain the split when it affects "I", as opposed to what happens with other splitting expressions. Rather, in the painting example, it is Jane's ignorance of the identity of the *other* person involved, the person visually presented to her in the mirror (Rose), that leads her to mistakenly identify that person as herself, and explains the split. That is, although one may use "I" to attempt reference to another person because one does not know who *that person* is, and one wrongly takes them to be identical to oneself, one is never under any misapprehension about *one's own identity* in doing so (Coliva 2003: 428). Another way of putting this (which ventures beyond Coliva's own analysis) is to say that in the case of "I", splits occur by virtue of *over-identification* (wrongly extending one's self-conception to include another person), rather than the simple *misidentification* involved in the Dennett example.

⁷ Intuitions may differ as to whether the "me" in "There is a tasteless painting behind me" can really be intended to refer to someone who happens to be distinct from the subject. If one agrees that it can, then the painting scenario is a counter-example to No-Split. If one disagrees, it is, instead, an example of EM with respect to the first-person. If the former (referential split), then *I believe something true of someone else*. If the latter (EM), then *I believe something false of myself*.

The sure-fire guarantee that all comprehending uses of the word “I” or the self-concept enjoy is thus neither IEM nor No-Split, but what Coliva calls “the Real Guarantee”:

The Real Guarantee (at the level of thought): The possession of the first-person concept guarantees that the subject knows which person that concept is a concept of. (Coliva 2003: 429.)

1.4 Taking stock

Coliva (2003) contributes a useful three-way distinction to first-person epistemology. The article leaves open some further questions, however. First, why is the conflation so pervasive, including in authors writing after a consensus on the nature of IEM crystallised around Shoemaker’s definition? Second, why does each of the three epistemic properties arise in connection with I-thoughts, and why are they distributed as they are, with the Real Guarantee attaching to all I-thoughts, but IEM and No-Split only attaching to a subset of them?

The rest of this chapter proposes some answers. In my view, the reason for the confusion comes from the fact that all three properties derive in various ways from a common source in the nature of the I-concept. Namely, they arise because the I-concept is a kind of *phenomenal concept*.

What I will not do here is to compare the view with other imaginable explanations of the confusion. To keep the discussion focussed, I will also leave aside the independent justifications for the phenomenal-concept view, possible objections and alternatives, which I discuss in other places (Guillot 2016 and ms). Support for the view will instead be provided, here, by its explanatory virtues: it explains each of Coliva’s three epistemic properties, their distribution across I-thoughts, and the fact that they are commonly conflated.

2 The phenomenal-concept model of the self-concept

One motivation for the phenomenal-concept view of the self-concept is that thoughts composed of phenomenal concepts and I-thoughts exhibit similar epistemic profiles.

2.1 Phenomenal concepts

Phenomenal concepts⁸ are experience-dependent concepts. By an experience, I mean, following Gertler (2012)⁹, a particular type of event: an instantiation of a particular

⁸ See e.g. Chalmers (2003), Alter and Walter eds (2007).

⁹ See also Nida-Rümelin, this volume.

phenomenal property F , for instance pain, in a subject at a time. I will also sometimes speak of an experience, however, to refer to an experience-type (a category of instantiations of the same phenomenal property F); or even as shorthand for F itself. While there is much controversy as to the exact nature and mode of reference of phenomenal concepts¹⁰, one point of agreement is that a subject cannot acquire a phenomenal concept C unless she has instantiated the relevant phenomenal property F .

A particularly clear case is that of *pure* phenomenal concepts. By this I mean those concepts we use to reflect on and classify our experiences themselves, based on what it is like to have them. Take the concept of having “pins and needles” in one’s legs. When a child who has never had this experience asks what the concept means, it is characteristically difficult to enlighten her. She must wait until she first gets pins and needles in her own limbs to acquire the relevant phenomenal concept. She might then have an *aha* moment as she reflects on the curious new experience – “So *this* is what it is to have pins and needles!” But until then, the best she can do to think of the experience is to use a descriptive concept as a proxy – something like “what people feel when they say they have pins and needles”.

Now a certain kind of empiricist might retort that *all* concepts are ultimately experience-dependent in the sense that their acquisition requires having had experiences of some sort. This is controversial; but even on such a supposition, phenomenal concepts would still stand out. On acquiring a phenomenal concept C , we acquire the capacity to apply it on the mere basis of having the experience on which its acquisition is dependent, without the mediation of any (subjectively accessible) inference¹¹. Once a subject possesses, say, the phenomenal concept of nausea, it is sufficient that she instantiates the phenomenal property to be able to apply the concept directly, based on attending to the experience. As Shea (2014: 556) puts it, “exercises of a phenomenal concept [...] show up at the personal level as *immediate* — the move from having the phenomenal experience to applying the phenomenal concept does not seem to be mediated by any intervening mental process.” (My emphasis.) This is unlike what happens with concepts such as that of a *prime number*, for example, or that of a *native of Birmingham*, concepts whose application to a given object typically require various checks or calculations, beyond the mere attention to how things look.

Finally, a *pure* phenomenal concept, in addition to being dependent on exposure to a certain phenomenal property F for its acquisition, and for the capacity to apply the concept non-inferentially, also has F (or one of its instantiations) as its referent. The subjective impression of boredom itself, or a particular episode of being bored, is what

¹⁰ See Papineau (2007) for a survey.

¹¹ This formulation is neutral as to whether phenomenal concepts must *always* be applied in this way. Perhaps we can sometimes apply correctly the phenomenal concept “grief”, say, without currently feeling grief or even simulating it. In Guillot (ms), I discuss the implications of various options on this issue in relation with the self-concept.

the pure phenomenal concept of boredom refers to.

Combining those three features, I will use the minimal definition of Shea (2014) for (what I call) pure phenomenal concepts, adapting it slightly¹²:

[Pure] Phenomenal Concept A person X 's concept C is a [pure] phenomenal concept iff_{df}

- (i) acquiring C requires X to instantiate or to have instantiated phenomenal property F ;
- (ii) on acquiring C , X has the disposition to apply C [...] non-inferentially in virtue of instantiating the phenomenal property F ;
- (iii) C refers to F [or particular instantiations of F]. (Shea 2014: 556)

Removing condition (iii) gives the definition of a phenomenal concept more generally (more about the distinction shortly):

Phenomenal Concept A person X 's concept C is a phenomenal concept iff_{df}

- (i) acquiring C requires X to instantiate or to have instantiated phenomenal property F ;
- (ii) on acquiring C , X has the disposition to apply C non-inferentially in virtue of instantiating the phenomenal property F .

I call property F in the above definitions the *phenomenal basis* of a phenomenal concept:

Phenomenal Basis The phenomenal basis of a phenomenal concept C is the phenomenal property F on which C is dependent (in the ways specified in the definitions above).

The phenomenal basis F of a given phenomenal concept C is the piece of phenomenology that contributes to individuating it: the sensation of nausea for the concept “nauseated”, for instance. Condition (i) says that this phenomenal basis is a necessary basis for acquiring the concept. Condition (ii) says that is also a sufficient basis for applying it. Condition (iii), restricted to pure phenomenal concepts, says that the phenomenal basis is also the referent of the concept.

2.2 Three epistemological properties of phenomenal thoughts

Phenomenal judgments, namely judgments composed of phenomenal concepts, display forms of resistance to error that are close counterparts to those Coliva distinguishes in the case of I-thoughts.

Let us start with the Real Guarantee. Intuitively, possessing the phenomenal concept of pins and needles guarantees that the subject knows which aspect of reality that concept is a concept of: it is *this* phenomenal property, this strange bodily sensation. This is plausibly a consequence of the fact that phenomenal concepts are based on acquaintance. The acquisition of the phenomenal concept of pins and needles, as we have

¹² Shea uses the expression “phenomenal concepts” to refer to what I call “pure phenomenal concepts” here. The difference is purely terminological. My preference for using the expression “phenomenal concepts” as I do is that conditions (i) and (ii) below capture a robust, and epistemically very significant, sense of experience-dependence.

seen, requires the subject to *have* (or to have had) the experience of pins and needles, F, to which it refers; which is just another way to say that she must have been acquainted with F.

Chalmers (2003) proposes that the phenomenal quality itself can directly enter the content of a phenomenal judgment that contains the corresponding phenomenal concept:

The clearest cases of direct phenomenal concepts arise when a subject attends to the quality of an experience, and forms a concept wholly based on the attention to the quality, “taking up” the quality into the concept. (Chalmers 2003, 235)

One possible way to understand this claim is to say that a pure phenomenal concept (what Chalmers calls a “direct phenomenal concept”) canonically refers by “sampling” the phenomenal property it is a concept of, in something like the way that a sample of a certain shade of blue in a paint manufacturer’s catalogue can refer to that very shade of blue. On this view, to refer to the sensation of pins and needles under the corresponding phenomenal concept is to refer to it by *exemplification*. In a typical use of a phenomenal concept, the referent itself is present *in* the representation, as a sample experience.¹³ This direct presence of the referent rules out any doubt as to its identity for the subject who grasps the representation.

This view of the mode of reference of phenomenal concepts is not uncontroversial. However, whether or not typical *uses* of a phenomenal concept select their referents in the way just sketched, the more basic fact that the *possession* of the concept is based on acquaintance suffices to entail the Real Guarantee. To be acquainted with something is surely a way to know it – indeed, a particularly secure way to know it (Russell 1910). This is not to say that being acquainted with something reveals everything about its nature (Duncan 2015). But it is a way for that thing to be *present* to us, which rules out ignorance as to which one of the various things in the world it is. So a subject’s possession of a phenomenal concept *C* guarantees that she knows which phenomenal property *F* it is a concept of.

Let us turn to No-Split. I’ll call “native” an application of a pure phenomenal concept that is grounded only on an occurrent experience (whether full-blown or imaginatively re-created), to refer to that very (type of) experience. At least in those cases, it is very difficult to imagine how the semantic referent could diverge from the intended referent. They are one and the same thing, presented in the same way at the same time. Suppose I am looking at colour samples and asking myself which one I prefer. When I finally decide, “This one”, there seems to be no room for the phenomenal concept of a colour sensation generated there and then to latch on to something else than the colour experience I am presently focussing on: the experience serves as its own label.

But it seems possible to apply phenomenal concepts in more complex ways,

¹³ The “quotational” model of Papineau (2007) has the same consequence.

however, which opens the door to referential splits. Suppose my friend Rose suddenly starts looking bored, eyes glazing over, and she loses track of the group conversation we're in. I judge, "her boredom is making her drift off". In fact, what she feels isn't boredom, but annoyance at something someone said. Arguably, I intended my tokening of the phenomenal concept of boredom to refer to the phenomenal quality of her current emotion, the one that caused her to drift off (as it happens, the feeling of annoyance). But the semantic referent of the concept is a different phenomenal quality, the one I instantiated when I acquired the concept "boredom".¹⁴ Perhaps what opens the door to the split, here, is that the application of the concept is partly grounded on third-person perceptual information. So, while native uses of phenomenal concepts are covered by No-Split, not all of their extended uses are.

What about IEM? When a phenomenal judgment is formed purely on the basis of an occurrent experience, and when phenomenal concepts are used in both the subject and predicate positions – as in "this headache is painful", or "this feeling of nostalgia is pleasant" – it seems impossible that one should judge correctly, on that basis, that something or other has the property in question, while being wrong about what has it. After all, the judgment is merely articulating various aspects of the total phenomenal quality of one's present experience.

If, however, the judgment doesn't apply to my own occurrent experiences, or if only one of the subject or the predicate is a phenomenal concept, then EM seems possible. Suppose I overhear small children playing in the next room, then one of them crying. Based on the sound, I judge "Olivia is upset". In fact, Anna is the one who is feeling upset, and crying. Here, my perception (of the sound of distress) grounds my knowledge that a phenomenal property (feeling upset) is instantiated by someone, but a different ground (the memory of what Olivia sounds like) makes me misjudge who that someone is. In this judgement, only the predicate (being upset) is expressed with a phenomenal concept¹⁵; the subject (Olivia) isn't.

It thus appears that the epistemological properties found in connection with I-judgments are also found in connection with phenomenal judgments, and with a similar distribution. As in I-thoughts, the phenomenal counterpart to the Real Guarantee attaches to all occurrences of a given phenomenal concept. No-Split and IEM, on the other hand, may only be associated with some of its occurrences.

2.3 Self-experience

My guiding intuition is that this is more than an analogy. First-person epistemology, I suggest, has the same source as the epistemology of phenomenal judgments; it comes

¹⁴ I thank an anonymous reviewer for suggesting third-person illustrations for referential splits and EM.

¹⁵ Note that even if the concept "being upset" isn't used in the native way here, it still counts as a pure phenomenal concept in the sense captured by conditions (i)-(iii) in Section 2.1.

from the fact that the self-concept, too, is a phenomenal concept.

Phenomenal concepts are experience-dependent concepts. But notice that there is an experience that comes with being a subject: self-experience. Kriegel (2009) and Levine (2001) observe that when we experience the world, we generally experience ourselves, too. When I stroke the cat, there is a certain *qualitative character* to my experience: a way, silky and warm, that it is like to feel the cat. But the experience also has a *subjective character*: there is a way that I am phenomenally aware of myself, as myself, in being aware of the cat.

Whether this subjective character is present in every experience is a matter of dispute (see McClelland and Billon, this volume). The notion of subjective character is also interpreted in different ways by various authors, which has introduced some confusion in the debate, as I argue in Guillot (2017)¹⁶. To clarify, I will use instead the expression “self-experience”¹⁷. I will assume that at least *some* experiences of the world come with self-experience, by which I mean a phenomenal way that I am presented with myself *as myself*¹⁸. This is not just a phenomenal awareness of what *happens to be* myself. To illustrate, consider the case of somatoparaphrenia (Vallar and Ronchi 2009, de Vignemont 2013). Sufferers have an altered awareness of one of their limbs, for instance a hand or an arm. The limb comes to be felt as alien, and no longer as “me”. For example, a patient with somatoparaphrenia quoted in Maravita (2008:102), also quoted in de Vignemont (2013), asked his doctor: “Once home could I ask my wife, from time to time, to remove this left arm and put it in the cupboard for a few hours in order to have some relief from pain?” As this report makes clear, the patient can still have sensations (such as pain) in the alien limb; so the issue is not an absence of phenomenology altogether. He has experiences of what is in fact (a part of) himself; these experiences, however, are not self-experiences. This is because the patient’s experience of his alien arm lacks a certain “way it is like” that attaches to his experience of his “good” arm.

I don’t mean this example to suggest that self-experience is limited to (normal) inner bodily awareness. Some phenomenal properties are cross-modal. It might be argued that roundness, for instance, is a phenomenal character that can be shared by visual and tactile experiences alike. Painfulness, and affective phenomenology more generally, also appear to be cross-modal; perceiving one’s body from the inside can be painful, but so can be remembering a misfortune. I will assume that self-experience is the same across all the modalities in which one can be self-aware: by feeling one’s body from the inside, but also by having intentions, perceiving objects in the outer world, or being in any other

¹⁶ See also Howell and Thompson (2016) and Salje and Geddes, this volume.

¹⁷ In Guillot (2017), I distinguished between what I called “for-me-ness”, “me-ness” and “mineness”. My use of “self-experience”, here, coincides with “me-ness”. My thanks to Lucy O’Brien for suggesting this helpful term. A case might be made that “mineness” would also be a suitable phenomenal basis for the concept of self, but see Guillot (2017), Section 4.2, for an objection. “For-me-ness” could be a phenomenal basis for a generic phenomenal concept of experience, if such exists.

¹⁸ I am neutral, here, as to whether this is a *sui generis* dimension of our phenomenology, or one that is reducible to other aspects of it.

kind of conscious mental state.

This claim of cross-modality is controversial¹⁹, and so is the more basic claim that such a thing as self-experience exists at all.²⁰ This is not the place for reviewing this debate. I will therefore put my main claim in a conditional form: *if* there is such a thing as self-experience, *then* it is available as a phenomenal basis for a distinctive phenomenal concept; and this concept is our ordinary self-concept.

Let me, however, try to give a measure of initial plausibility to the antecedent of that conditional. In Guillot (2017: 45-6), I offer an argument from justification to the effect that all normal experiences, at least, come with self-experience. Here is the argument in condensed form. Ordinary experiences, such as a tactile experience of the cat's soft fur, give me an *immediate justification* to make certain judgments (Soldati 2012)²¹. One is the judgment that the cat has soft fur. That this judgment is immediately justified by the experience means, as Soldati understands it, that the experience gives me "a kind of warrant that does not depend on, for instance, any further inferentially acquired justification." The tactile experience, in other words, is a self-standing²² reason for me to judge that the cat's fur is soft: it is not itself dependent on any further reason. If I am asked why I judge as I do, the experience is all the evidence I need (and am generally able) to cite. Now, note that the experience of the cat's soft fur also gives me immediate justification, in this sense, to form other judgments²³: among others, that an experience is occurring; that it is an experience of mine; and, most relevant for my purposes, that *I, myself*, am present. Having *any* (normal) experience is enough on its own to justify my judging that I am present. And if asked for my *reason* to judge this, it is hard to see what else I could invoke than the experience itself. So the relevant epistemic warrant backing the judgment that *I* am present is likely to be something *in* the experience: something about its phenomenal character.²⁴ This gives some plausibility to the claim that we typically have *experiential* access to ourselves; or, in my terminology here, that a normal experience includes self-experience.

But *if* this self-experience exists, then it is available as a potential phenomenal basis for forming a phenomenal concept in its own right. The obvious candidate is the self-

¹⁹ See Billon, this volume, for a defence.

²⁰ See Part I of this volume for sceptical considerations. Replies include Zahavi and Kriegel (2015), as well as Billon, this volume.

²¹ See also Pryor (2013). The notion of immediate justification has been challenged (see e.g. Comesaña's 2013 reply to Pryor). It would take me too far to engage with this debate here; the role of this paragraph is only to indicate that the case for self-experience is at least *arguable*, sufficiently so to use it as a *hypothesis*. Thanks to Manuel García-Carpintero for pressing me on this.

²² This doesn't mean that the reason in question is *infallible*. The point is that the experience is what alone justifies the judgment, to the extent, large or small, that it is indeed justified.

²³ I am not suggesting that Soldati himself would endorse the argument I build on his notion of immediate justification.

²⁴ This could be challenged. For reasons of space, I won't repeat the defence of this specific claim in Guillot (2017: 45-6). Remember that the argument in this paragraph is not meant to be decisive, as my claims based on self-experience are merely hypothetical.

concept. The next section considers how this might go.

2.4 The self-concept as a phenomenal concept

Based on the supposition that there is such a thing as self-experience, my first claim is that it is the phenomenal basis of a particular phenomenal concept: the self-concept.²⁵ I can think of the phenomenal impression of the sour by experiencing the phenomenal quality of sourness; in the same way, I can think of myself, under the self-concept, by experiencing what it is like to be me.

The suitability of the phenomenal-concept model to the self-concept might not seem immediately obvious. Perhaps this is because the discussion of phenomenal concepts has largely been dominated by examples of *pure* phenomenal concepts, such as the concepts of reddishness and pain. These concepts refer to experiences (or the phenomenal properties they manifest). But the self-concept refers to something of a different metaphysical kind: a subject, namely *me*.

However, not all phenomenal concepts (as I defined them above) refer to experiences. We can use our experiences to think about those experiences themselves, but we can also use them to think about objects we are presented with through the experiences. One example is their causal origin (Papineau 2007). This is what we do with a kind of phenomenal concept Papineau (2007) calls “perceptual concepts”, and Gertler (2012) “perceptual demonstratives”; I will call them *phenomenal-causal concepts*. Those concepts can be concepts of kinds, or kind-members, or individuals; what they have in common is that they are concepts of empirical objects classified on the basis of the phenomenal impressions they cause in us.

An example of a phenomenal-causal concept referring to a kind or kind-members is our naïve concept of a palm-tree. On seeing a palm-tree for the first time, I took in a new phenomenal “template” (Papineau’s term): the characteristic look of a palm tree, with the elegant arc of the trunk and the crenelated leaves crowning the top. This complex phenomenal profile, when I re-encounter, remember or re-imagine it, will enable me to refer to the *genus* “palm tree”, or to individual palm-trees, which are the causal origin of that experience.

An example of a phenomenal-causal concept referring to an individual is the concept I formed as I kept catching a glimpse of a particular celebrity on the front page of magazines (it turned out later she is called Kim Kardashian, although I didn’t need to know that, or to have any other biographical information, to form my concept of her). There is a distinctive way that she looks: glossy black hair, amber eyes, etc. Without

²⁵ In Guillot (2016), I explore the possibility of applying the phenomenal-concept model to the self-concept with a different phenomenal basis: namely, the generic type of cognitive phenomenology that attaches to thoughts. For reasons that it would take too much space to discuss here, however, the pathological case of thought-insertion now inclines me towards the present variant.

knowing anything specific about her beyond this – I am vaguely aware that she is famous, but wouldn't be able to say on what account – I am able to use this particular phenomenal profile, this “look of her”, to refer specifically to *her*, the individual who is the causal origin of that experience.

Both my palm-tree concept and my Kim Kardashian concept are phenomenal concepts in the sense captured by conditions (i)-(ii) in Section 2.1. Each is dependent on a particular phenomenal basis. A particular (visual) experience was necessary to acquire the concepts. Once acquired, they enable me to refer to their objects on the mere basis of having the experience again, without any perceived gap between attending to the experience and applying the concept. In addition, the reference-rule which individuates each of those concepts has to be spelled out in terms of its phenomenal basis: the referent, as Gertler puts it, is whatever (appropriately) caused the current instantiation of the relevant phenomenology (Gertler 2012:104). And that reference is “direct”, in the sense that it is causally-determined rather than descriptively-determined; relational, rather than satisfactional (Bach 1987).

Now, phenomenal-causal concepts have presumably evolved to enable us to track stable natural objects, such as natural kinds and individuals. It is reasonable to further suppose, then, that concepts of this kind have the *function* to refer to the kind of entity that regularly causes the relevant phenomenal character; i.e., to palm trees in the case of the concept of a palm-tree.

Because there happens to be a reliable connection between the presence of certain individuals, or of members of a certain natural kind, and the occurrence in perceivers of a certain kind of impression, phenomenal appearances (and hence the phenomenal-causal concepts that are grounded in those appearances) are a good tool for tracking those stable natural objects. But most of those concepts can occasionally fail to perform this function: sometimes what causes a palm-tree-like impression is not a palm-tree (but a fake tree on a film set, say). As we will see in Section 3, there may be metaphysical reasons why the same cannot be said of the self-concept, and this will have consequences for the understanding of first-person epistemology.

Before moving to this next step, however, let me specify a bit the conditional claim towards which this section builds. Supposing self-experience exists, the self-concept is a phenomenal concept having self-experience as its phenomenal basis. Self-experience is typically part of what I grasp when I grasp the self-concept. To the extent that the self-concept refers to an individual rather than to the experience itself, it has similarities with my phenomenal concept for Kim Kardashian (as well as important disanalogies, discussed in Section 3). The proposal is that once I have acquired the self-concept on the basis of attending to my self-experience, I can apply it, on the mere basis of having this experience, to refer to a specific individual; namely the individual (me) who *is the subject*

of²⁶ the self-experience used on this occasion to activate the concept.

It is important to note that this differs from a descriptive view of the self-concept. The relation “being the subject of *this*” (where “*this*” captures a self-experience) doesn’t have to be represented; it is only exploited to determine a referent. Similarly, my concept of Kim Kardashian is not descriptive. What little descriptive information I have about her is not individuating. I refer, when using it, to *whatever individual appropriately caused this* (where “this” captures an instantiation of the “look of her”), but the italicized description needn’t be part of what I grasp in so doing, any more than it generally is when I think of the person pictured in a photograph (Evans 1982).

There is a second point to stress about the formulation above. Once I have acquired the self-concept on the basis of attending to my self-experience, I can apply it to myself on the mere basis of having this experience. The claim here is merely dispositional: saying that possessing this phenomenal self-concept *enables* me to self-refer with it on the mere basis of having a self-experience doesn’t mean that I *always* have to base my self-references on occurrent self-experiences. Here I remain neutral with respect to this stronger claim, whose implications are discussed in Guillot (ms).

Let me now move to the way I think the view can explain the three properties Coliva distinguishes in her discussion of first-person epistemology, and help understand why they are so often conflated.

3 A unified account for IEM, No-Split and the Real Guarantee

3.1 Explaining the Real guarantee

The view gives a simple explanation of why the self-concept enjoys the Real Guarantee. One cannot acquire the self-concept construed as a phenomenal concept without having had self-experience. Which means that I cannot possess the self-concept without grasping *what it’s like to be me*. This amounts to saying that I cannot have the self-concept without knowing which person I am.

To explain: self-experience is a way I am phenomenally presented with myself, as myself. In order to possess the self-concept, a subject must have had self-experience. So, a subject’s possession of the self-concept guarantees that she has been phenomenally presented with herself, as herself. But being presented with something leaves no doubt as to which of the various things in the world it is.²⁷ Another way to phrase the point is that

²⁶ I won’t discuss whether being the subject of an experience is reducible to being its proximate causal origin. What matters is that the subject of experience is one of the things I am presented with by virtue of having the experience. If there is nothing more to being the subject of an experience than being a certain kind of causal origin for it, then the self-concept is a kind of phenomenal-causal concept. Otherwise, it is a different sort of phenomenal concept for an individual, which also satisfies the experience-dependence conditions (i) and (ii) in Section 2.1.

²⁷ The stronger (and more controversial) claim by Duncan (2015) and others that we are not just presented with ourselves through phenomenology, but directly *acquainted* with ourselves, would provide an

being presented with an object puts one in a position to think singular thoughts about that object; thoughts whose content will be about *it*, specifically. In being presented with something, one doesn't necessarily get to know anything substantive about its nature; but one does have a way to latch onto it and to single it out among other things. Possessing the self-concept, then, guarantees that the subject knows which of the individuals in the world it is a concept of.

3.2 Explaining No-Split, I

The phenomenal-concept view also explains why splits between semantic and intended reference are possible, but comparatively rare, in the case of I-judgments.

Note that splits are excluded when only one way of thinking of the referent is active in the act of referring, and when this way of thinking of the referent coincides with the mere grasp of the reference-rule associated with the representation: that is, the grasp of the condition an object has to satisfy to count as the referent (Peacocke 2008). In short, no-split obtains when the object to which the referring term or concept is applied is only thought about, at least tacitly, as whatever satisfies the reference-rule.

A good example is the word (or underlying concept) “this”, when applied to a perceptually salient object, like the cup in front of me: no split between semantic and intended reference seems possible in such cases, because “this” semantically refers to whatever object I am perceptually focused on, and I use “this” precisely to speaker- (or thinker-) refer to the focus of my perceptual attention²⁸.

When more than one way of thinking of the object is active, however, there is a risk of a split between semantic and intended reference. Splits happen in circumstances where a referring term or concept is being applied to an object that is simultaneously singled out in reality *via* another epistemic route than a grasp of the condition encoded in the reference-rule.

Let me illustrate this contrast with a familiar example from Donnellan (1966): the utterance “John’s murderer must be insane”. Donnellan distinguishes two cases: (a) that in which I use the description “John’s murderer” *descriptively*, while looking at the body of his victim on the crime scene; (b) that in which I use the same description *referentially*, while looking at the suspect in the courtroom.

In the first case (a), I am not thinking of “John’s murderer” in any other way than as the person who murdered John. The description “John’s murderer” determines its semantic reference satisfactorily, as whatever object satisfies a descriptive condition encoded by the referring expression; and I also intend to use the description to speaker-refer to whoever satisfies the description. There is thus only one epistemic route to the

even simpler explanation. See Section 3.3 for a possible defense of that claim.

²⁸ The Spiro Agnew painting case in Kaplan (1989) does not count as a counter-example, because it isn't a case where “this” is applied to a perceptually salient object.

referent. Hence, the description is immune to referential split.

By contrast, when I exclaim “John’s murderer must be insane” while looking at the suspect in the courtroom, the description “John’s murderer” still determines its semantic reference satisfactorily; but I intend to use that description to designate a man I am being visually presented with. As there are now *two* distinct routes to the referent, a descriptive one and a perceptual one, there is a risk of their failing to converge on the same object, hence a risk of mismatch between speaker (demonstrative) reference and semantic (descriptive) reference (as in the case where the suspect I see in the courtroom is not, in fact, the murderer).

Now let us return to the case of phenomenal concepts. Just by virtue of having the concept “pins and needles”, you already know which aspect of reality the referent is; in fact, having been acquainted with the referent is a necessary condition for having the concept at all. Possessing the concept enables you to apply it on the mere basis of experiencing the relevant quality (that particular sensation); whether or not all uses of the concept are of that kind, that is at least the *typical* way such a concept is applied. In that typical case, there is *no need* to have an independent epistemic route to the referent to be able to pick out the aspect of reality one intends to apply the concept to. The same condition that is necessary for possessing the concept of pins and needles, and grasping what counts as falling under that concept, namely *being acquainted with that sensation*, is also sufficient for picking out the aspect of reality to apply the concept to.

This contrasts with, among others, descriptive concepts and names. In those two cases, mastering the referring device does not entail that you are acquainted with the referent, or that you know what object the referent is in reality. In the typical case, when a descriptive concept is applied to an individual object encountered in reality (like the suspect in the trial for John’s murder), you have in mind both the description and a perceptual presentation, giving at best defeasible evidence that the object is what satisfies the description. Similarly with names: using a name *a* to pick out correctly the thing that is *a* in reality takes more than just knowing the reference-rule for *a*, namely that it applies to whatever was baptised with the name *a*. It also takes knowing which object that is.

Thus, when one uses phenomenal concepts, splits can happen in cases where more than one epistemic route to the referent is active; but that is not the default case because there is *no need* for it. This correctly predicts that No-Split doesn’t cover the self-concept across all its occurrences (as Coliva observes), but that splits are contingently rare. In most cases, when I deploy the self-concept, I apply it to something that is singled out in reality on the same basis of self-experience that also grounds my competence with the concept. Splits occur in connexion with the self-concept only in more complex cases where more than one epistemic route to the purported referent is active. This is precisely what happens in the painting example, where the self-concept is applied to something that is picked out in reality through a third-person visual presentation, a route that is distinct from the self-experience that determines the semantic referent of the concept and grounds

its acquisition, as well as its typical uses.

3.3 Explaining No-Split, II

There is a deeper sense in which the self-concept is resistant to referential splits, however. I argued in Section 2 that the self-concept works in something like the way a phenomenal-causal concept works. But there is a particular type of referential split that ordinary phenomenal-causal concepts are vulnerable to, and that is ruled out by the self-concept. Here is an example involving the phenomenal-causal concept of a panda. I am taking a child on a visit to the zoo. Something that looks to me like a panda is playing with a ball. In fact, it is a human being in a panda costume. I draw the child's attention and point: "Look, a panda!" Intuitively, something has gone wrong.

The semantic reference of a particular *occurrence* of the phenomenal-causal concept panda is whatever causes the panda-like impression on that occasion. But presumably, the concept as a *type* also has the *function* to track pandas, in referring to the kind of things that look panda-like. And in referring to whatever is causing a panda-like impression, one does, usually, refer to a panda. However, a token of the concept might occasionally refer to something different – a human being wearing a panda costume, in the zoo example. In that case, we might say that a referential split of a kind has occurred. The agent, in mastering the concept, grasps that tracking pandas is what it is for; so part of her intention at the zoo is to refer to a thing of the kind in question. On that occasion, however, this intended reference fails to coincide with the semantic reference of her use of the panda-concept, which is an object of a different kind.

This, however, is one way the self-concept differs from phenomenal-causal concepts. Split scenarios of the kind just described can't occur in connection with the self-concept. I speculate that this has to do with the metaphysics of the entity referred to by the self-concept. Experiencing the phenomenal profile of a panda isn't always to be presented with a panda. But having a self-experience is always to be presented with *me*. Pandas themselves are not *defined* as what causes panda-like impressions. On the other hand, I find it plausible that being phenomenally aware of myself as myself, as I am in self-experience, contributes to defining what and who *I* am. Subjects are things that are aware of themselves as themselves; and being this particular subject is being what is aware of this subject as myself.

Behind these intuitions is a metaphysical assumption I find compelling: if I have self-experience, then it is an essential property of mine, and it is key to individuating me. A prominent strand in the philosophy of personal identity, illustrated by Locke and Shoemaker among others, assumes that "a human person would go where his or her mind goes" (Johnston 2010: 47). Dainton (2005) makes the intuition a bit more precise, by noting that the sort of psychological continuity that matters is *phenomenal* continuity:

[C]ould your stream of consciousness continue on but fail to take you with it? [...] Whatever else we might be, we are not the kind of being that can cease to exist whilst our consciousness continues on.” (2005, 1-2)

Relatedly,

an account of our existence and persistence conditions that is rooted in purely phenomenal unity and continuity – the sorts of unity and continuity found in our streams of consciousness – will be more compelling, more believable, than any account that is not. (2005: 2)

If we accept the existence of self-experience, the intuition, I think, can be specified further. Could the flow of self-experience continue on but fail to take me with it? Whatever else I might be, I would add, I am not the kind of being that can cease to exist whilst my self-experience continues on.²⁹

If one is prepared to go along with this supposition, a first consequence is a point of nomenclature. If self-experience is an essential property of mine, there is a sense in which to be acquainted with it is to be acquainted with me. The self-concept thus has features in common with both pure phenomenal concepts and phenomenal-causal concepts. As the latter can, it refers to an individual, on the model of direct causal reference operative for my Kim Kardashian concept. As with the former, however, it leaves no gap between being acquainted with the concept’s phenomenal basis (self-experience) and being acquainted with its referent (myself).

A second consequence is that the self-concept is immune to the type of split scenario sketched above. There is accord between the reference-rule for the self-concept and the individuation-conditions for the entity that the concept refers to. On the semantic side, a typical token of the concept exploits an occurrent self-experience to refer to whatever is the subject of it. On the metaphysical side, I am individuated as whatever is the subject of self-experience. This predicts that the self-concept is protected from empty reference, as it is indeed. This also blocks the specific way in which phenomenal-causal concepts are vulnerable to occasional referential splits. Pandas are not defined as the sources of panda-like impressions; other things, like puppets and decoys and people wearing panda costumes, can give off panda-like impressions; and pandas can also give impressions that are not panda-like. So I may use my phenomenal-causal panda concept in an occasion in which it will refer to something that, unbeknownst to me, isn’t a panda (for instance a person wearing a panda costume), or in which there will be nothing at all to be referred to (as when a panda-like impression is just an illusion). But no similar split can arise in connection with the self-concept. Since I, the subject, am metaphysically individuated as what has self-experience, and as this is just what constitutes the concept’s semantic reference, it is not the case that I could, along similar lines, mis-refer or make a

²⁹ Dainton, who is skeptical about self-experience as I define it, would not endorse this.

failed attempt to refer when I deploy the self-concept, at least when its use is based only on its phenomenal basis. Whatever appears to me via the subjective character of my experiences is me, because that is just what a self is.

In short: the semantic reference of an occurrence of the phenomenal concept of a panda is whatever causes the panda-like impression. But the concept as a type also has the function to track pandas, in referring to what causes panda-like impressions. However, a given token of the concept can fail to perform that function, causing a form of referential split. By contrast, the semantic reference of an occurrence of the phenomenal concept of self is whatever has the self-experience on which it depends. But the concept as a type also has the function to track *me*, in referring to the subject of self-experience. And as long as the concept is applied on the sole basis of self-experience, it never fails to perform that function.³⁰

3.4 Explaining IEM

Let us turn, finally, to IEM. In Section 1.2 above, $IEM_{de\ se}$ is defined as the property that attaches to an occurrent first-person judgement of the form “I am F” based on grounds G if the following holds:

$IEM_{de\ se}$: one way I can’t be wrong is by knowing (based on G) that property F is instantiated, where the bearer of F is a person *s* that is distinct from me.

IEM and “identification-freedom”

I will assume here that something like what is known as the “Simple Account” of IEM ³¹ is correct. Some authors, following the ideas of Evans (1982), further elaborated by Wright (2012), hold that IEM judgments are so immune because of their underlying justificational architecture. The inference on which they are grounded, by contrast with the type of inference behind EM (see Section 1), doesn’t involve a separate identification premise. This makes it *a fortiori* impossible that the identification premise should be false; which leaves no room for a mistaken identification. The absence of an identification premise, in turn, reflects the fact that IEM judgments are based on a type of epistemic source that simultaneously yields information about the instantiation of a property, and about the *locus* of that instantiation. I know directly that *a* is F, without having to judge antecedently that *b* is F and that *a* is *b*. The key of this now-standard account of IEM is the *unity of the grounds* for tokening each of the two components (the subject component *a* and the predicative component F) of the judgment that *a* is F.

³⁰ Not all readers will accept the metaphysical assumption on which this account turns. But of those, some will accept the self-experience assumption, and this leaves intact the more generic account of No-Split in Section 3.2.

³¹ See e.g. Wright (2012) and García-Carpintero (2018) for detailed analyses of the account.

Experience is one epistemic source that offers the required sort of unity of grounds. It yields some experiential content (e.g. the presence of hunger or pain) together with the information that *I* am the experiencer. IEM_{de se}, for this reason, is found only (with some very limited exceptions) in connection with self-attributions of experiential properties, based on experience.

The two sides of experience as a source of identification-freedom

The picture I propose correctly predicts that when it comes to I-judgments, IEM occurs almost only³² when those judgements are grounded in experience; and it explains why the unity of grounds is satisfied in this case.

Following Levine (2001) and Kriegel (2009), I have been assuming that there are two dimensions in the phenomenal character of normal experiences: (i) a *qualitative character*, such as hunger; and (ii) a *subjective character*. Having a certain experience, for instance as of hunger, is thus in itself sufficient grounds for doing two things:

- Tokening the phenomenal concept of hunger (based on the qualitative character of the experience);
- Tokening the phenomenal concept of self (based on the subjective character of the very same experience).

So, to judge “I am hungry”, there is no need for an error-vulnerable transition via a judgment of identity based on separate grounds. Hence, there is no room for error through misidentification.

This account is in line with the “Simple Account”, according to which the source of IEM is the absence of an identification premise. But it goes one small step further, in explaining *why*, in the core class of I-judgments enjoying IEM_{de se} (i.e. I-judgments grounded in experience), no such identification premise is needed. That is because an ordinary experience, on its own, provides a sufficient basis for tokening the concepts involved in both the subject-side and the predicate-side of an I-judgment formed directly on the basis of this experience, thus also providing the required unity of epistemic grounds.³³

The account also makes the correct predictions as regards the distribution of IEM across I-thoughts. As noted in Section 1, not all I-thoughts are IEM: “I am hungry”, judged on the basis of experience, is IEM; but “I am the lottery winner”, judged on the basis of testimony, is not. This is exactly what the phenomenal-concept view would lead us to expect. On this account, only those I-judgments where the predicate, as well as the

³² Not all I-judgments enjoying IEM_{de se} are experiential self-attributions. A very small class of non-experiential I-judgments also enjoy it. The judgment “I am saying this”, for instance, is IEM, whether it is based on experience or on mere reflection on the descriptive content of the concepts involved. But this is in some ways a limiting case. My claim here is just that experience is the most important (and probably also the most interesting) source of IEM_{de se}, for reasons having to do with the phenomenal nature of the self-concept.

³³ For a much more detailed account of IEM_{de se} which also appeals to the special justification given by a phenomenal state to both the predicate and subject side of first-person experiential judgements, see García-Carpintero (2018).

subject, is a phenomenal concept, and where both concepts are activated on the sole basis of a single occurrent experience (here, hunger), are IEM.

Could this account of IEM_{de se} be generalised to other cases of IEM? I suspect that it might, at least for some of the other kinds of indexical judgments, such as temporal and simple demonstrative judgments, which may also plausibly be phenomenally-grounded. This, however, is left for further work. It might be that *some* types of judgments exhibiting IEM, on the other hand, will turn out not to fit well with the phenomenal-concept account. If Wright (2012) is right that some mathematical judgments are IEM, for instance, this is likely to be a case that is out of its reach. But we should bear in mind that the two key components of the Simple Account of IEM, namely the unity of epistemic grounds and the resulting identification-freedom, are very abstract and high-level properties. They might come to be instantiated in different ways and for different reasons in different cases, which would leave room for a degree of pluralism in accounting for IEM in different domains. My proposal here is merely that in the case of IEM_{de se}, the required unity of grounds and identification-freedom have a phenomenal origin. Thus, even if the account I offer should form only a part of a general theory of IEM, it does provide an explanation, both complete and germane to the standard account, in the limited case at hand.

4 Conclusion

To conclude, let me rehearse my five main claims, all conditional on the hypothesis that there is such a thing as self-experience as I defined it.

First, the self-concept is a phenomenal concept of a special kind, with common features both with pure phenomenal concepts and with phenomenal-causal concepts, and it is based on self-experience.

This explains all three of the epistemological properties distinguished by Coliva.

It explains the Real Guarantee, and why this guarantee applies to every tokening of the self-concept. This is because possessing the self-concept requires that you have the specific experience of what it's like to be you; you thus can't have the concept without knowing which person you are.

It explains No-Split, and correctly predicts that splits between the intended reference and the semantic reference of the self-concept are possible, but atypical. This is because intended reference with the self-concept is, by default, directly based on an occurrent self-experience, through which one is presented with oneself. But this is also how the semantic referent is fixed. This means that intended and semantic reference don't take distinct epistemic routes, which could fail to converge on the same object.

It explains IEM_{de se}, and correctly predicts that not all self-attributions, but all basic experiential self-attributions, are IEM. This is because in my normal experiences, the total phenomenal character I am aware of has two components: a qualitative character, and a subjective character, or self-experience. Each justifies the application of a dedicated phenomenal concept: respectively, the concept of a quality like hunger, and the self-

concept. The same basis – this single experience – is thus sufficient for grounding the tokening of the concepts corresponding to the predicate and subject components of a judgment of the form “I am F”. But not all judgments of the form “I am F” are based solely on an experience. So while basic experiential I-judgments are IEM, not all I-judgments are.

Finally, and as a consequence of the previous four claims: all three of the epistemological properties observed in connection with I-thoughts derive from the same source; namely, the fact that the self-concept is a phenomenal concept. This common source makes it more intelligible why the conflation of the three properties is widespread, and in particular why some writers, as Coliva is right to point out, take IEM to be a property of self-reference.

Understanding the self-concept as a phenomenal concept thus provides an explanation of all three epistemic properties, and makes the correct predictions as to their frequency: all acts of self-reference under the self-concept enjoy the real guarantee, many but not all are immune to a split between intended and semantic reference, and many but not all I-judgments are IEM. It also makes intelligible a long-standing confusion in the debate on first-person epistemology. These explanatory virtues of the phenomenal-concept model of the self-concept give some additional plausibility to the claim that there is self-experience.

Acknowledgements

Financial support was provided by the DGI, Spanish Government, research project FFI2016-80588-R, a Juan de la Cierva fellowship (Spanish Government), n°JCI-2012-13921, and the Marie Skłodowska Curie project PHENOSELF, FP7-IEF-622127. This work benefited from discussion of related ideas with many colleagues, including Uriah Kriegel, Rory Madden, Lucy O’Brien and Peter Pagin, as well as audiences in Paris and Barcelona. I am very grateful to Julien Bugnon, Santiago Echeverri, Manuel García-Carpintero and two anonymous reviewers for extremely helpful comments.

References

- Alter, T. and S. Walter (eds) (2007). *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.
- Anscombe, Gertrude E. M. (1975). “The First Person.” In *Mind and Language*, edited by S.D. Guttenplan, 45-65. Oxford: Clarendon Press.
- Bach, K. (1987). *Thought and Reference*. Oxford, England: Oxford University Press.
- Castañeda, Hector-Neri. (1966). “‘He’: A Study in the Logic of Self-Consciousness”. *Ratio (Misc.)* 8 (December):130-157.
- Chalmers, David. (2003). “The Content and Epistemology of Phenomenal Belief”. In Q. Smith, and A. Jokic, eds., *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.

- Christofidou, A. (1995) "The First Person: The Demand for Identification-Free Self-Reference", *The Journal of Philosophy*, 92, pp. 223-34.
- Coliva, A. (2017). "Stopping points: 'I', immunity and the real guarantee", *Inquiry*, 60:3, 233-252.
- Coliva, A. (2003). "The first person: error through misidentification, the split between speaker's and semantic reference and the real guarantee". In *Journal of Philosophy* vol. 100 pp. 416-31
- Comesaña, J. (2013). "There Is no Immediate Justification". In M. Steup, J. Turri and E. Sosa (eds), *Contemporary Debates in Epistemology*, 2nd Edition. Wiley-Blackwell.
- Dainton, B. (2005). "The Self and the Phenomenal". In G. Strawson (ed), *The Self? Ratio* special issue, Wiley-Blackwell.
- Duncan, M. (2015). We are acquainted with ourselves. *Philosophical Studies* 172(9), 2531–2549.
- Donnellan, K. S. (1966). Reference and Definite Descriptions. *The Philosophical Review*, 75(3), 281–304.
- Evans, G. (1982). *The Varieties of Reference*. Posthumous edition by J. McDowell. Oxford: Clarendon Press.
- García-Carpintero, M. (2018) *De se* thoughts and immunity to error through misidentification. *Synthese* 195, 3311–3333 (2018)
- Gertler, Brie. (2012). "Renewed Acquaintance." In *Introspection and Consciousness*, edited by Declan Smithies and Daniel Stoljar, 89-123. Oxford: Oxford University Press.
- Guillot, M. (ms). Which kind of phenomenal concept might the concept "self" be?"
- Guillot, M. (2017). "I Me Mine: on a Confusion Concerning the Subjective Character of Experience." *Review of Philosophy and Psychology*, 8, 23–53.
- Guillot, M. (2016) "Thinking of oneself as the thinker: the concept of self and the phenomenology of intellection", *Philosophical Explorations*, 19:2, 138-160.
- Howell, R. & Thompson, B. (2017): "Phenomenally Mine: In Search of the Subjective Character of Consciousness", *Review of Philosophy and Psychology* 8, 103-127.
- Johnston, M. (2010). *Surviving Death*. Princeton University Press.
- Kaplan, D. (1989). "Demonstratives", in *Themes from Kaplan*, edited by J. Almog, J. Perry and H. Wettstein, 481–563. Oxford: Oxford University Press.
- Kriegel, Uriah (2009). *Subjective Consciousness. A Self-Representational Theory*. Oxford: Oxford University Press.
- Kripke, S. (1977), Speaker's Reference and Semantic Reference. *Midwest Studies In Philosophy*, 2: 255-276.
- Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Maravita, A. (2008). Spatial disorders. In S.F. Cappa, J. Abutalebi, J.F. Demonet, P.C. Fletcher, P. Garrard, eds., *Cognitive Neurology: A Clinical Textbook*. New York: Oxford University Press, 89-118.
- McGinn, C. (1983). *The Subjective View: Secondary Qualities and Indexical Thoughts*. Oxford: Clarendon Press.
- Nida-Rümelin, M. (2007). "Grasping Phenomenal Properties". In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. Oxford: Oxford University Press.
- Papineau, David. (2007). "Phenomenal and Perceptual Concepts." In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. Oxford: Oxford University Press.

- Peacocke, C. (2008). *Truly understood*. Oxford: Oxford University Press.
- Pryor, J. (2013). "There Is Immediate Justification". In M. Steup, J. Turri and E. Sosa (eds), *Contemporary Debates in Epistemology*, 2nd Edition. Wiley-Blackwell.
- Pryor, J. (1999). Immunity to Error through Misidentification. *Philosophical Topics*, 26(1/2), 271–304.
- Rovane, C. (1987). "The Epistemology of First Person Reference", *The Journal of Philosophy*, 84, pp. 147-67.
- Russell, B. (1910). Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society*, 11, 108–128.
- Shea, N. (2014) Using phenomenal concepts to explain away the intuition of contingency, *Philosophical Psychology*, 27:4, 553-570.
- Shoemaker, Sydney. 1968. "Self-Reference and Self-Awareness." *Journal of Philosophy* 65 (19), Oct. 1968: 555-67.
- Soldati, G. 2012. Direct realism and immediate justification. *Proceedings of the Aristotelian Society* 112: 29–44.
- Vallar, G. & Ronchi, R. (2009). Somatoparaphrenia: A body delusion. A review of the neuropsychological literature. *Experimental Brain Research*, 192(3), 533–551.
- de Vignemont, F. (2013). The mark of bodily ownership. *Analysis*, 73(4), 643–651.
- Wright, C. (2012) Reflections on François Recanati's, 'Immunity to error through misidentification: what it is and where it comes from'. In Simon Prosser & François Recanati (eds.), *Immunity to Error Through Misidentification: New Essays*. Cambridge University Press. pp. 247-80.
- Zahavi, D., and U. Kriegel (2015). For-me-ness: What it is and what it is not. In *Philosophy of mind and phenomenology: Conceptual and empirical approaches*, ed. D.O. Dahlstrom, A. Elpidorou, and W. Hopp. London: Routledge.