

Brentano's Most Striking Thesis: No Representation Without Self-Representation

Uriah Kriegel University of Arizona

Forthcoming in D. Fisette and G. Fréchet (eds.), *Themes from Brentano*. Amsterdam: Rodopi

1. Introduction: Theses from Brentano

It is well-known that Brentano took intentionality to be the mark of the mental (Brentano 1874 Book II Ch.1), though it is less clear whether 'mark' talk is supposed to underlie merely an extensional claim or a stronger, "intensional" claim. The extensional claim would be that all and only mental states are intentional states. The stronger claim would be that mental states are mental *in virtue of* being intentional. Unfortunately, the text itself does not appear to decisively support one interpretation over another, though arguably the tenor of Brentano's discussion suggests the stronger claim: intentionality is the essence of mentality.

It is much less known that Brentano also held a thesis of coextension between the mental and the conscious. Obviously, no non-mental state could be conscious. But interestingly, Brentano also argues for the converse: no mental state can be non-conscious. Thus the final Section 13 of Chapter 2 of Book II of the *Psychology* reads as simply: "There is no unconscious mental activity." This was not a striking thesis at the time, mind you: Freud's "The Unconscious" was not to appear for another forty years (Freud 1915).¹ It follows, in any case, that the mental, the conscious, and the intentional (or representational) were all co-extensive for Brentano.

A much more striking Brentanian thesis concerns the nature of consciousness itself. Over the past decade, several partially interpretive works have appeared that cast Brentano as a forerunner of current-day *self-representational* theories of consciousness,

according to which a mental state is conscious when, and only when, it represents itself in the right way (see, e.g., Caston 2002, Kriegel 2003). Brentano himself does not typically put things in terms of representation, but instead in terms of *presentation* (*Vorstellung*) and intentionality. In keeping with the current-day tendency to use the terms “intentionality” and “representation” synonymously, however, we may safely take a passage as this to mean that every conscious state self-represents (Brentano 1874: 153-4):

[Every conscious act] includes within it a consciousness of itself. Therefore, every [conscious] act, no matter how simple, has a double object, a primary and a secondary object. The simplest act, for example the act of hearing, has as its primary object the sound, and for its secondary object, itself, the mental phenomenon in which the sound is heard.

Here as before, there is no decisive textual evidence that the claim is not merely extensional – that all and only conscious states happen to self-represent.² But the tenor of the discussion again suggests a stronger thesis – that conscious states are conscious *in virtue of* self-representing (and to that extent that self-representation is the essence of consciousness).

So now we have a four-way coextension in Brentano: intentionality/representation, mentality, consciousness, and self-representation all coextend for him. The coextension of the first and last of these is particularly surprising, however: Why would representation coextend with self-representation? Why would the very existence of one order of representation (namely, first-order) be conditioned by the existence of another order of representation (second-order)? Brentano is not shy about this claim of coextension. Section 8 of Chapter 2 of Part II of the *Psychology* reads: “A Presentation and the Presentation of that Presentation are Given in One Mental Act.” Thus Brentano appears fully committed, in a considered way, to the idea that there could be no representation without self-representation. To my mind, this is Brentano’s most striking thesis: that the very possibility of a mental state representing an apple, say, depends on the state representing itself to represent an apple.

In what follows, I want to argue that this claim, which sounds odd to our modern sensibilities, is actually deeply insightful. The insight can be briefly illustrated as follows. When I consciously think of the Sydney Opera House, I am in an internal state that

represents the Sydney Opera House to me. The state thus instantiates two representation relations: it bears a representation-*of* relation to the Opera House and a representation-*to* relation to me. Although virtually all work on mental representation in the past century has focused on the nature of the representation-*of* relation, little if any illumination has been offered for the representation-*to* relation and its connection to representation-*of*. I want to argue that close examination of the connection between these two relations reveals, first, that in some sense there could be no representation-*of* without representation-*to*, and secondly, that representation-*to* is plausibly implemented in self-representation.

The plan for the rest of the paper is as follows. In §2, I will argue that there could be no representation-*of* without representation-*to*. There can be *token* representations-*of* that are not representations-*to*, but they must betoken a type of representation some tokens of which are both representations-*of* and representations-*to*. In §3, I will offer an analysis of “*x* represents *y* to *z*” according to which it means (more or less) that *z* has a representation of *x* representing *y*. In §4, I will note that this generates a regress of representations which can only end with self-representing representations. If my thought of the Sydney Opera House represents both the Opera House and itself, then it is both a representation-*of* and a representation-*to* without requiring the postulation of any further representation. The upshot is that there could be no representation without self-representation: in a world without self-representing representations there would be no representation at all. Brentano’s most striking thesis is true.

2. No Representation-*of* without Representation-*to*

The notion that there could be no representation-*of* without representation-*to* might strike the reader as a non-starter. The world is awash with natural signs, and these can legitimately be said to represent what they do even if there is nobody *to whom* they represent. The rings on the tree’s trunk represent the tree’s age, the traces in the snow represent the culprit’s path, etc. Moreover, as Dretske (1988: 55) notes, what such natural signs represent is independent of what they are *taken* to represent. The tree rings would represent the tree’s age even if nobody took them to represent the tree’s age (or anything

at all for that matter). They would represent the tree's age even if everybody took them to represent the tree's height (or took them to be non-representational).

There may be a subtle fallacy in this reasoning, however. An ant strolling about a deserted island's beach may form the inscription "cat" in the sand, and some rocks on Venus may accidentally be arranged "cat"-wise. We have no trouble saying that these structures represent cats, but we realize that, in and of themselves, there is nothing about these "cat" structures to make them mean cats in English rather than lizards in a language yet to be invented. These are instances of what we may call "representation by courtesy."³

What is representation by courtesy? The "cat"-ly inscriptions represent what they do, and at all, only because they are graphically type-identical to other inscriptions. There is, moreover, an asymmetry between these inscriptions and, say, my own deliberate "cat" inscriptions. The former represent because the latter do, but not conversely. In other words, the former represent by courtesy of the latter. Thus the token Venutian rock formation represents cats only because it betokens a type of representation some tokens of which represent cats. It follows that if those other tokens did not represent, the Venutian rock formation would not represent either.

In general, we may construe representation-by-courtesy as follows: R_1, \dots, R_n represent by courtesy of R_{n+1}, \dots, R_{n+n} just in case (i) R_1, \dots, R_n belong to the same representation type as R_{n+1}, \dots, R_{n+n} , (ii) R_1, \dots, R_n would not represent if R_{n+1}, \dots, R_{n+n} did not, and (iii) it is not the case that R_{n+1}, \dots, R_{n+n} would not represent if R_1, \dots, R_n did not. It is very likely to turn out that (iv) R_{n+1}, \dots, R_{n+n} are paradigmatic or "core" tokens of the representation type to which they and R_1, \dots, R_n belong. This may or may not force us to add this explicitly as a fourth condition, a condition of "paradigmaticness."

With the notion of representation-by-courtesy in mind, we may wonder whether unseen, undiscovered tree rings represent tree ages merely by courtesy of the discovered and seen tree rings. This would mean that representations-of that are not also representations-to represent only by courtesy of representations-of that *are* also representations-to. If so, the former would not represent if the latter did not.

The slogan "No representation-of without representation-to" should not be read, then, as the thesis that every token representation-of must also be a representation-to.

Rather, the thesis is that every token representation-of must betoken a representation type some tokens of which (perhaps paradigmatic ones) are both representations-of and representations-to.⁴ In the remainder of this section, I argue for this thesis.

When we look at a duck-rabbit picture, it represents to us a duck and a rabbit (though not simultaneously). Suppose Smith has suffered a duck-related childhood trauma that causes her to repress all incoming duck-related information. It is reasonable to say that the same picture represents a rabbit to Smith, but does not represent a duck to her. Unlike Smith, Jones simply suffered a brain lesion and consequently cannot recognize ducks. To her too, the picture represents a rabbit but not a duck. Importantly, in the case of Jones we may suppose that it is *nomologically impossible* for her to recognize the duck representation in the picture.

Suppose now that there were pictures of an animals which it was nomologically impossible for *any* human to recognize as such. More generally, imagine a natural sign S of a signified entity E, such that it is nomologically impossible for all humans to recognize that S signifies E. Worse, imagine it is nomologically impossible for *any* nomologically possible sentient creature to recognize S as signifying E. Does S represent E? I entreat the reader to consult her own intuition, but mine is a categorical No.

If it be accepted that S does not represent E, despite bearing the right informational/teleological relation to E, the question arises, Why is S not a representation of E? The only plausible answer seems to be that S is not a representation of E because it is not, and cannot in principle be, a representation of E *to* anyone. It appears, then, that S cannot be a representation of E if it is nomologically impossible for S to be a representation of E to some nomologically possible sentient creature. Thus the nomological possibility of representation-to is a necessary condition for actual representation-of.

It is quite improbable that there is a natural sign type all of whose actual tokens do not represent to anyone even though its nomologically possible tokens do. But suppose there is. It is still the case that these tokens are representations-of only because they belong to the same type as the nomologically possible tokens that are also representation-to.⁵

Furthermore, the converse does not hold: it is not the case that the nomologically possible tokens that are also representations-to are representations-of because they belong to the same type as the actual representation-of that are not also representations-to. It follows that the actual representations-of represent only *by courtesy* of the relevant nomologically possible representations-of.

It might be objected that our tentative fourth condition on representation-by-courtesy does not hold here. The relevant nomologically possible tokens cannot be paradigmatic, precisely because they are merely possible.

However, the assumption that actuality is necessary for the status of paradigm is misplaced. There are no actual perfect circles, only possible ones. Yet perfect circles are surely the paradigms of circularity. In the case of natural representations, it likewise seems that those that represent what they do *to someone* are paradigmatic.

Another objection would be that there are different notions of representation, some of which may require little more than brute covariation relations. In philosophical contexts, “representation” is used technically, and so can surely be used, in some context, in that sort of thin sense.

This objection is grounded in a correct observation but misapplies it. It is correct that the notion of representation is technical and can be used in any number of ways. But there is only one sense of representation that concerns me here, namely, the sense in which it can be said to capture *intentionality*. The discussion in the present paper is not interested in every possible notion of representation, but in the notion that is of relevance to Brentano’s most striking thesis, which is a thesis about intentionality.⁶

There is also a more principled methodological reason for giving center stage to this thicker notion of representation. After all, this is the only form of representation we really know from the first-person point of view – the only form of representation we are *personally acquainted with*, if you will.⁷ Other forms of representation are theoretical constructs and can thus deploy all manners of technical notion of representation. But the notion of representation that captures intentionality is not merely a theoretical construct. It is something we are directly acquainted with, and which we take to have the status of representation not merely in a purely theoretical or technical sense.

3. What is Representation-to?

As mentioned in §1, discussions of representation have focused on representation-of to the virtual exclusion of representation-to. As a result, we have a very murky notion of representation-to to work with. In this section, I will try to say something constructive about representation-to.

A minimalist might attempt to account for representation-to simply in terms of *ownership*. For a state or feature represents something to someone is simply for that state or feature to represent what it does and be a state or feature *of* the relevant “someone.” On this account, “ x to represent y to z ” can be analyzed as follows: (i) x represents (is a representation of) y and (ii) x occurs in z (or: x “belongs to” z in the appropriate sense).

The minimalist account is highly implausible. The tree rings (i) represent the tree’s age and (ii) are a feature of the tree, but they do not represent the tree’s age *to the tree*. So conditions (i) and (ii) can be fulfilled even when representation-to does not occur. Conversely, the tree rings do not represent the tree’s age to the tree, but they do represent the tree’s age to Tania the botanist. Yet they are not a state or feature of Tania. So representation-to can occur even when condition (ii) is not satisfied.

Note, more generally, that representation-to comes in two varieties: one where x is an internal state of z and one where it is not. As I think of the Sydney Opera House, my thought represents the Opera House to me. And as I look at my postcard of the Sydney Opera House, the postcard represents the Opera House to me. Both cases involve representation-to, but clearly they are crucially different. We may call the first variety *first-person* representation-to and the second *third-person* representation-to.

It is reasonable to expect that the two varieties would require two different accounts. After all, they seem to be slightly different phenomena. In particular, the first-person variety seems somehow more *fundamental* than the third-person one. At the same time, it is reasonable to expect that there will be something in common in the two accounts, since there is something in common in the two phenomena accounted for. It should therefore be a constraint on the adequacy of an account of representation-to that it account for this commonality.⁸

A modified minimalist account might restrict the analysans to conscious and/or sentient creatures. So “ x represents y to z ” is to be analyzed as (i) x represent y , (ii) x occurs in z , and (iii) z is a conscious and/or sentient creature.⁹ Nothing represents anything to trees, but internal states of conscious/sentient creatures may represent to them.

This analysis is still implausible. For one thing, it does not address the possibility of third-person representation-to. As we noted, in such representation-to condition (ii) is not satisfied. In fact, however, the account is unsuccessful even as an account of first-person representation-to only. States of a conscious and sentient creature’s skin, say, can represent some environmental feature without representing it *to* the creature. Thus, Brown’s goosebumps represent that the ambient temperature is below 6C. The state of being goosbumped thus (i) represents the ambient temperature and (ii) occurs in Brown, and (iii) Brown is a conscious and sentient creature. Yet it does not represent the temperature *to* Brown. (At least this is so if we stipulate that, throughout her life, Brown remains unaware that her goosebumps constitute the natural sign they do.)

Moreover, even if we restrict ourselves to *mental* internal states, conscious and sentient creatures have unconscious states that represent but not *to* them. Blain the blindsight patient has an unconscious perceptual representation of an orange in his left visual field. Blain’s perception (i) represents the orange and (ii) occurs in Blain, and (iii) Blain is a conscious and sentient creature. Yet clearly it does not represent the orange *to* Blain.¹⁰

A more sophisticated account of representation-to will focus on the *use* to which representations are put. When the tree rings represent the tree’s age to Tania, at least part of what is involved is that Tania can *use* the tree rings *as* a representation of the tree’s age. She can use that information, e.g., to make inferences, decisions, etc. Likewise, my thought, which represents the Sydney Opera House to me, is poised for free use in action guidance, reasoning, deliberation, etc. An internal state can thus be said to represent *to* the subject when it is poised for this sort of use *by* the subject. The analysis here offered is basically this: x represents y to z iff (i) x represents y and (ii) z can use x as a representation of y .

One advantage of this use-based account of representation-to is that it identifies a commonality of the first-person and third-person varieties. At the same time, the use will be cashed out differently in each case. In the case of third-person representation-to, the representation is “used” somewhat as a tool is used. The subject “uses” the representation as she may use a hammer. In the case of first-person representation-to, the use is rather a matter of the representation’s *functional role*: x represents y to z when x has the right functional role in z ’s mental life.

Interestingly, it was once popular to account for representation-of in terms of functional role. That was the idea in *functional role semantics*.¹¹ Beyond the variety of technical problems this approach faced, ultimately it came up against the principled problem that “functional role semantics” appears to be an oxymoron: functional role is a matter of relations *among* mental states, whereas semantics, representation-of, is a matter of relations to (typically) *extra-mental* entities. It appears impossible to get the latter out of the former.

Since representation-to is not fundamentally a matter of relation to extra-mental entities, a functional role account of representation-to is much more plausible. In particular, one may employ functional role semantics that incorporate a “consumer semantics” component.¹² This view considers as the key aspect of a representation’s functional role its availability for consumption by certain cognitive modules. Presumably, the account would designate as central the representation’s availability to an executive control module, or perhaps focus on the representation’s *global* availability to a number of high-level modules.

The use-based account does run into a principled problem of its own. Having a certain functional role is a *dispositional* property, but being a representation-to seems to be a *manifest*, hence non-dispositional, property. Thus, for a mental state to be *poised* for use, it need not actually *be* used; for it to be *available* to certain modules, it need not actually be *availed of* by them. But representing something to someone is not just a matter of things possibly happening. Something does actually happen when my thought represents the Sydney Opera House to me. So representation-to cannot amount to nothing more than functional role, since a non-dispositional property cannot amount to a mere disposition. In fact, it is quite likely that representation-to is the *categorical basis* of the

relevant functional role. My thought of the Sydney Opera House is poised for free use *by me because* it represents what it does *to me*, not the other way round.¹³

Another problem with the use-based account is that, despite initial appearances, it actually fails to account for the commonality of first- and third-person representation-to. Although we can employ the word “use” for both, it does not seem to be in the same sense. In the case of third-person representation-to, “use” is employed literally: the tree rings’ representation of the tree’s age is indeed used as a hammer might. But in the case of first-person representation-to, “use” is employed most certainly *metaphorically*. I do not “use” my thought’s representation of the Opera House as I might a hammer. Perhaps when I consciously and deliberately make inferences on the basis of my thought, I could be said to “use” it in a literal sense. But such conscious, personal-level inferences account for a marginal part of the relevant functional role.

In the face of these problems, I would like to make what may sound on first hearing like a silly suggestion: a representational account of representation-to. The idea is to account for representation-to purely in terms of representation-of. Thus, the fact that the tree rings represent the tree’s age to Tania lends itself quite naturally to the following interpretation: Tania has a representation of the tree rings’ representation of the tree’s age. More generally, x represents y to z iff (i) x is a representation of y and (ii) z has a representation of x ’s representation of y . Now, for z to have a representation of x ’s representation of y is for z to have, or be in, some internal state w , such that w represents x ’s representation of y . So the overall analysis may be stated as follows: x represents y to z iff (i) x is a representation of y and (ii) z has a(n internal state) w , such that w is a representation of x ’s representation of y .

A general argument for the representational account is this: x may be said to represent y to z only if z is *aware* of x ’s representation of y ; awareness of something requires representation of it (for example, I cannot be aware of a table without having an internal representation of it); therefore, x may be said to represent y to z only if z has a representation of x ’s representation of y .

Condition (ii) in the representational account does not require – though it may be strengthened to do so – that z ’s representation be propositional or conceptual. As (ii) stands, it does not require that z have a representation *that* x represents y , nor even that z

have a representation of x as a representation of y . The representation of x 's representation of y may perfectly well be altogether non-conceptual. But should we come upon a good reason to introduce a stronger requirement – say, that the representation be conceptual – we could readily tweak (ii) to make it so.

4. The Necessity of Self-Representation

One advantage of the representational account is that it successfully accounts for the commonality between first- and third-person representation-to. When my postcard represents the Opera House to me, this is because I have a representation of the postcard's representation of the Opera House. And when my thought represents the Opera House to me, this is because I have a representation of my thought's representation of the Opera House.

This raises an immediate worry, however. The account appears to lead straightforwardly to a vicious regress, at least for first-person representation-to. When x represents y to z , this is because z has (or is in) an internal state w , such that w represents (is a representation of) x 's representation of y . But given that representation-of requires representation-to, we may expect w not just to represent what it does, but to do so *to z*. And off we are on a regress.

Once the specter of regress afflicts first-person representation-to, it is bound to haunt third-person representation-to as well. Let us suppose that the tree rings represent the tree's age to a machine whose function is to garner and process information on trees' age. The rings represent as they do because the machine has a representation of the tree rings' representation of the tree's age. The machine may in its turn represent what it does *to Blain the blindsighted*. It would do so in virtue of Blain's unconscious perceptual representation of the machine's representation of the tree rings' representation of the tree's age. This unconscious perceptual representation may represent what it does *to Nora the neuroscientist*, who is observing Blain's brain with functional Magnetic Resonance Imaging (fMRI). The unconscious representation represents as it does because Nora has a representation of Blain's representation of the machine's representation of the tree rings' representation of the tree's age. But it may then be the case that Nora's representation

represents what it does *to* Nora. The chain of third-person representations-to thus comes to an end with a first-person representation-to. So if a vicious regress does afflict Nora's first-person representation-to, the affliction will transmit to all those third-person representations-to chainlinked thereto.

It might be thought that this looming regress problem could be skirted by purely technical considerations. Recall, the thesis argued for in §2 was *not* that every *token* representation-of is also a representation-to. We have allowed there to be token representations-of, perhaps very many of them, that are not representations-to. If so, *z*'s internal state *w*, which is a representation of *x*'s representation of *y*, need not necessarily be a representation *to z*. Thus, when my thought represents the Opera House to me, I must have a representation of my thought of the Opera House, but that second-order representation need not represent what it does *to me*.

This technical resolution of the regress problem seems somehow unsatisfactory, however – as though it does not get to the root of the issue. After all, if the second-order representation does not represent anything *to me*, if it is characterized by the sort of 'blindness' of brute representations-of, it is hard to see how the first-order representation would inherit its representation-to component therefrom. Internal states that represent but not *to* their subject are akin to the representations the blindsighted hosts somewhat impersonally. Such an impersonal representation seems ill-suited to bestow the representation-to status on my thought of the Opera House.

Relatedly, recall that although we allowed representations-of to be such even when they are not representations-to, we allowed them to be representations-of only *by courtesy*. It is difficult to see how a full-blown representation, which is both a representation-of and a representation-to, can be such in virtue of being 'represented' by something that is a representation only by courtesy. (I use inverted commas around "represented" precisely because the representation in question is only such by courtesy.)

Moreover, there is nothing in the above technical consideration to preclude the chance occurrence of an infinite chain of representations. Although we have allowed representations-of to *not be* representations-to, we have also allowed them to *be* representations-to. Even if the likelihood that a particular thought of mine would implicate an infinity of higher-order representations is negligible, this is a purely

contingent matter. There is thus a nomologically possible world in which my thought represents the Opera House to me in virtue of a second-order representation-of that happens to be also a representation-to, the second-order representation is a representation-to in virtue of a third-order representation-of that also happens to be a representation-to, and so on *ad infinitum*. The infinite regress is thus nomologically possible.

Furthermore, a regress may be vicious enough without being infinite. The possibility, whose likelihood is *not* negligible, that a thought of mine would implicate, say, six orders of representation, is problematic. It is problematic because it is empirically implausible that we may happen to entertain six orders of representation on a particular occasion of thinking of an opera house.

For all these reasons, we should not rest content with the technical resolution of the regress problem, but instead seek a resolution that is not so shrouded in suspicion. I turn now to such a resolution.

We said above that when x represents y to z , z must have a token representation w of x 's representation of y . Our discussion thus far has assumed that it must be the case that $w \neq x$. But this is precisely the assumption that generates the regress. If we allow that in (some) cases of first-person representation-to $w = x$, the regress does not get started. For then x represents y to z because (i) x represents y and (ii) x represents x 's representation of y . No second-order representation is introduced, so the regress is stopped at its second step. Thus, my thought represents the Sydney Opera House *to me* in virtue of the facts that (i) the thought represents the Sydney Opera House and (ii) the thought represents that it, itself, represents the Sydney Opera House.¹⁴

On this view, what is involved in a token x representing y to z is x 's falling under two distinct types of representation-of: the type Representation-of- y and the type Representation-of-representation-of- y .¹⁵ The latter type can perfectly well be betokened by a token state that does *not* betoken the former type. In that case, no self-representation comes into play. That is in fact what happens when $w \neq x$. There, w betokens the type Representation-of-representation-of- y but not the type Representation-of- y , and so does not self-represent.

The emerging picture involves representations-of that are also third-person representations-to in virtue of being represented by further representations-of in a chain of representations that ends in representations-of that are first-person representations-to; these are first-person representations-to in virtue of being representations of themselves. Thus representation-to bottoms out in representation-of, but of a very special kind, namely, self-representation. The outcome can thus be captured in the slogan “No representation without self-representation.”

This picture raises problems of its own. An immediate set of questions arises in connection with the very notion of self-representation. What is self-representation? What does it mean for a thought to represent itself? What representations are self-representing? How does a representation come to self-represent?

Addressing all these questions would be impossible here (see Kriegel 2009 for partial addressing of many). But note that the Brentanian thesis that all and only conscious states are self-representing would be eminently relevant here. For it would cast conscious experiences as the stopgaps of regress, bestowing representational status on all those representations-of chainlinked to them. This line of reasoning suggests a tight connection between consciousness and representation, of the sort proponents of the so-called “Phenomenal Intentionality Research Program” seem to have in mind (see Horgan and Kriegel 2012 for a collection of relevant essays).

Other kinds of mental states have at times been offered as self-representing. Burge (1988) argued that “cogito-like” thoughts are self-representing (or have a second-order content “locked onto” their first-order content). Shoemaker (1996) claimed that certain “available” beliefs self-represent (they perform “double duty” as first-order beliefs that p and second-order beliefs that one believes that p). Such cogito-like thoughts or available beliefs may turn out to be conscious states, but in any case they could serve as stopgaps themselves.¹⁶

If there are mental states that are self-representing, there must be a way they come to self-represent and there must be a possible account of what it means for them to self-represent. Such an account would ultimately have to be provided if we are to fully understand representation-to and representation-of. But for present purposes it suffices we note simply that the account must exist. That would suggest that the questions raised

above do not point to inherent defects in the picture proposed here, but should rather be thought of as invitations to elaborate the model.

Another sort of objection may be that the original regress problem has only been replaced by another. Consider x 's representation, not of y , but of its own representation of y . Does x represent its own representation of y to z , or not? If it does, then it would seem that it must do so in virtue of representing its own representation of its own representation of y . If it does not, then it is still hard to see how such an 'impersonal' representation of the fact that x represents y can bestow a representation-to component on x .

One response to this dilemma is to embrace the new regress but note that it is only a regress of *types*, not of *tokens*, and as such is less vicious. Even if we agree that x represents not only its own representation of y , but also its own representation of its own representation of y , its own representation of its own representation of its own representation of y , and so on, it still remains the case that x is a single token state. And we must realize that falling under a type, or betokening a type, is not some sort of activity that a token might be said to *expend energy* on. There is no extra "cognitive burden" involved in betokening three types rather than two or an infinity rather than five.

At the same time, there is certainly something odd about the notion of an infinitely typed thought. Intuitively, there has to be something about a token in virtue of which it betokens the types it does. Thus if a token mental state betokens infinitely many representational types, there should be something about that token that makes it the case that it betokens the representational type of the 837th order, and it is hard to see what that would be.¹⁷

A better response employs the token/type distinction as well, but differently. When we contemplated the "technical" resolution of the regress problem above, we were uncomfortable allowing a representation-of to qualify as a representation-to in virtue of being represented by a representation-of that is itself *not* a representation-to. We were uncomfortable allowing this mainly because representations-of which are not also representations-to represent only *by courtesy*. Observe, however, that being a representation by courtesy (or not) is not a property of representational *types*, but rather of representational *tokens*. So, once a token becomes a representation-to in virtue of

betokening a second-order representation type in a way that makes it self-representing, it does not represent only by courtesy, but in and of itself. We may thus allow that x represents y to z in virtue of representing its own representation of y , since it represents its own representation of y in and of itself and not merely by courtesy.

This may suggest an “attenuation” of the representational analysis of representation-to offered in §2: x represents y to z iff (i) x represents y , (ii) z has a(n internal state) w that represents x 's representation of y , and (iii) w represents not merely by courtesy. On the other hand, this analysis would strip some third-person representations-to of their status as such. A more cautious analysis would rest on a disjunctive third condition: x represents y to z iff (i) x represents y , (ii) z has a(n internal state) w that represents x 's representation of y , and (iii) either (a) w represents not merely by courtesy, or (b) there is a z^* , such that w represents x 's representation of y to z^* .

Conclusion: Brentano the Teacher

Husserl once wrote a touching essay entitled “Reminiscences of Franz Brentano” (see McAlister 1977 Ch.2), which offers a warm and respectful portrait of someone who for him was clearly first and foremost a teacher. It is well-known that Brentano taught a whole generation of leading philosophers, including Husserl, Meinong, Twardowski, Stumpf, Marty, and others. At one time, later generations were interested in Brentano mostly for his thesis that intentionality is the mark of the mental, and sometimes also for his insistence that intentionality cannot be understood in physical terms. In recent years, however, it has turned out that Brentano still has much more to teach us. His account of consciousness in terms of self-directed intentionality has been developed and defended by a variety of authors, and is now among the leading options for a philosophical theory of consciousness in analytic philosophy of mind. His mereologically sophisticated account of the unity of consciousness has the potential to illuminate many recalcitrant puzzles (see Textor, this volume). His taxonomy of conscious phenomena is bound to be rediscovered as the field of consciousness studies develops (see Kriegel forthcoming). In this paper, I have argued that there may be yet another deep and insightful lesson we have in store from Brentano, one that connects the very possibility of representation to self-

representation. Unlike most work on mental representation over the past century, Brentano's thinking on the subject is attune to the fact that mental representation involves crucially not only a representation-of relation, but also a representation-to relation. Attention to this relation of representation-to opens the door to a number of surprising and important theses about the nature of mental representation; not least the notion that there would be no representation without self-representation.

References

- Block, N. 1986. "Advertisement for a Semantics for Psychology." *Midwest Studies in Philosophy* 10: 615-677.
- Brandl, J.L. This volume. What is Pre-Reflective Self-Awareness? Brentano's Theory of Inner Consciousness Revisited."
- Brandom, R. 1994. "Reasoning and Representing." In *Philosophy of Mind*, ed. M. Michael and J. O'Leary-Hawthorne. Dordrecht: Kluwer Academic Publishers.
- Brentano, F. 1874. *Psychology from Empirical Standpoint*. Edited by O. Kraus. English edition L. L. McAlister. Translated by A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge and Kegan Paul, 1973.
- Burge, T. 1988. "Individualism and Self-Knowledge." *Journal of Philosophy* 85: 649-63.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge UP.
- Caston, V. 2002. "Aristotle on Consciousness." *Mind* 111: 751-815.
- Field, H. 1978. "Mental Representation." *Erkenntnis* 13: 9-61.
- Georgalis, N. 2005. "Representation and the First-Person Perspective." Forthcoming in *Synthese*.
- Harman, G. 1982. "Conceptual Role Semantics." *Notre Dame Journal of Formal Logic* 23: 242-256.
- Horgan, T. and U. Kriegel 2012. *Phenomenal Intentionality: New Essays*. Oxford and New York: Oxford UP.
- Hossack, K. 2002. "Self-Knowledge and Consciousness." *Proceedings of the Aristotelian Society* 102: 163-181.
- Kriegel, U. 2003. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy* 33: 103-132.
- Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford UP.
- Kriegel, U. Forthcoming. *The Varieties of Consciousness*. Oxford and New York: Oxford UP.
- Lehrer, K. 1989. *Thomas Reid*. London: Routledge.

- Loar, B. 1981. *Mind and Meaning*. London: Cambridge UP.
- McAlister, L. 1977. *The Philosophy of Brentano*. New York: Humanities Press.
- McGinn, C. 1988. "Consciousness and Content." In *Proceedings of the British Academy* 76: 219-239. Reprinted in N.J. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press, 1997.
- Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge MA: MIT Press.
- Millikan, R. G. 1989. "Biosemantics." *Journal of Philosophy* 86: 281-297.
- Searle, J. R. 1992. *The Rediscovery of Mind*. Cambridge MA: MIT Press.
- Smith, D. W. 1986. "The Structure of (Self-)Consciousness." *Topoi* 5: 149-156.
- Textor, M. This volume. "Unity Without Self: Brentano on the Unity of Consciousness."

¹ It is sometime claimed that Leibniz's *petites perceptions* were a pre-Freudian introduction of the unconscious, but upon examination this is not all that plausible. It is much more plausible that the *petites perceptions* were perceptual experiences in the periphery or "fringe" of consciousness.

² For a fuller discussion of Brentano on these matters, see Brandl (this volume).

³ The fitting label is due to Jenann Ismael.

⁴ If we accept the fourth condition on representation-by-courtesy, we may also formulate the thesis (or a nearby one), more economically, as follows the claim that every paradigmatic token of representation-of is also a representation-to.

⁵ One view, by no means implausible, is that tokens of a natural sign type which has no actual tokens that are representations-to are not a representations-of. On this view, representation-of requires not only nomologically possible representation-to, but also actual representation-to. Although I find this view plausible, I will not commit to here.

⁶ Moreover, to the extent that we are liberal in our usage of "representation," just as we can devise a thin sense of representation that would require very little, so we can devise a thick sense on which representation-of is impossible without representation-to *by stipulation*.

⁷ This consideration is inspired by a similar methodological principle due to Georgalis (2005).

⁸ Thus, if someone held that the minimalist account is at least a good analysis of first-person representation-to, we would protest that it makes no room for a commonality with third-person representation-to. There is, of course, also the problem discussed above, that the tree rings do not represent the tree's age to the tree even though the minimalist analysis' conditions are satisfied.

⁹ I use "and/or" here as short for formulating what are in fact two separate analyses, one requiring *z* to be conscious, the other requiring him or her to be sentient.

¹⁰ The same holds for tacit beliefs. Tacitus believed that 1374.67 is greater than 873.92. His belief (i) represented the fact that 1374.67 > 873.92 and (ii) occurred in Tacitus, and (iii) Tacitus was a conscious and sentient creature. Yet the belief did not represent the fact that 1374.67 > 873.92 *to* Tacitus in any non-

technical sense. At least this is so if we stipulate that, throughout his life, Tacitus never became aware of this tacit belief. Searle (1992) claims that tacit beliefs could not be anything but brute neurophysiological states. If so, Tacitus' tacit belief is no different from his goosebumps. Unconscious brain states and skin states are exactly the same in the respect.

¹¹ See Field 1977, Loar 1981, Harman 1982, Block 1986, Brandom 1994.

¹² This is the semantics favored by Carruthers (2000 Ch. 9); see Millikan (1984, 1989) for consumer semantics proper.

¹³ To my mind, the same problem afflicts functional role accounts of representation-of. Functionalists assume that a representation acquires its representational content in virtue of its functional role. But this quite obviously puts the cart before the horse. More plausibly, a representation acquires the functional role it does precisely because of the representational content it has.

¹⁴ I do not mean to suggest that the representation must represent itself *as* itself. (Perhaps it will turn out that it does, but we ought not commit to this until a compelling argument is offered.) Also, although I am using the propositional form here (with the 'that' clause), this is inessential and can be rid of readily.

¹⁵ This is in fact Caston's (2002) view of conscious representations: they are token representations that betoken two representational types in such a way that they become self-representing.

¹⁶ A view of this sort was also held by Thomas Reid (see Lehrer 1989, Hossack 2002), who held that all *sensations* are self-representing.

¹⁷ Nonetheless, the regress of types must be deemed at least more acceptable than the regress of tokens. Thus, some people may hold that every token necessarily betokens an infinity of types, since it betokens indefinitely many *disjunctive* type. Thus, if a creature betokens the type Cat, it also betokens the type Cat-or-dog. Others, however, may reject this notion, claiming either that there *are* no disjunctive types or that the tokening of a disjunctive type is not just a matter of tokening one of the disjunct types.