

The Self Deflated¹

Takashi Yagisawa

Abstract

The first-person singular notion MYSELF is divided into two parts: ME and SELF. SELF is shown to be applicable to all sorts of things, not just human beings or persons, and analyzable in terms of the relations of anaphoric dependence and identity. The so-called problems of the self are recast as having little to do with SELF but as having everything to do with ME, and a version of adverbialism concerning so-called attitudes *de se* is offered.

0. Introduction

It is commonly thought that the self poses deep *sui generis* problems in philosophy, and many philosophers speak of the “problems of the self.”² But it is a mistake to regard those deep *sui generis* problems as dealing in the notion SELF. They instead deal in a different notion, which is almost universally conflated and confused with the notion SELF. This other notion is the first-person singular notion ME. The notion SELF is associated with interrelated philosophical issues, which are instances of wider issues in philosophy of logic, philosophy of language, and metaphysics, and are subject to discussion quite independently of discussion of issues concerning the notion ME.

I have argued for a certain analysis of ME, which secures an *a priori* and metaphysically necessary connection between ME and a representational frame of epistemic access to reality (Yagisawa, 2017). In this paper, I distinguish and separate ME from SELF. Non-philosophers typically conflate the two notions, and philosophers are hardly any better. When philosophical theorizing is done under the presumption that SELF and ME are inseparable, deep conceptual confusions ensue, argumentative clarity is lost, and serious philosophical mistakes are likely to result. Such mistakes muddle one’s understanding of the metaphysics and epistemology of the self and of the first-person singular.

The claim that SELF and ME are two separable notions consists of three sub-claims: (a) that the function of SELF is exhausted by the indication of the anaphoric co-designation relation (semantically speaking) and the expression of the numerical identity relation (metaphysically speaking), (b) that the conflation of the two notions is due to the conflation of the two words “self” and “myself,” and (c) that the function of ME does not include the indication of the anaphoric co-designation relation or the expression of the numerical identity relation. I will critically examine issues widely known among philosophers of language and mind about attitudinal attribution concerning oneself under the reflexive conception of oneself as such (so-called attitudes *de se*) and explanation of behavior in terms of such attribution.

1. Semantics

Let us start with the semantics of the word “self.” Once we become clear about it, the epistemological and metaphysical importance of the notion SELF will be evident. The most common use of the word “self” is as a pronominal suffix: “herself,” “yourself,” “itself,” “themselves,” etc. Suppose that Bianca is looking at a white box in front of her. The box is lit with red light but she does not know it. Pointing at the box, Bianca says to herself, “This is a red box.” The word “herself” in the preceding sentence co-designates with the name “Bianca” in the same sentence, designating Bianca. This illustrates a typical function of the suffix “-self.”³ Another illustration of the same function occurs when Bianca describes the box by saying, “It is darker than my hat but not darker than itself.” Here the word “itself” co-designates with the word “it” at the beginning of the same sentence, designating the white box.

In general, the expression “self” as suffixed to a pronoun signifies co-designation with the appropriate antecedently occurring or understood term (proper name, pronoun, or definite description).

Consider the following sentences:

- (1) Bianca talks to her.
- (2) The box in front of Bianca is not darker than it.
- (3) The tallest student in the class is as tall as she.

What is the relation between the *designata* of the proper name “Bianca” and the pronoun “her” in (1)? We do not know the answer to this question until sufficient information is provided concerning the relevant context of utterance of (1). The speaker may be pointing to Bianca—the same person as the *designatum* of the first word of (1)—as he utters the last word of (1), in which case the relation between the *designata* is identity. Alternatively, the speaker may be pointing to some other person than Bianca as he utters “her,” in which case the relation is (at least) non-identity. Similar observations can be made about the relation between the *designata* of “the box” and “it” in (2), and about the relation between the *designata* of “the tallest student in the class” and “she” in (3). In each of these examples, in the absence of sufficient information concerning the relevant context of utterance, we do not know whether the *designata* are one and the same.

The situation changes dramatically with the addition of the suffix “-self”:

- (4) Bianca talks to herself.
- (5) The box in front of Bianca is not darker than itself.
- (6) The tallest student in the class is as tall as herself.

In (4) the pronoun “her” is suffixed with “self,” becoming “herself,” which co-designates with the antecedently occurring proper name “Bianca.” To know this, we need no contextual information about the utterance of (4). Similarly, “itself” co-designates with the definite description “the box in front of Bianca” in (5), and “herself” co-designates with the definite description “the tallest student in the class” in (6). In all of these examples, no other interpretation of the pronoun is possible. Such is the semantic function of the suffix “-self.” By illustrating this key semantic function of “self,” which we may call *indication of anaphoric co-designation*, these examples teach us two closely related lessons.

First, it is not the case that the semantic function of the expression “self” is to indicate consciousness, animateness, or even life. This is clear from the example (5). The box in front of Bianca is neither conscious, animate, nor alive. The semantic function of “self” makes the suffixed pronoun “it” designate the box. “Self” works just as well as a suffix to the pronoun “it” as it does as a suffix to a personal pronoun like “her.”

Second, the pronominal suffix “-self” can be understood to express the relation of (numerical) identity. Of course, “self” is not a predicate, so it cannot express any relation at all in the way in which a (relational) predicate does, let alone express the particular relation of identity in the way in which the predicate “is identical with” and the symbol “=” do. Nevertheless, there clearly is a sense in which “self” may be said to express identity, namely the following:

In “... *N* ... *PNself* ...” the *designatum* of *PNself* is identical with the *designatum* of *N*.

N is a proper name or a pronoun or a definite description, *PN* is a matching pronoun, and *PNself* is the result of suffixing *PN* with “self.”

The close connection between this point and the previous point is obvious. The word “self” can function as a suffix to the non-personal pronoun “it,” and when it does, the *designatum* of the resulting pronoun “itself” is identical with the *designatum* of the appropriate antecedently occurring singular term, and the latter *designatum* may well be non-conscious, inanimate, or lifeless.⁴

The use of the word “self” standing alone is derivative of its use as a suffix. In examples like (4) – (6), the occurrence of the suffix “-self” cannot be replaced by a free-standing word “self” without loss of grammaticality.

(4*) Bianca talks to self.

(5*) The box in front of Bianca is not darker than self.

(6*) The tallest student in the class is as tall as self.

Inserting the definite determiner in front of “self” hardly helps:

(4**) Bianca talks to the self.

(5**) The box in front of Bianca is not darker than the self.

(6**) The tallest student in the class is as tall as the self.

At best, (4**) – (6**) are inarticulate substitutes for (4) – (6). The statement of the form “The self is Φ ” is meaningless unless it is understood to be short for “() herself is Φ ,” “() himself is Φ ,” or “() itself is Φ ,” where the blank is filled with a singular term whose designation is determined independently, or “You yourself are Φ ” or “I myself am Φ ,”⁵ where the designation of “you” or “I” is determined in a given context of utterance in the usual way.

2. Metaphysics

The above semantic point about the suffix “-self” can be converted to a metaphysical point about the self. The metaphysical point has two parts, corresponding to the two semantic lessons just learned.

First, in order to make sense of speaking of what is called “the self,” we need to ascertain that there is an antecedently given or understood object. In other words, we should be clear about the relativity of the self and not forget that the notion SELF can be deployed intelligibly only relative to some object or other. Unless some object is given or understood prior to its deployment, the notion is useless and it makes no sense to speak of the self.

Second, relative to an object x , the self is x . That is, relative to x , the notion SELF delivers us the relational property *being identical with* x , thus in effect delivering the object x . It is important to note not only that the notion SELF does this but also that this is all it does. The notion SELF does not deliver consciousness or animateness or life or any other metaphysically substantive item. The truism “Everything is identical with itself” encapsulates the metaphysical significance of the self exhaustively. As already observed, in the example illustrated by (5) the self is simply the box in front of Bianca, and the self is not conscious, animate, or alive. If there were any philosophical problems to be properly called the “problems of the self,” the problems would have nothing essential to do with consciousness, animateness, or life. Any indication of consciousness, animateness, or life must come, if at all, not from “self” but from the pronoun to which “self” is suffixed.

It is worth repeating that the notion SELF delivers nothing non-relatively, that is, there is no such thing as the self in the absolute sense. The notion SELF makes sense only in relation to an already given something or rather. Relative to a given box, the self is the

box; relative to a given bat, the self is the bat; relative to a given person, the self is the person. Relative to the individual Bianca sees in the mirror, the self is the individual Bianca sees in the mirror. Relative to you, the self is you. Relative to me, the self is me. These statements are no more than statements of identity: the individual Bianca sees in the mirror is identical with the individual Bianca sees in the mirror, you are identical with you, and I am identical with me.

Unlike the suffix “-self,” the pronoun “her” does seem to indicate life, animateness, and consciousness. Does this lead to a formulation of any metaphysical problem in terms of “her”? In (4) the pronoun “her” (in “herself”) designates Bianca, so the *designatum* of “her(self)” is Bianca; she is identical with Bianca. What is problematic about this identity? Is the relation of identity itself problematic? Are the “problems of the self” really problems of identity? If so, they are not problems in philosophy of mind, as usually thought. They are problems in philosophy of logic and metaphysics, for identity is a logically important metaphysical relation. Bianca is herself. Bianca alone is herself. Identity holds between Bianca and herself, and not between Bianca and some individual other than herself. All this is trivial logically, is true metaphysically necessarily, and holds irrespective of whether Bianca is capable of carrying out any mental act or being in any mental state.

3. Epistemology

So far, we have seen that semantically the word “self” as a suffix works to indicate anaphoric co-designation with an appropriately salient preceding name, pronoun, or description, and that metaphysically it expresses the relation of identity. Many other words and phrases work to indicate anaphoric co-designation just as well, and many words and phrases express the relation of identity just as well, or perhaps even better. Thus, there is nothing particularly special or *sui generis* about the semantic or metaphysical significance of the notion SELF.

But perhaps there is something epistemologically special about the notion SELF. Perhaps when philosophers speak of the “problems of the self” and take them seriously, seeking answers by means of careful philosophical investigation, their concern may be epistemological rather than semantic or metaphysical.

Consider:

- (7) Bianca believes the pants she is wearing are on fire.
- (8) Bianca believes the pants Bianca is wearing are on fire.
- (9) Bianca believes the pants she herself is wearing are on fire.⁶

Assume that “she” in (7) designates someone other than Bianca, say Monica. Bianca sees Monica wearing pants and sees the pants being on fire. Bianca need not be able to see Monica clearly in order for this scenario to work. Bianca may be moderately nearsighted just so that she is unable to make out Monica’s facial features in detail, yet able to see her and her pants on fire. Also, nothing important hinges on Bianca and Monica being strangers to each other. The scenario works just as well even if they are postulated to be identical twins, for example.

Suppose then that they are identical twins and that Bianca is moderately nearsighted. Now let us change the scenario and suppose that, instead of looking at her twin sister Monica, Bianca sees herself in the mirror but is unaware of this fact and thinks she is looking at her twin sister. From her point of view, this new scenario is indistinguishable from the previous scenario. Bianca believes of the person she is looking at, viz. Bianca, that the pants she is wearing are on fire.⁷ Bianca would not say, “The pants Bianca is wearing are on fire” or “The pants I am wearing are on fire,” to express her belief. This illustrates the truth of (8).⁸

Now Bianca steps forward and realizes that there is a mirror in front of her and that she is looking at her own reflection in the mirror. She does not withdraw her assessment of the condition of the pants worn by the person she sees (in the mirror) and continues to believe Bianca’s pants are on fire; (8) remains true. What is different now is that she may now be said to have a belief whose content can be expressed by using the word “herself”; in addition to (8), (9) is now also true. It is important not to miss the point that (9) was not true before, when Bianca thought she was seeing Monica. (8) was true before and is true now, but (9) was untrue before and is true now. The identity relation between Bianca and the object (*res*) of her belief (*de re*) held before and holds now. Bianca’s doxastic situation before is not the same as her doxastic situation now. The only relevant corresponding difference between (8) and (9) is the presence of the suffix “-self” in (9). Does this not show some *sui generis* doxastic significance of the notion SELF?

No, it does not. The exact function of our use of “herself” in (9) is to distinguish the following (11) from (10):

(10) Bianca is an x such that Bianca is a y such that x believes the pants y is wearing are on fire.

(11) Bianca is an x such that x believes the pants x is wearing are on fire.

The difference “herself” in (9) makes is that the identity of the believer and the *res* of the belief *de re* in (11) but not in (10). This is clearly indicated by the two occurrences of one and the same variable “ x ” after “such that” in (11), whereas in (10) two distinct variables “ x ” and “ y ” occur and the identity “ $x = y$ ” does not occur.⁹ Thus, the word “herself” indicates identity between the believer and the *res* of the belief *de re*. This confirms my point that the function of the notion SELF is to indicate identity.

And yet, the effect of this functioning of the notion SELF falls short of necessitating the interpretation of (9) according to which Bianca would, for example, jump into the pool instead of pouring water on the mirror. The word “self” guarantees the identity between the believer and the *res*, but does not force that reading of (9) which describes the “Aha!” moment for Bianca. The possible impression to the contrary is the result of conflating the situation (11) describes with the situation described by the following sentence:

(12) Bianca expresses her belief correctly by saying, “My pants are on fire.”

This exhibits the unique epistemically relevant semantic function of the word “my,” as we shall see more carefully in the next section.

4. The First-Person Singular

In Bianca’s own words, the contrast between (8) and (9) is exhibited as follows:

(13) The pants she is wearing are on fire.

(14) The pants I myself am wearing are on fire.

When expressed from the vantage point of Bianca, the believer, “herself” gives way to “myself” but the force of the suffix “-self” can be observed clearly to vanish, for the contrast (14) carries with (13) is not diminished even when “myself” is deleted:

(15) The pants I am wearing are on fire.

Corresponding to “she is” in (13) we have “I am” in (15), and this syntactic difference exactly marks the contrast between (8) and (9). It is clear that the crucial word here is the first-person singular pronoun “I”; the nominative case is not necessary, as is made clear by the following rephrasing of (15):

(16) My pants are on fire.

The work “myself” appears to be doing in (14) is in fact done equally well by “I” in (15) and by “my” in (16). We can appreciate this fact even more clearly by noting that the transition from (13) to (14) may be expressed by Bianca’s succinct exclamation, “It’s me!” This invokes the first-person singular notion ME. It is the notion ME that is doing all the heavy lifting. The notion SELF is otiose.

Let us call the kind of awareness which a person x has when x is aware of someone who is identical with x and which x would express by saying, “It is me,” *me-awareness*. The awareness expressed by (13) need not be *me-awareness*, but the awareness expressed by (14) – (16) has to be *me-awareness*. *Me-awareness* underlies what is commonly referred to by philosophers as “attitudes *de se*,” and distinguishes them from mere attitudes *de re* where the *res* is the attitude holder.¹⁰ To avoid the association the phrase “attitude *de se*” has with the English word “self,” I recommend that we use the phrase “attitude *de me*” instead. We may then say that (14) – (16) express an attitude *de me* and that awareness *de me* underlies that attitude.¹¹

Some might object to my “*de me*” terminology. After all, when Bianca realizes that her own pants are on fire and jumps into the pool, her awareness is not of me and her belief about the flaming pants is not about me; she does not push me into the pool.

This complaint is clearly based on a simple misunderstanding. The terminology of “*de me*” is intended to replace the terminology of “*de se*,” as the latter literally means

“about the self.” I have argued that the word “self” only expresses identity (manifested linguistically as anaphoric co-designation) and therefore is too weak for the intended philosophical purposes. Attitudes *de se* in the literal sense are nothing more than attitudes about the attitude holder oneself, that is, attitudes *de re* where the *res* is the attitude holder. We need a term that captures the first-person singular mode. This is the spirit in which the terminology of “*de me*” is proposed. The target is not me, Takashi Yagisawa.

Though the complaint is based on an easily correctable simple misunderstanding, it does help reveal an important point about the *de me*, namely that the *de me*-ness of a belief *de me* is best understood meta-linguistically by noting the use of the first-person singular pronoun. When Bianca has the belief *de me* about having pants on fire, she would say, “My pants are on fire” rather than “Her pants are on fire,” if she were to express it. If I express my belief by saying, “My pants are on fire” rather than “His pants are on fire,” my belief is as *de me* as Bianca’s belief. The *res* is different, of course, but the way the *res* is presented to the believer is the same. If the “*de me*” terminology failed to capture the common way in which Bianca jumps into the pool as a result of her belief *de me* and in which I jump into the pool as a result of my belief *de me*, then my terminological recommendation would be undermined. The “*de me*” terminology presupposes the existence of one common immediate way in which different individuals are presented to themselves. Note that I do not presuppose that we are consciously aware of that way. It is sufficient for my purposes to presuppose that we are presented to ourselves in that common immediate way, whether or not the way itself is presented to us in any (meta-)way at all, let alone in any common (meta-)way.¹²

Note that it is possible to have an attitude *de me* via misidentification *de me*. It is possible for Bianca to believe that she is looking at herself in the mirror, even though there is no mirror and she is in fact looking at her twin sister Monica, thus misidentifying herself as the object of perception in the mirror. In such a case, Bianca has a belief *de me*, which she is ready to express by saying, “I am looking at myself,” while she is in fact not looking at herself.¹³

Let us not forget that in this case of misidentification *de me*, Bianca also has beliefs *de re* where the *res* is not herself but Monica. The belief she would express by saying, “She is looking at me, and she is me,” is doubly *de re*. First, it is *de me* and every belief

de me is *de re*. Second, it is *de re* where the *res* is Monica; Bianca's use of "she" in fact refers to Monica.

5. Subjectivity

The "problems of the self" may involve issues concerning the subjective/objective distinction, which is at the core of at least three major topics in philosophy: the Cartesian *cogito*, the Lockean secondary qualities, and the Nagelian qualia. They also have little to do with the notion SELF.

René Descartes intended his *cogito* argument to yield a secure foundation for knowledge without presupposing any pre-ordained outside authority.¹⁴ One of the crucial factors for this epistemological project was *a priori* certainty. An important necessary condition for the desired *a priori* certainty was that the foundation for knowledge be forthcoming from Descartes himself rather than from any other source. Would this important necessary condition be satisfied simply by having the source of the foundation for knowledge be identical with Descartes? No, it would not, any more than Bianca would jump into the pool simply by seeing the burning pants worn by a person who is identical with Bianca. Identity of the perceiver and the perceived is not enough. Bianca would need to be in a position to characterize the content of what she sees as "My pants are on fire." Her awareness of the wearer of the pants in flame needs to be awareness *de me*. Likewise, for the satisfaction of his quest for certainty, Descartes would need to be in a position to characterize the person who is the source of whatever appropriate foundational knowledge not simply as such-and-such a person, where such-and-such a person is identical with Descartes, but as "Me" ("*Moi*," "*Ego*"). It is crucial to him that the premise and the conclusion of his famous argument be put in the first-person singular form, not in the third-person form—"cogito" and "sum," not "cogitat" and "est." The identity between the argument-giver and the subject of *cogitare* and of *esse* is not enough. The key notion in the Cartesian *cogito* argument is ME, not SELF.

John Locke's secondary qualities are contrasted with primary qualities.¹⁵ Both kinds of qualities may be perceived by our senses, but primary qualities are supposed to be inherent in the external objects of which they are qualities and to be independent of the perceiver, while the secondary qualities are supposed to be not inherent in the external objects and to be dependent on the perceiver. In this sense the primary qualities are

supposed to be objective, while the secondary qualities are supposed to be subjective. What makes the latter (supposedly) subjective as opposed to objective has little to do with the identity of the perceiver. It is certainly not true that the perceiver who perceives a particular external object as red is identical with what is perceived, either the object or the color. Furthermore, the supposed subjectivity is not mere relationality. Whether someone is a mother is determined not just by her inherent properties but by an external factor, namely the existence of someone of whom she is the mother; motherhood is a relational matter. Redness is (supposedly) a relational matter in the same sense; whether some object is red is determined not just by its inherent properties but also by its relation to a perceiver; the existence of someone who has perceptual experience is crucial. But redness, unlike motherhood, is (supposedly) more than that. Whether a given object is red or not (supposedly) depends on the character of the perceptual experience of the perceiver, where the experience is (supposedly) private and irreducible. The subjectivity in question is inseparable from the phenomenology of qualia. The identity of the perceiver—the experiencer of the qualia—with herself has little to do with the supposed subjectivity of the Lockean secondary qualities in contrast to the primary qualities.

The topic of qualia is at the heart of Thomas Nagel's well-known discussion of what it is like to be a bat.¹⁶ Knowing all objective facts—facts determinable from the third-person point of view—about a bat would not suffice for knowing what it is like to be a bat. This is meant to illustrate the point that the first-person point of view is irreducible to the third-person point of view and is primitive. The first-person point of view Nagel is eager to emphasize is not captured by the mere identity of the person who attempts to imagine being a bat and the bat she is imagining herself to be. Suppose, as before, that Bianca thinks she sees Monica, while in fact she is seeing herself in the mirror. Suppose further that Bianca attempts to imagine the individual she sees, who she thinks is Monica, as a bat and fails. This is not what Nagel has in mind when he claims that we cannot imagine what it is like to be a bat. For the Nagelian purposes, Bianca needs to attempt to imagine herself as a bat under awareness *de me*. She needs to attempt to imagine a situation she would describe as “a situation in which I am a bat.” Again, the operative notion is ME, not SELF.

6. Ways of Belief

I have said that the function of the notion SELF is to indicate identity via anaphoric coreference. What is the function of the notion ME? Consider:

(17) Lois Lane says, “Superman is terrific.”

(18) Lois Lane says, “Clark Kent is terrific.”

“Superman” and “Clark Kent” rigidly designate the same individual (in fiction), and neither name is more special or privileged than the other, semantically speaking. Let us assume that Lois’s utterances reported by (17) and by (18) express the contents of her beliefs. Then these belief contents are, respectively, the singular proposition (represented by the ordered pair), $\langle \alpha, \text{being terrific} \rangle$, where $\alpha = \text{Superman}$ and the singular proposition, $\langle \alpha, \text{being terrific} \rangle$, where $\alpha = \text{Clark Kent}$. Since $\text{Superman} = \text{Clark Kent}$, these belief contents are one and the same.¹⁷ Now consider the following pair from the pants-on-fire example where Bianca sees herself in the mirror:

(19) Bianca says, “Her pants are on fire.”

(20) Bianca says, “My pants are on fire.”

Assume that (19) and (20) report Bianca’s utterances expressing the contents of her beliefs. Then these belief contents are, respectively, the singular proposition, $\langle \beta, \text{wearing pants on fire} \rangle$, where $\beta = \text{she}$, and the singular proposition, $\langle \beta, \text{wearing pants on fire} \rangle$, where $\beta = \text{Bianca}$. As with the (17) – (18) example in which $\text{Superman} = \text{Clark Kent}$, $\text{she} = \text{Bianca}$. Does this mean that as with the (17) – (18) example, the belief contents are one and the same? It is tempting to answer “No” by pointing out the different behavioral output of the respective beliefs: pouring water on the mirror versus jumping into the pool. But it is premature to yield to the temptation, for the case of Lois Lane also involves different behavioral output: admiring behavior toward the superhero wearing a cape and tights versus admiring behavior toward the reporter wearing a business suit and glasses.¹⁸ The “No” answer is motivated by the thought that it is impossible to account for the behavioral difference without postulating different belief contents, but it is this thought that is premature to accept at this point. It might be possible to account for the behavioral

difference resulting from the same belief content by introducing different ways of believing the same content.¹⁹

In cases like the Superman/Clark Kent example the different ways correspond (in semantically or pragmatically complicated manners) to different descriptions—not necessarily purely qualitative descriptions, but descriptions nonetheless. But it is unclear how to mimic this move in the case of Bianca, for it is unclear what description characterizes the way in which Bianca believes *de me* that her own pants are on fire. Perhaps “the subject of *this* experience”? But how should “*this* experience” be understood? Should it be understood as designating an experience *e* if and only if Bianca is in a position to designate it by saying, “I am having *e*”? If so, “*this* experience” cannot be understood prior to “I” being understood, that is, without the notion ME. As an account of the way characteristic of beliefs *de me*, as opposed to beliefs *de re* where the *res* is the believer, this is circular. At this point two sub-options are available: to specify the designation of “*this* experience” independently of the notion ME, or else to regard the designation of “*this* experience” as primitive.

The first sub-option is hopeless, for *thisness* in “*this* experience” is relative to the subject of experience. Like leftness and behindness, *thisness* is a relation rather than a property—or a relational property rather than a non-relational property.²⁰ When Bianca has an experience, which she designates by the phrase “*this* experience” as she is having it, she is in effect designating the experience by means of the notion ME as “*this* experience of *mine*.”

This consideration also shows that the second sub-option should be rejected, too. Since *thisness* is a relation and the designation of “*this* experience” must be determined under the activation of the notion ME, the designation cannot be primitive but must involve the relation of *being had by*, and we are back to Bianca’s designating the experience *e* in question by saying, “I am having *e*.”

A widely known general account of ways of believing is to reduce ways to objects, postulating intermediary objects of one kind or another; to believe a content in a certain way is to believe that content via a certain object, which works as an intermediary between the believer and the believed content. Obvious candidates include mental sentences, mental files, and guises.²¹ Let us take mental sentences as our representative example. Assume that we think in an internal representational system that is linguistic in

nature—a *language of thought*—and that to believe that p is to have a sentence of the language of thought meaning that p in the mind’s *belief box*. We may then identify ways of believing with sentences in the belief box; when Bianca merely has the belief *de re* with the *res* being Bianca, the sentence is (the translation into the language of thought of) the sentence “Her pants are on fire,” whereas when she has the belief *de me*, the sentence is (the translation into the language of thought of) “My pants are on fire.”

Unfortunately, this does not really solve our problem, for the same question comes back in a slightly different form: What is the relevant difference between (the language-of-thought translation of) the word “her” designating Bianca and (the language-of-thought translation of) the word “my” designating the same Bianca? This question mirrors the metalinguistic version of the original question about the difference between Bianca’s mere belief *de re* that her pants are on fire and Bianca’s belief *de me* that her own pants are on fire. We appear to have made little progress by introducing the extra machinery of the intermediary. Switching to any of the remaining intermediary candidates will not change the situation in any essential manner.

Let us therefore return to the idea that Bianca’s belief *de me* has a different content from her mere belief *de re* before the realization that it is her own pants that are on fire. The best known proposal of this type is due to David Lewis.²² Lewis makes his proposal within the possible worlds framework. The usual move within the framework is to define a proposition as a set of possible worlds.²³ The propositional content of Bianca’s utterance reported by (19), where “her” refers to Bianca, is the set of all and only those possible worlds at which Bianca’s pants are on fire. Since the propositional content of Bianca’s utterance reported by (20) is the same set, we do not get the required differentiation of content. To get the differentiation, the notion of *centeredness* of a world is introduced and belief as a relation to a proposition is redefined (or replaced by what is defined) as self-ascription of the property of inhabiting a member of a set of centered possible worlds.

To see how this is supposed to work, take the purely *de dicto* belief that someone’s pants are on fire. To have this belief is to self-ascribe the property of inhabiting a member of the set S_1 of centered worlds, where any world w is a member of S_1 if and only if at w someone’s pants are on fire. The idea of centeredness does no work. Next, take the belief *de re* about Monica that her pants are on fire. To have this belief is to self-ascribe the property of inhabiting a member of the set S_2 of centered worlds, where w is a member of

S₂ if and only if at w Monica's pants are on fire. Centeredness may appear to do no work here, either. But the appearance is illusory. The specification of Bianca's belief *de re* about Monica must be understood to involve Bianca's belief *de me* to the effect that the individual seen in front has pants on fire, and this requires Bianca to self-ascribe the property of inhabiting a world whose center bears such-and-such a (*de-re*-ness generating) relation, e.g. a direct perceptual relation, to someone in front whose pants are on fire. When Bianca believes *de me* that her own pants are on fire, she self-ascribes the property of inhabiting a world in which the center's pants are on fire. Centeredness has an obvious crucial role to play here.

This Lewisian idea supports the claim that Bianca's belief *de me* has a different content from her belief *de re* before the realization that it is her own pants that are on fire, insofar as we accept the centered worlds framework for belief content specification and identify the belief content with the self-ascribed property. This, however, does not mean that the idea of a way of believing has no role to play. On the contrary, it must play the important role of supplementing the Lewisian idea and filling a gap the Lewisian idea inevitably leaves.

Earlier in this section, we discussed a way of believing construed as an intermediary object intervening between the believer and the believed content. We dismissed this type of construal just before turning our attention to the Lewisian framework of centered worlds. But there is a different construal of a way of believing, which is not subject to the objections we saw against the objective construals and which is compatible with the centered worlds framework. It is the adverbial construal. It does not reify a way of believing as any kind of object at all but construes a way adverbially.

7. Primitive *De Me* and Adverbialism

There is a sense in which Bianca self-ascribes the property of wearing pants on fire when she merely believes *de re* of the woman she sees that her pants are on fire, where the woman, unbeknownst to Bianca, is in fact Bianca herself. Bianca does not realize that she is self-ascribing the property. This is mere "self-ascription *de re*," which is not self-ascription in the sense Lewis intends. When Bianca believes *de me* that her own pants are on fire, her self-ascription of the property is not merely self-ascription *de re*. The obvious way to distinguish mere self-ascription *de re* from self-ascription in the sense intended by

Lewis is to call the latter “self-ascription *de me*.” This terminological move highlights the fact that Lewis does not analyze belief *de me* reductively.²⁴

What is the relation between the *de-me*-self-ascribing believer and the center that makes the belief *de me*? It is not sufficient to say that the relation is identity, for the *de-me*-self-ascribing believer is the center by definition; irrespective of whether the belief is purely *de dicto* or merely *de re* or genuinely *de me*, the *de-me*-self-ascribing believer simply is the center. The identity relation needs to be built into the self-ascription *de me*. It is part of what it is for Bianca to *de-me*-self-ascribe the property of inhabiting a centered world at which the center’s pants are on fire that she identifies herself *de me* as the center. When put this way, the ineliminability of the *de me* is clear.²⁵ Thus, the *de me* is not reducible to any notion that is not *de me*. This, however, does not mean that we can say nothing further about the *de me* that is not trivial.

The only object proprietarily associated with the *de me* is the believer, but since the *de re* where the *res* is the believer need not be *de me*, the object associated with the *de me* does not determine the nature of the *de me*. Thus the nature of the *de me* is not object-wise but adverbial, i.e. it consists not in the “what” but in the “how” of representation.²⁶ What distinguishes the *de me* from the merely *de re* where the *res* is the believer is the subsumption of the *res* under the notion ME, and this subsumption is to be understood adverbially. The subsumption introduces no novel object but consists in *re-identification* of the *res* without any extra intermediary representation. When Bianca believes *de re* where the *res* is Bianca, but not *de me*, that her pants are on fire, she identifies the *res* merely as the individual she takes to be visually represented in front. When she comes to believe *de me* that her own pants are on fire, she comes to identify the *res* differently, i.e. in a way that is proprietarily *de me*, resulting in her coming to be disposed to use the words “my pants” rather than “her pants.”²⁷ This way can only be understood adverbially and primitively.

References

- Anscombe, G. E. M. (1975). ‘The First Person,’ in Samuel D. Guttenplan (Ed.), *Mind and Language* (pp. 45–65), Oxford: Oxford University Press.
- Babb, M. (2016). ‘The Essential Indexicality of Intentional Action,’ *The Philosophical Quarterly*, 66, 264, 439–57.
- Cappelen, H. & Dever, J. (2013). *The Inessential Indexical*, Oxford: Oxford University Press.
- Castañeda, H. N. (1966). ‘He’: A Study in the Logic of Self-Consciousness,’ *Ratio*, 8, 130–57.
- (1967). ‘Indicators and Quasi-Indicators,’ *American Philosophical Quarterly*, 4, 85–100.
- Chisholm, R. (1981). *The First Person*, Minneapolis: University of Minnesota Press.

- Descartes, R. (1641). *Meditations on First Philosophy*, Michael Moriarty (Trans.), Oxford: Oxford University Press (2008).
- Fodor, J. (1975). *The Language of Thought*, Cambridge, MA: Harvard University Press.
- Kaplan, D. (1989). ‘Demonstratives,’ in Almog, Perry & Wettstein (Eds.), *Themes From Kaplan* (pp. 481–563), Oxford: Oxford University Press.
- Kripke, S. A. (2011). ‘The First Person,’ in *Philosophical Troubles. Collected Papers Vol I*, Chapter 10, Oxford: Oxford University Press.
- Lewis, D. (1979). ‘Attitudes *De Dicto* and *De Se*,’ *The Philosophical Review*, 88, 513–43.
- Locke, J. (1689). *An Essay Concerning Human Understanding*, Peter H. Nidditch (Ed.), Oxford: Oxford University Press (1975).
- Magidor, O. (2015). ‘The Myth of the *De Se*,’ *Philosophical Perspectives*, 29, 249–83.
- Nagel, T. (1974). ‘What Is It Like to Be a Bat?’ *The Philosophical Review*, 83, 435–50.
- Perry, J. (1979). ‘The Problem of the Essential Indexical,’ *Nous*, 24/5, 723–34.
- Prosser, S. (2015). ‘Why Are Indexicals Essential?’ *Proceedings of the Aristotelian Society*, 115, 211–33.
- Quine, W. V. (1969). ‘Propositional Objects,’ in W. V. Quine, *Ontological Relativity and Other Essays*, New York: Columbia University Press.
- Ricanati, F. (2012). *Mental Files*, Oxford: Oxford University Press.
- Salmon, N. (1983). *Frege’s Puzzle*, Cambridge, MA: The MIT Press.
- (1986). ‘Reflexivity,’ *Notre Dame Journal of Formal Logic*, 27, 3, 401–29.
- Stalnaker, R. (1984). *Inquiry*, Cambridge, MA: MIT Press.
- Williams, B. (1973). *Problems of the Self*, Cambridge: Cambridge University Press.
- Yagisawa, T. (2010). *Worlds and Individuals, Possible and Otherwise*, Oxford: Oxford University Press.
- (2017). ‘The Meta-Representational *Me*,’ in Boran Berčić (Ed.), *Perspectives on the Self* (pp. 355–69), Rijeka, Croatia: University of Rijeka Press.

Footnotes

- ¹ This is an expanded version of the talk, “A Deflationary Conception of the Self,” which I gave at the conference, *Aspects of Self: A Workshop*, at Kyoto University on May 14, 2018. I thank the organizer, Professor Yasuo Deguchi, and the audience at the talk.
- ² The best known publication with this phrase as its title is Williams (1973).
- ³ Other typical functions of the suffix “-self” include emphasis, as in “Never mind about Monica. I heard it from Bianca herself,” and contrast, as in “I don’t know about you but I like Dadaism myself.” When used in these ways, the suffix “-self” does not even appear to give rise to any serious philosophical problem.
- ⁴ A similar but more limited point is made by Elizabeth Anscombe in Anscombe (1975). Anscombe’s point is limited in that (i) it is made only by means of illustrative examples without a general formulation like the one I give in the last paragraph, and (ii) it fails to say that selfhood has nothing to do with animateness or consciousness.
- ⁵ Or their stylistic variants: “() is Φ herself,” “() is Φ himself,” “() is Φ itself,” “You are Φ yourself,” or “I am Φ myself.”
- ⁶ The pants example is due to David Kaplan in Kaplan (1989, pp. 533–537).
- ⁷ This illustrates the point Anscombe makes in the work mentioned in footnote 4.
- ⁸ Of course, this imagined scenario is not the only kind of scenario illustrative of the truth of (8), but it works well to demonstrate the difference “self” apparently makes.
- ⁹ See Salmon (1986).
- ¹⁰ The phrase “attitude *de se*” is due to David Lewis in Lewis (1979). What are called “attitudes *de se*” have been extensively discussed by philosophers; in addition to Lewis (1979), the classic articles on the topic include Catañeda (1966, 1967), Perry (1979), and Kaplan (1989).

¹¹ David Kaplan's celebrated semantics for the first-person singular pronoun in Kaplan (1989) may be a satisfactory account of the linguistic meaning of "I" ("me," "my," "mine") capable of explaining the *a priority* and contingency of "I exist" among other things, but as observed by Saul Kripke in Kripke (2011), it leaves room for the separate issue of the first-person singular *concept*, which is the issue I am addressing as the issue of the *de me*. Kripke argues that the first-person singular concept is ineliminable and irreducible. I agree.

¹² This is borne out by the arguments by a number of philosophers concerning the close conceptual link between awareness, attitudes, and intentions *de me*. See the literature cited in footnote 10. See Cappelen & Dever (2013) and Magidor (2015) for skepticism about the link, and Babb (2016) for a response.

¹³ This does not work with realization, for the verb "realize" is factive. If Bianca realizes she is looking at herself in the mirror, then it follows that her identification *de me* is correct, i.e. she is in fact looking at herself in the mirror.

¹⁴ See Descartes (1641).

¹⁵ See Locke (1689).

¹⁶ See Nagel (1974).

¹⁷ This presents a problem if you think that (17) and (18) exhibit a failure of substitutivity *salva veritate*.

¹⁸ The same basic point is made and elaborated in Cappelen & Diver (2013).

¹⁹ See Salmon (1983) for a general strategy of this sort.

²⁰ Compare s-e relations in Prosser (2015).

²¹ See for example, Fodor (1975), Recanati (2012), and Salmon (1983).

²² In Lewis (1979), following Quine (1969).

²³ See Stalnaker (1984) for a defense of this move, and Yagisawa (2010) for arguments for the need to include not only possible worlds but impossible worlds as well.

²⁴ Nor does he intend to offer such a reductive analysis.

²⁵ Lewis would agree that his proposal is not intended to eliminate what I am calling the *de me*. Roderick Chisholm makes a non-reductive proposal concerning the first-person singular in the same spirit as Lewis in Chisholm (1981).

²⁶ See Yagisawa (2017) for detailed consideration of this point.

²⁷ I call this the "adverb to noun transition" in Yagisawa (2017).