CrossMark

Explaining the reified notion of representation from a linguistic perspective

Farid Zahnoun[1]

Abstract

Despite the growing popularity of nonrepresentationalist approaches to cognition, and especially of those coming from the enactivist corner, positing internal representations is still the order of the day in mainstream cognitive science. Indeed, the idea that we have to invoke internal content-carrying, thing-like entities to account for the workings of mind and cognition proves to be particularly resilient. In this paper, my aim is to explain at least partially where this resilience of the reified notion of representation comes from. What I want to show is that, crucially, positing inner representations isn't so much warranted by the scientific practice itself – as is commonly held – but much more motivated by nonscientific and pre-theoretical elements that largely stem from, what I will call, linguistic contingencies. Otherwise put, much of what makes the reified notion of representation an attractive posit can be explained, not by the science, but by the way we, including cognitive scientists, speak. What I want to do here, then, is first, rehearse what reification means in the context of representationalism (and why it is problematic) and, second, specify which linguistic contingencies can (partially) account for why the idea of positing representations remains for many not only a viable option, but an indispensability for anyone interested in explanations of mind and cognition.

Keywords Enactivism · Anti-representationalism · Mainstream cognitive science · Reification · Linguistic contingency · Hypostatization · Mental states · Manifest vs. scientific image · Content/vehicle distinction · Semantic properties

1 Introduction

The days when representationalist cognitivism could lay claim on being the only game in town are now well behind us. Not only have we seen, in the past two decades, the

✉ Farid Zahnoun
   faridzahnoun@gmail.com

[1] Center for Philosophical Psychology, University of Antwerp (UA), D. 223, Grote Kauwenberg 18, 2000 Antwerp, Belgium

rise of various alternative approaches to mind and cognition (the so called E-approaches), recently, it has even been suggested that playing the mental representation game should never have been considered a viable tactic to begin with. In their recent book, enactivists Ezequiel Di Paolo, Xabier Barandiaran and Thomas Buhrmann put it as follows:

> One cannot claim that there is no alternative to explain cognition when representationalism itself is not an alternative in the first place. (Di Paolo et al. 2017: 29)

The idea of invoking internal representations for explanations of cognition has for many now indeed become part of a 'no-go zone'. It is felt that the notion of internal representation, traditional cognitive science's central explanatory posit, is simply untenable. For a growing group of researchers, the notion is considered to be problematic for reasons that simply can't be ignored, the most notorious being perhaps the problem of how to naturalistically account for these entities' mysterious property of carrying a certain content. Furthermore, and adding insult to injury, it has repeatedly been argued that, even if we somehow *would* be able to make ontological sense of there existing internal content-carrying entities in the head, we are still left with the epistemological question of how these entities are supposed to be doing any explanatory work.[1] How, exactly, are internal content-carrying entities supposed to be explanatorily relevant within a causal framework?

Despite growing skepticisms regarding representations, however, these are questions that remain within mainstream cognitive science not so much unanswered, as simply unasked. Here, internal representations are still considered to be indispensable theoretical entities. What explains representation's resilience? The main purpose of this paper is to try and shed some light on certain important, yet underexamined elements that might co-explain why the so-called 'representational pull' (Di Paolo et al. 2017: 11) remains for many so hard to escape. Interestingly, the abovementioned enactive authors already provide us with their own (partial) answer to this issue:

> [t]he pull of representationalism, the sustained pervasiveness of the mediational epistemology, may in part be attributable to the ubiquitous amount of [*images*] and symbols that populate our habitat. Behind each of these images and symbols lie acts of producing representations and interpreting them. In fact, representations are nothing but the reification of these acts. (Di Paolo et al. 2017: 27, my addition)

The above passage contains some valuable insights and I am deeply sympathetic with Di Paolo et al.'s emphasis on the element of reification here. I do feel, however, that the point of reification deserves more attention. For my purposes at least, it is important that we get as clear as possible on what reification means in the context of internal representation. The quoted passage above is in this respect too tentative, and in fact appears to reiterate the reified conception of representation: saying that we reify acts of producing representations still contains the idea that representations are the kind of

---

[1] See Hutto and Myin (2013), Chemero (2009) and van Gelder (1995) for a more elaborated account of why representations might be explanatorily superfluous.

things that can be produced, which is in tension with what these authors (rightly, I think) want to deny, namely that representations are a kind of thing. And just to be clear, it is this reified notion of representation that still holds sway over contemporary theorizing about mind and cognition. On, what Frances Egan calls, the 'Standard View',[2] a representation is understood as a kind of distinct internal item that somehow carries semantically evaluable content. Assuming that these internal entities must exist in order to account for cognitive phenomena is to be in the grip of the representational pull.

The reified understanding of representation is so ubiquitous (it is, after all, the Standard View) that all exemplification appears somewhat random. One of many instances can be found in the work of William Bechtel. In discussing Newell and Simon's influential physical symbol theory (which is itself rife with reification), Bechtel agrees with authors like John Haugeland and Tim van Gelder that, in general, representation is to be characterized in terms of an entity X standing in for some entity Y. The fact that this 'entity X' is to be understood as a thing (reification) becomes very clear when Bechtel continues:

> While there is a good deal of agreement about the importance of the standing-in aspect of representations, it is considerably more difficult to explicate what it is for one *thing* to stand in for another. Philosophers who have tried to explicate this notion have looked in two different directions: back to the object or event for which *the object* is to stand in and forward to the process which will use the representation in lieu of that for which it stands in. (Bechtel 1998: 298; my emphases).

Again, this is but one of countless Standard View examples in which representation is being reified.[3] So when I say that I want to contribute to a better understanding of the representational pull, I first of all mean to say that I want to clarify why this specific idea of internal representations as thing-like entities remains for many an attractive theoretical assumption. My point is that (at least part of) what makes the idea of there being internal content-carrying things in the head so readily acceptable, is for the bigger part not attributable to the scientific practice itself, but to a large extent the result of certain language related contingencies. Laying bare these contingencies is supposed to show that, in the end, reasons for assuming the existence of internal representations are not nearly as good as might initially be thought. Before I discuss these linguistic contingencies, however, I will first say a bit more about reification, and how it relates to representation.

2 Getting clear on reification

Reification, as the etymology of the term indicates, refers to the act of conceiving of something which is not a thing as a thing (the Latin 'res' means 'thing'). In German, we

---

[2] See Egan (2012).

[3] Another example would be Millikan's influential notion of 'pushmi-pullyu representations', where these entities are presented as the kind of 'things' that can be produced or consumed. See, for instance, Millikan (2005).

get 'Verdinglichung' ('Ding' meaning 'thing'), in French we get 'chosisme' ('chose' meaning 'thing'), and in English, we also find some authors using 'thingification'[4] as an alternative to reification. I am explicitly defining reification here because there appears to be some divergence of opinion as to what 'reification' is supposed to denote exactly. Michael Rescorla, for instance, uses the notion of reification as synonymous with making something countable, assuming that this is the original sense in which Quine used the notion.[5] I think Rescorla is mistaken, both in believing that reification equates to making something countable[6] and in thinking that this is also what Quine had in mind. Conceiving of something as 'some thing' involves much more than countability. Typically, 'things' are thought to be characterized, not only by countability, but also by independent spatiotemporal existence, by being in principle observable, by having physical properties (e.g. mass), by physical property-based classifiability, by localizability, by finiteness in both space and time, by causal efficaciousness and so on. So reification with regard to non-things such as actions, events, relations, norms, rules, qualities, and, as I'll argue, representations, refers to the practice of conceiving of these non-things as attributable with this list of properties. In any case, this is how I will understand reification.[7]

It is important to stress that the psychological phenomenon of thinking of 'non-things' as thing-like is not some marginal phenomenon, but a prevalent feature of human thought itself, and one which lies at the heart of some fundamental philosophical problems. For instance, when Heidegger criticizes traditional ontology for having overlooked the difference between 'Being' and 'beings', the target of his critique can be read in terms of the reification of Being. In discussing Heidegger's 'ontologische Differenz', Jan Slaby correctly observes that this reification of Being

> is not philosophy's fault alone, as it rather testifies to a pervasive condition of human existence in general. It is a structural self-misunderstanding build into all mundane self-relations— namely, *the tendency to understand oneself in terms of the modes of being of the things and the stuff that one routinely deals with in everyday comportment*, a kind of default self-objectification. (Slaby 2017, p. 7., my emphasis)

Slaby's thoughts become very pertinent when we bring them to bear on representationalism within contemporary cognitive science. The very fact that the idea that brains are in the business of representing needs to be cashed out in terms of the workings of

---

[4] See, for instance Barad (2003): 812.

[5] I am referring here to Rescorla's guest-lecture *Reifying Representations*, which was held September 12th 2017 at UCL.

[6] In addition, it is also unclear why we would need to reify entities so as to make them countable. We usually have no problem counting, for instance, events, without having to conceive of them as thing-like. We count the two World Wars, but nobody thinks of them as having thing-like properties. There seems to be no reason why this would be any different for neural events, as Rescorla seems to assume.

[7] In this sense, my use of the term reification is also different from Ludger van Dijk's, who uses the notion to denote the bigger psychological phenomenon of backwards causation. It would seem, however, that what van Dijk calls 'concretization' (which is one of three aspects of backwards causation) is precisely what I mean by reification. Terminological matters aside, van Dijk's analysis of the three-step phenomenon he calls reification deserves close attention as it lays bare uncritical tendencies in our explanations about behavior which require close attention themselves. See van Dijk (2016).

internal things (representations), and that, furthermore, these things are themselves being modeled on actual things which are ubiquitous in our sociocultural environment (ordinary public representations), bears testimony to that same 'pervasive condition of human existence' Slaby wants to reveal here. Trying to understand the mind in the same ways the mind understands its daily environment is not at all an attractive vantage point. More generally, Slaby's remarks point us to an important question, namely the question of whether it is at all plausible that we can understand ourselves in the same ways we understand the world. Put in cognitive science terms, is it at all plausible that we can understand human cognition by the same means with which human cognition 'understands' the world? We will have to leave these questions for another day.

3 Representation and reification

Now, I have just said that 'representation' should not be understood as referring to things. This might sound counter-intuitive, if not downright wrong, considering the fact that we prima facie do ordinarily use the term to pick out *things* like (especially written) linguistic symbols, pictures, traffic signs, road maps, scale models, statues, and so on. And indeed, it is undeniable that on the Standard View, virtually all technical notions are modelled on these more ordinary public items. Depending on the representationalist theory, internal representations are sometimes conceived of as symbol-like, picture-like, map-like, model-like, and so on.[8]

On the Standard View, then, the question is not *whether* internal representations are thing-like, but *what kind of thing* they are like. Their thing-like character is part of the definiendum, not the definiens. Further characterization of internal representation typically becomes a matter of relating them to other thing-like representations of which we already have an ordinary understanding. In fact, it is this connection to our more ordinary understanding of representation that precisely warrants the use of the term 'representation'. As Dennett remarked already forty years ago:

> Whatever *mental* representations are, they must be understood by analogy to *non*mental representations, such as words, sentences, maps, graphs, pictures, charts, statues, telegrams, etc. (Dennett 1978, p. 175)

Similarly, William Ramsey has more recently claimed that any theory that posits internal representations has to meet the challenge of showing how these entities are actually doing something "recognizably representational in nature" (Ramsey 2007, p. 28), meaning that the theoretical notion of a representation should be sufficiently similar to our pre-theoretical notion of what it means to represent something else. I agree with Dennett and Ramsey on this point. The problem, however, is that representation isn't only being reified in theories of

---

[8] For the idea that representations are distinct physical symbols, see for, instance Newell (1980) or Newell and Simon (1976). For the idea that these entities are akin to pictures, maps or models, see for instance, Kosslyn et al. (2006), Gładziejewski (2015), O'Brien and Opie (2015), Ramsey (2007), Waskan (2006), Braddon-Mitchell and Jackson (1996), McGinn (1989), Craik (1967), Godfrey-Smith (2014).

cognition, but that we already, and more originally, encounter this reification in our pre-theoretical understanding of representation as well. In other words, our pre-theoretical conception of representation is *already* characterized by the tendency to think of representations as thing-like, and this reification has been inherited by the more technical, theoretical notions serving as explanatory posits within cognitive science.[9] It is fairly obvious that the examples of representations given by Dennett seemingly qualify as 'thing-like'. So, thinking of 'representation' as essentially a name by which we denote discrete items in the public external sphere seems to warrant the idea that we can also use this term to denote private, internal items. Why, then, do I maintain that representation, whether internal or external, is not a thing, nor a discrete item? As I'll argue below, representation is a complex phenomenon that involves, but is not reducible to things.

4 "**What makes something a representation?**"?

A philosophical inquiry into the nature of representation typically proceeds along this line of reasoning: Certain physical structures (like words, maps, pictures, statues, neural states…) are representations because they represent. Like other functional notions like a pump or a chair, these things are defined by their function. But now a problem arises: for most functionally defined objects, we have a pretty clear idea of how to naturalistically account for what it is that makes these objects the kind of functional objects they are (e.g., what makes a heart a pump, or some wooden structure a chair). Accounting for the functional nature of the object can usually be done by referring to some of its properties. This, however, appears to be not so easy for those things serving the function of representing, understood as specifying how things stand with the world such that they could be different. So the philosopher of cognitive science is now confronted with a very specific question, which has recently been reformulated by Alex Morgan and Gualtiero Piccinini:

> anyone seriously interested in the conceptual foundations of cognitive science must eventually grapple with *what makes something a representation*. (Morgan and Piccinini 2017, p. 11, m.e.)

The problem, however, is that the question 'what makes something a representation?' is almost always taken as an invitation to further examine the properties of these things we identify as representations: what is it about certain things that gives them their special property of standing for something else? By asking such questions, the inquiry is setting off on the wrong foot as the reified notion of representation – which gave rise to the question to begin with

---

[9] I do say 'notions', for it would be a mistake to assume that there is only one notion of representation at work within cognitive science. But what all of these notions do have in common is the assumption that, if 'internal representations' exist, they are a kind of internal physical structure that somehow manages to say how things stand with the world. We could call this the generic reified notion of representation.

– is further reaffirmed and sustained: representations are assumed to be a kind of thing, and something about these things must account for their representational properties. Otherwise put: something about these things must make them about other things. Contrary to Morgan and Piccinini's suggestion, however, perhaps the question we "must eventually grapple with" is not *what* it is that makes something a representation. To understand the nature of representation, we really ought to consider the possibility that it might not be a 'what', but a 'who' that makes something a representation. Representation simply cannot be understood if we keep assuming it to be a kind of functional physical thing with thing-like properties. Representation is a complex phenomenon involving, indeed, physical things, but also individual cognitive capacities on the one hand, as well as certain socio-normative practices that regulate these capacities. So representation *involves* physical structures ('things'), but that doesn't mean that it can be understood or explained in terms of those physical structures. I would argue that representation necessarily involves a prescriptive element: a thing acquires the status of a representation when we *ought* to see it as standing for something else. To be sure, this 'ought' is a socio-normative ought. A representation is something *we*, within a given socio-normative community, are *supposed* to relate to as something else which is in some sense absent. Importantly, the objective status of something's being a representation, as well as what the thing represents (its content), follows from the prescriptive fact that we ought to see it as such. This is what consolidates the representational status of a physical object. Otherwise put, representation is socio-normatively constituted. Elsewhere, I have referred to this as an instance of an 'ought' determining an 'is'. But, in addition to this prescriptive element, representation also presupposes a specific individual cognitive capacity, that is, the capacity to relate to something as something else (which is something *persons*, not brains, do). In short, the fact that some physical thing can be properly identified as a representation relies on the individual cognitive capacity to see something as something else, as well as the socio-normative regulation of this capacity. Reification, then, occurs when these cognitive and socio-normative preconditions are being dispensed with altogether so that what remains of 'representation' is not much more than a count noun referring to some (internal or external) physical structure that is assumed to manage, in virtue of its properties, to specify how things stand with the world such that they could be different. This thing-like notion of representation is by many – including myself – now felt to be untenable, and we know of no clear example that might even begin to make intelligible the idea that some physical entity can carry a content in virtue of the way it is. Approaching things functionally doesn't help, for the idea that the verb 'to represent' literally refers to an object's activity (as in, 'neural structure x is engaged in the activity of representing') seems to be simply an instance of a metaphor being interpreted literally. 'No thing' literally performs the activity of representing so that, for instance, we could say that, just as hearts pump and livers filter blood, brains represent the world. Again, when we ordinarily use the expression 'x represents y', it is simply an error to think the x is literally doing something, something which can be described as representational activity. The only clear instance of

representational activity refers to an activity of persons, not brains, and certainly not external physical structures. But then this representational activity, as an instance of a cognitive activity, should not be taken as an explanans, but as itself in need of an explanation. Unfortunately, attributing personal level cognitive activity to subpersonal structures has become part and parcel of mainstream cognitive science' explanatory strategy.[10] How this is supposed to work exactly remains unclear.

We should wonder why it is, then, that the reified notion of representation perseveres. What I want to do is reveal at least part of the answer by showing how, within mainstream cognitive science, certain language-related elements are co-responsible for the still widespread reluctance of giving up on internal representations. More specifically, I want to show how the reified notion of representation is being kept alive by discursive, rather than scientific practices. In the following, I will discuss a number of pre-theoretical language-related elements that might partially explain why the idea of internal content-carrying entities, despite its critical shortcomings, is so hard to eradicate. In addition, I will also discuss one theoretical, non-language related motive for wanting to hold on to the reified notion of representation. As I will show, this motive is itself ill-conceived.

5 Reification through hypostatization

All languages are rife with idioms, appearing on both the semantic and the syntactic level. For instance, sentences in which nouns or substantives literally refer to physical objects or 'substances' are much less current than one might suspect. A sentence like 'The cat is on the mat', in which both substantives refer to physical things (cat, mat), may be the handbook's preferred example of an assertion, it is hardly representative of the idiomatic complexities of ordinary language. One such idiomatic phenomenon is that of 'hypostatization'. Hypostatization can be defined as the linguistic practice of placing non-substantive entities like actions or qualities in the grammatical category of nouns, so as when an activity-indicating verb like 'to walk' becomes a noun in the expression 'going for a walk'. I want to claim that the count noun 'representation' should also be viewed as an example of hypostatization. I'll first say a bit more about hypostatization, and its relation to reification.

Hypostatization is an instance of the much broader linguistic phenomenon of grammatical derivation, which isn't necessarily a derivation to a noun; verbs or qualities can be derived *from* nouns as well (e.g. 'chairing' or 'friendly'), and we also find derivations from verbs to adjectives (e.g. 'talkative') or from adjectives to verbs

---

[10] One recent example of this widespread tendency to attribute subpersonal-level structures with personal-level activities can be found in the popular Predictive Processing accounts. Jacob Hohwy, for instance, writes: "…the problem of perception is the problem of using the effects – that is, the sensory data, that is all the brain has access to – to figure out the causes. It is then a problem of causal inference for the brain analogous in many respects to our everyday reasoning about cause and effect, and to the scientific methods of causal inference (Hohwy 2013: 13). And, of course, Predictive Processing's core idea of brains making predictions already exemplifies such a projection from the personal to the subpersonal level

(e.g. 'to sweeten').[11] Hypostatization is a remarkably common form of derivation, and it can be found, at least in the English language, virtually everywhere.[12] Crucially, however, what makes hypostatization such a unique form of grammatical derivation is that, unlike other derivations, the derived 'noun' term does not literally refer to a substance. In contrast, when verbs or adjectives are derived from nouns (e.g. 'hammering' or 'manly'), the terms *do* refer to an action or a quality respectively. Apparently, it is only with derived noun-terms that we encounter the phenomenon of an entity being referred to with a grammatical category to which it does not literally belong. In a sense, then, it is only here that grammar might set us semantically on the wrong foot.[13] My point is that this has also happened, and continues to happen, with the notion of representation. But before I elaborate this point, one quick preliminary remark regarding the difference between hypostatization and reification. Indeed, hypostatization is sometimes defined as synonymous with reification. Here, however, I will consider both terms as referring to two different, though very closely related phenomena. I will take 'hypostatization' as referring to the just described linguistic phenomenon of categorical substitution, whereas I will use 'reification' as referring to the earlier explicated psychological phenomenon of conceptually treating non-things as things. Both phenomena are closely connected in the sense that one can give rise to the other, but they are not, strictly speaking, identical. We may hypostatize in language without therefore reifying in thought.

My point with regard to the notion of representation, then, is that the fact that mainstream cognitive science discourse has adopted, and continually reiterates the habit of referring to 'internal representations' in substance terms, has a psychological effect on the way these hypothetical entities are conceived of ontologically. It is only by reifying representation that the notion of an *internal* representation can start to make sense. And one particularly good strategy to keep the reified notion alive is by adopting the grammatical category of a noun to refer to these 'things'. And the scientific mind is susceptible to the psychological effects of language as well, even though some scientific minded readers will doubtlessly find that the above does not apply to them.[14] They are invited to interpret the above as a cautionary reminder not to fall into the

---

[11] The phenomenon of linguistic derivation is, of course, not an exclusive feature of the English language. So, for instance, in Spanish we find, as an instance of a verb derived from an adjective 'verdear' (meaning something like 'becoming green'), or in Russian, as an instance of a verb derived from a noun, we find the verb 'solit', which means 'to cover with salt' ('sol' meaning 'salt'). These examples are borrowed from Lachlan Mackenzie (Mackenzie, private correspondence, see also: Mackenzie 2004).

[12] In fact, the expression 'it can be found everywhere in the English language' already contains two hypostatizations: first, the 'it' as referring to 'hypostatization', for the noun 'hypostatization' is itself an example of hypostatization; and second, 'the English language', for languages aren't things, and they certainly aren't things in which other things can literally be found.

[13] We may wonder why this is so. According to Mackenzie, this is probably related to the fact that only nouns can denote independently, whereas verbs or adjectives require a broader semantic-syntactic context. (Mackenzie, private correspondence).

[14] An objection may be that representation is a functional notion, and that 'representation' first of all refers to the activity of, say, a neural state. It may be argued that it is because certain neural states are in the business of, or, to use a more common rhetoric, fulfill the role of representing, that we may correctly identify them as representations. Note, however, that this functional-role interpretation does not in any way escape the reified picture, and all the problems associated with it. It is still some physical object that is supposed to be fulfilling the role of representing and it still needs to be shown in a naturalistically credible way how such an object (presumably some internal neural structure) can literally be fulfilling the role of representing or standing for something else.

pitfalls of ordinary language. The take-home message here is that adopting a reified notion of representations comes with the responsibility of being able to give language independent reasons for this reified conception, and not simply assume it based on the already installed practice of grammatically referring to these entities by a noun.

6 Reification through ambiguity: 'state' as an ambiguous term

Next to the reified notion of representation itself, within mainstream cognitive science, there is another closely associated notion at work which is ambiguous, and manages on the basis of this ambiguity to further endorse a reified conception of representation. The notion I'm thinking of here is that of a *state*. Cognitive science literature, as well as traditional philosophy of mind, is rife with talk of mental states, brain/neural states, intentional states, representational states, and so on. What is typically overlooked here, however, is that the meaning of 'state' changes considerably depending on the combinations it is used in. In this regard, it is worth bringing up the oft-quoted § 308 of Wittgenstein's *Philosophical Investigations*:

> How does the philosophical problem about mental processes and states and about behaviourism arise? – The first step is the one that altogether escapes notice. We talk of processes and states and leave their nature undecided. Sometime perhaps we shall know more about them – we think. But that is just what commits us to a particular way of looking at the matter…. (The decisive movement in the conjuring trick has been made, and it was the very one that we thought quite innocent.)

The quote's particular relevancy for our discussion should be clear, as Wittgenstein is explicitly referring to states here. More than half a century later, cognitive science has still left the nature of states 'undecided', allowing theorists to use 'mental states' and 'brain states' interchangeably. But what, on closer inspection, allows us to treat these various uses of 'state' as equivalent, or even similar? After all, the way 'states' are ascribed to brains is entirely different from the way 'states' are ascribed to a person. When we are ordinarily speaking of a mental state, we are using the term 'state' in the meaning of 'condition'. 'Mental states' first of all pick out the different psychological conditions in which a person (or an animal) can be said to be. When we ordinarily inform about mental states, we first of all want to know how *someone* is doing. What we are *not* interested in is a certain physical *configuration*, which is a whole different sense of the term 'state'. In a second, perhaps less ordinary sense, a person's mental state might also be used to refer to, not *how* the person is doing, but *what* he is doing in his mind. For instance, when we ascribe assertoric thoughts to a person, we could say that the person is in a certain mental state, namely that of thinking. And to the extent that thinking is a representational activity with intentionality, we might say that he is in an intentional mental state. But notice how, also in common speech, the *is* has changed into *has*. Rather than saying that a person *is* in an intentional state, we tend to say that he *has* intentional states. Saying that a person is thinking or saying that he has thoughts is in ordinary speech perfectly interchangeable. However, the move from *being* in an intentional state to *having* an intentional state is a step towards reification, and one that

is unmotivated by science. The ambiguity of 'state' further completes the reifying process. With some conceptual leniency 'brain states' can indeed be said to be things, or at least configurations of things (neurons). But now, the conjuring trick is easily accomplished. Because brain states are physically localizable things with physical properties, mental states in general, and intentional/representational states in particular, must also be like that, hereby conveniently overlooking the ambiguity of 'state'. Furthermore, since they are now conceived of as 'thing-like', they can also be *had*. *Being* in a representational state (thinking) has now become synonymous with *having* a representation (a thought). And having a representation *just is* having a brain state.

7 Reification through the notion of property

As we've seen, cognitive science's generic notion of internal representations as truth evaluable entities (entities that somehow specify how things stand with the world such that they could be different) is modeled on our pre-theoretical familiarity with such entities. But this means that internal representations are first of all modeled on assertions, claims, judgments, and other linguistic items with truth evaluable content, for these are the *clearest* instances of entities that can unproblematically be said to have truth conditions. The question then becomes: does it make sense to think there can be assertions in the brain? Now, perhaps one may object to this formulation and say that, literally speaking, it would be absurd to expect to find assertions in the brain. He or she might add that perhaps we might find something sufficiently analogous to assertions, so that we can still say that it shares the relevant property of semantic evaluability with actual assertions, without it literally being an assertion. This line of reasoning, however, hinges on the assumption that truth evaluability is, first, a kind of *property*, and, second, a property that can be ascribed, not only to certain linguistic entities like assertions, but to extra-linguistic entities as well (maps, models, or neural states, for instance). As I want to show, the assumption that we can invoke property talk in relation to semantic issues is problematic, and also tends to promote reification.

Within analytic philosophy, the use of the notion of property abounds. Likewise, talk of *semantic* or *representational* properties has become an almost routinous affair. Yet, as anyone with a background in analytic philosophy knows, the philosophical notion of property is highly debated. In its most ordinary, more narrow sense, 'property' is used to refer to an object's features or characteristics. Roughly, they are the 'things' we would be mentioning if we were to describe the object. However, in its wider sense, 'property' is understood as synonymous with truthful predication. Anything that can be truthfully predicated of an entity can be said to be a property of that entity. So when it is asserted that an assertion like 'the cat is on the mat' has semantic properties, what is usually meant is that its semantic evaluability – i.e. the entity's having truth conditions – can be truthfully predicated. Of course, no one believes that the sentence 'the cat is on the mat' has the property of being truth evaluable merely in virtue of that sentence's physical properties. But note, however, that regardless of whether one adopts the narrow notion of property or the wide notion preferred by analytic philosophy, what remains in place is the idea that 'things' like assertions can be incorporated in the same metaphysical picture we use to classify the whole of nature, namely in terms of objects (in the broad sense of the term) and their properties. On this construal, semantic

properties may be different from ordinary, material properties, but the ontological foundation on which it is construed remains untouched, namely the object/property distinction itself, and its logical reflection in the subject/predicate distinction. By simply invoking the notion of property in relation to semantic issues, one is framing these issues in a metaphysical picture in which they might not belong. As already discussed, using nouns to refer to abstract entities like assertions is one thing (hypostatization), but to conceive of these 'things' as *being* physical objects with properties is another (reification).[15] And although we endow non-thinglike objects like actions, events and qualities with properties as well, when it comes to internal representations, the object-property structure we encounter here is nothing but the ordinary 'physical-thing-with-physical-properties' structure. It is this conception of representations as physical things with properties we keep encountering in cognitive science literature, not only in the more traditional computationalist approaches (e.g. Newell and Simon's *physical symbols* system), but in newer approaches as well. Take the signal system approach to the brain, for example. Here, the semantic notions on duty are codes, information and "signals running around a very complicated signaling network" (Skyrms 2010). But no matter what nouns are being used to denote the supposedly semantic-content-carrying entities, the same familiar picture keeps recurring, namely that of a physical thing with both physical and semantic *properties*. Consider how Rosa Cao, for instance, discusses the role of the receiver in a signal system:

> To interact with signals in the right way (so as to be a receiver of semantic information) is, roughly, for the receiver to have some degree of flexibility in its response to a signal. At first pass, this means that receivers will be aptly described as acting (at least in part, but perhaps primarily) on the basis of the semantic properties of a signal, *in addition to its material ones.* (Cao 2012: 53, my italics)

Conceiving of semantic properties as being on a par with material properties is one more expression of reification. It results from the mistaken assumption that truth evaluable representations can be conceived of as physical things with intrinsic properties, both material and semantic ones. But it is a mistake to think that semantic properties are somehow in line with material properties. In his paper *Naturalizing Representational Content*, Nick Shea provides us already in the opening lines with a clear instance of this error. He writes:

> Some things in the world have semantic properties. Spoken and written sentences are paradigm cases. They are perfectly ordinary particulars in the causal order: ink marks on the page and vibrations in the air. But they also have more exotic properties: they can be true or false, or, in the case of imperatives, they can be satisfied, or go unsatisfied. (Shea 2013: 496)

---

[15] Following Wittgenstein, McGinn (1989) makes a very similar point when he warns us against "… assimilating intentional properties to the properties characteristic of substances." (McGinn 1989: 29 ft. 40). And further he reminds us: "Wittgenstein, of course, had the idea that there is a persistent and rooted tendency to model the mind on the world of material objects. It is not as if we come to see that this is false and there's an end to the matter. Prolonged therapy may therefore be needed to dislodge wrong philosophical conceptions. We might have an internal fight on our hands." (McGinn 1989: 30 ft. 41)

Spoken and written sentences are *not* perfectly ordinary particulars in the causal order. They are, as I've argued above, socio-normatively constituted. A truth-evaluable assertion is not a material object which has, in addition to its material properties, semantic properties as well. An assertion *just is* truth-evaluability. And the fact that we can truthfully predicate of some physical thing that it is truth evaluable is not the result of the things properties (material or semantic). In other words, it is not because a thing has semantic properties that it becomes truth evaluable. It is truth-evaluable because we make it so. But nothing is gained by invoking talk of properties here. All we need to say is that it is true that certain physical things specify how things are such that they could be different. Explaining how this could be the case need not rely on property talk, but only on talk about cognitive capacities that allow us to relate in different ways to an object, and about certain socio-normative dynamics (on different timescales) that constrain and regulate this capacity.

All of this does not mean that talk of semantic properties is always and everywhere problematic. When an analytic philosopher like Davidson, for instance, uses the vocabulary of semantic properties, he does this exclusively in relation to *abstract* linguistic entities like sentences. The wide sense of property (true predication) allows for this: since it can be truthfully predicated or asserted of an assertoric sentence that it has truth conditions, it is by definition correct to say that the sentence has this property. But one can't simply transfer this language-philosophical approach to physical things, whether they are 'ink marks on the page', 'vibrations in the air', hand gestures or neural structures. Again, semantic properties are not on a par with physical properties. Conflating these very different properties is one more factor which facilitates and reinforces a reified understanding of internal representation.

8 Reification through the vehicle/content metaphor

One of the most reiterated, yet at the same time most underexamined idioms in cognitive science literature is the commonplace distinction between a vehicle and its content. On a standard interpretation, internal representations are conceived of as content carrying vehicles and most theorists simply adopt the distinction without explicitly thematizing it (for an exception to this, see for instance Hurley 1998). There is discussion about when exactly the distinction got introduced into philosophy of cognitive science (Dennett 1968 appears to be one of the earliest sources[16]), but it is important to note that the conceptual distinction itself already existed long before people like Dennett and Millikan[17] started to use it in discussions about consciousness or the naturalization of semantics. The idea of thinking about content as being carried by physical vehicles is not born within philosophy of cognitive science, and it certainly isn't motivated by empirical research (despite what some theorists seem to believe[18]).

---

[16] In *Content and Consciousness*, Dennett writes: "The crucial point that emerges from this is that the candidates for *vehicles of content* or significance in the brain are compound." (Dennett 1968: 56; my italics)

[17] See, for instance, Millikan 1993.

[18] See for instance Manzotti and Pepperell 2013. The authors criticize an anonymous referee, referred to as *A*, who apparently takes the content/vehicle distinction to be empirically established: "For *A*, the nature of mental vehicles, and thus their separation from mental content, is an empirical matter rather than a terminological one. Yet (…) we wonder where the empirical evidence is for this distinction existing anywhere other than the minds of those who believe in it?" (Manzotti and Pepperell 2013: 368)

We have good reasons to assume that the distinction long predates cognitive science, and that it has emerged as a result of certain specific historical developments, namely the practice of writing, printing and distributing books. To my knowledge, the rather obvious link between the vehicle/content distinction and these historical events has never been made explicit within cognitive science literature. Doing so, however, quickly reveals one more contingency supporting the reified picture of representation.

Before anything else, it should be again emphasized that the vehicle/content distinction is first of all a *metaphor*. Or, rather, it is a *double* metaphor in that both the term 'vehicle', as well as 'content' are used in a non-literal sense. For our discussion of reification, however, it is especially the metaphorical notion of content that is of interest here, and not so much that of a vehicle (vehicles are about as thing-like as it gets). In its literal sense, 'content' is that which is being contained by a *volume* of some kind. It is, then, not difficult to see how the literal idea of content came to be associated with, and ultimately metaphorical for, meaning, considered in relation to writing, and the practice of writing books in particular. In fact, the earliest reference we find to content in relation to meaning connects the word to books. In the Oxford English Dictionary, the first mention of content in this sense dates back to 1481:

> Here endeth the table of the content and chapytres nombred of this present book.[19]

It is also no coincidence, then, that we refer to books as *volumes*. And just as amphorae, cups, boxes and other volumes can contain fluids or other materials, so too can books contain 'material', i.e. meaning. Furthermore, if we want to know what the book is about, we typically do so by asking what's *in* it. This image is not restricted to the English language. We find that this metaphor of meaning as content not only has roots in Latin (*continere*), but in other Indo-European languages as well (German: 'Inhalt', Dutch: 'Inhoud', Danish: 'indholdet'…). And of course, since books – like other volumes – are the kind of things that can be carried around, the idea of it serving as a vehicle for its content follows quite naturally. The problem, however, is not that this picture reifies the 'carriers' of meaning, but meaning itself. It is, after all, hard to see how something can be *in* a volume and be carried around by a vehicle without it being a physical 'thing' itself. In other words, the idea of thinking about meaning as something that can be contained by a volume and transported by a vehicle supports a reified understanding of meaning. This does not have to be problematic in itself, as long as we keep in mind that we are dealing with a metaphor here, and that the question as to how the metaphor should be cashed out in naturalistically respectable terms still stands. Nevertheless, within contemporary cognitive science, and especially within standard informational-computational approaches, the idea that the brain performs computational operations on content carrying vehicles (representations) is, as a rule, taken quite literally.

---

[19] Caxton tr. *Siege & Conqueste Jerusalem*, 1481, edition of 1893.

9 Reification and scientific explanation

The focus of the above discussion has been on language-related elements which plausibly sustain and reinforce the reified concept of representation, yet which are themselves not grounded in scientific practice. It should be noted, however, that one important contribution to the reification of representation does indeed come from *within* scientific practice, and not just from linguistic contingencies. As already indicated at the beginning of this paper, one major reason for supporting the reified picture of representation is that it renders the notion deployable within the larger framework of causal explanation. It is our understanding of what it means to be a scientific explanation that itself motivates the idea of internal representations as thing-like entities with, therefore, thing-like causal properties. Viewed in light of the demands of a scientific explanation, it isn't hard to see what makes the reified construal an attractive hypothetical posit. It is an attempt of tying in two threads which are deemed indispensable for the explanation of psychological phenomena. On the one hand, cognitive scientists want to hold on to the idea that the assumed contents of our mental states are explanatorily essential. On the other hand, it has become a central assumption of scientific theorizing that, for something to be a scientific explanation, the explanation must be – in the final analysis – causal. The reified notion of internal representation, which, as we've seen, is inherently dual in nature, wants to accommodate these demands by *stipulatively* uniting and incorporating both elements. On closer examination, however, we come to see how incoherent such an idea really is. When we think about what it means for an account to qualify as an objective causal explanation, we see that one of the conditions is that we precisely *abstract away from*, or *leave out content*. Causal explanations are objective in the sense that they explain how and why things are the way they are *regardless* of how subjects think or feel about it. From the causal scientific perspective, content simply *cannot* play an explanatory role, because content is precisely what is abstracted away from when we want to explain a phenomenon causally. Conversely, explaining why someone acted the way she did in terms of content (e.g. "Because she believed it was the right thing to do.") requires us precisely to *not* give an account in terms of cause and effect. Of course, both explanations do not have to be incompatible, but you can't have them *both as one explanation*, even if it should turn out that one explanation (presumably the content-invoking one) is reducible to the other (presumably the causal one). The reified notion of representation, however, seems to think it can do just that by stipulatively fusing them into one hypothetical entity: an objective symbol with causal power. In a sense, then, internal representations are the reified and intracranial version of what Sellars once described as the fusion of two different perspectives into one "stereoscopic view" (Sellars 1963: 5). I'll explain this – admittedly suggestive – claim in a bit more detail.

In his *Philosophy and the Scientific Image of Man*, Sellars famously argues that the contemporary philosopher is confronted with two very different, yet "equally public, equally non-arbitrary conceptions of man-in-the-world" (Sellars 1963: 5). As is well-known, he refers to these two conceptions as the *Manifest Image* and the *Scientific Image* respectively, and he takes it to be one of the great challenges of philosophy to understand how both images hang together. For my purposes, I'll define the Manifest Image as that perspective from which the world appears as what it is *for us*, or more precisely, *for a community of persons*. It is our everyday-perspective in which a

collection of H_2O molecules first appears as drink water, or black ink lines on a piece of paper as saying that someone's going to be back in 5 min. The Scientific Image can be defined in contrast to the Manifest Image in that, from the scientific perspective, the world is approached, not as what it is for us, but as how it is in itself, that is, how it is *objectively*. Within the Scientific Image, drinkable water appears first of all as a collection of H_2O molecules, or a note saying that someone will be back in 5 min as black ink lines on a piece of paper. I'm oversimplifying matters here a bit, but this doesn't affect the point I'm trying to make. For it seems rather clear that those things we refer to as representations, that is, objects which are said to be carrying content or meaning, are firmly confined within the space of the Manifest Image, and have – quite literally – no meaning within the Scientific Image. Objects we call representations (linguistic items, traffic signs, maps, models…) but also all other normatively constituted things like coins, chess pieces or baseballs (to use John Haugeland's favorite example), simply lose their meaning, and therefore their identity, when viewed through the objective, disengaged lens of the scientist interested in providing causal explanations. To be more precise, an essential property of *any* representation (whether it is public or intracranial) is that it has content. But content is precisely what we lose when we look at these entities from the perspective of the Scientific Image. In short, then, the idea of a content carrying representation makes no sense outside the Manifest Image. This does not imply that it would be impossible to *explain* the existence of contentful representations in objective-causal terms, but it does seem conceptually confused to *use* contentful representations within objective-causal explanations. Yet this is precisely what the notion of internal representation is supposed to be doing in causal-mechanistic explanations of mind and cognition, for instance in computationalist explanations. In other words, what is seen by philosophers like Sellars[20] as perhaps the greatest philosophical challenge (fusing the Manifest and the Scientific Image together in one stereoscopic view), is by mainstream cognitive science already considered a *fait accompli* by postulating internal content carrying entities which are at once causally explanatory. The reified notion of representation wants to provide a 'best of both worlds' explanation by combining essential elements from both explanations within the Manifest Image, as well as the Scientific Image, i.e. content and physical causation, respectively. This requires the postulation of a physical entity (to accommodate the causal efficaciousness condition) that somehow carries content in a non socio-normatively constituted sense. But this fusion can only be a con-fusion. In light of the above, it should be clear that the idea of such *things* (reification) existing in the head or elsewhere appears to be incoherent. Ultimately, the incoherent nature of the idea is the result of confusing, or rather, conflating a prescriptive socio-normatively constituted entity – within the Manifest Image – with an objective-descriptive one – within the Scientific Image. Indeed, this is the same as saying that the reified notion of internal representation is the result of a category mistake.

---

[20] I could just as well have referred to Peter Strawson, who also highlights the distinction between two possible, yet different perspectives on the world. What Sellars captures in terms of Manifest vs. Scientific Image is very similar to what is by Strawson referred to as the possible "occupying" of two "alternative standpoints". See Strawson (1985): 55.

10 Conclusion

I have tried to show how certain discursive practices within mainstream cognitive science keep reinforcing the reified notion of representation. Crucially, these 'ways of talking' are not themselves grounded in scientific practices. They come from elsewhere. One important source can be located in our pre-scientific and pre-philosophical use of the term 'representation', which has already reified the notion. In case of our ordinary usage of the term, matters are rather complicated with 'representation' because the reification is somewhat cloaked by the fact that the term does quite unproblematically appear to pick our ordinary things (such as traffic signs, maps, models, statues and so on). On closer inspection, however, representation should be understood as referring to a complex cognitive and socio-normative phenomenon which *involves*, besides certain personal cognitive capacities and socio-normative practices, physical things. But it is a mistake to therefore think that representation is itself a physical thing. However, this appears to be still a widespread assumption of mainstream cognitive science. In the final analysis, however, the idea that some physical thing can be said to be carrying a content – that specifies how things are such that they could be different – outside a socio-normative context (for instance in a neuroscientific context) turns out to be incoherent. In Sellarsian terms, it is nothing but a futile attempt to smuggle a notion which is firmly rooted within the Manifest Image into the Scientific Image. This move is unwarranted, and appears to be mainly facilitated by the language of representation itself. Rather than assuming representations "even before the theory starts" (Tonneau 2011/2012: 338), cognitive scientists should adopt a critical attitude towards their own discursive practices and see how much of their representation-talk is actually motivated by the scientific practice itself, rather than by mere linguistic contingencies.

References

Barad, K. (2003). Posthumanist performativity: toward an understanding of how matter comes to matter. *Signs 28,* 801–831. Gender and Science: The University of Chicago Press. New Issues.

Bechtel, W. (1998). Representations and cognitive explanations: assessing the dynamicist's challenge in cognitive science. *Cognitive Science, 22*, 295–318.

Braddon-Mitchell, D., & Jackson, F. (1996). *Philosophy of mind and cognition*. Oxford: Blackwell.

Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology and Philosophy, 27*, 49–71.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge: MIT Press.

Craik, K. J. W. (1943/1967). *The nature of explanation*. Cambridge: Cambridge University Press.

Dennett, D. C. (1968). *Content and consciousness*. London: Routledge and Kegan Paul plc.

Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Cambridge: MIT Press.

Di Paolo, E., Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor Life: An enactive proposal*. Oxford: Oxford University Press.

Egan, F. (2012). Representationalism. In E. Margolis, R. Samuels, & S. Stich (Eds.), *The Oxford handbook of philosophy and cognitive science* (pp. 250–272). Oxford: Oxford University Press.

Gładziejewski, P. (2015). Explaining cognitive phenomena with internal representations: A mechanistic perspective. *Studies in Logic, Grammar and Rhetoric, 40*, 63–90.

Godfrey-Smith, P. (2014). Signs and symbolic behavior. *Biological Theory, 9*, 78–88.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hurley, S. (1998). Vehicles, contents, conceptual structure, and externalism. *Analysis, 58*, 1–6.

Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge: MIT Press.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford: Oxford University Press.

Mackenzie, J. L. (2004). Semantic categories and operations in morphology I: Entity concepts. In G. Booij et al. (Eds.), *Morphology. An international handbook on inflection and word-formation* (Vol. 2, pp. 973–983). Berlin: de Gruyter.

Manzotti, R., & Pepperell, R. (2013). Denying the content–vehicle distinction: a response to 'the new mind revisited'. *AI and Society, 28*, 467–470.

McGinn, C. (1989). *Mental content*. Oxford: Basil Blackwell.

Millikan, R. G. (1993). Content and vehicle. In N. Eilan, R. McCarthy, & B. Brewer (Eds.), *Spatial representation* (pp. 256–268). Oxford: Blackwell.

Millikan, R. G. (2005). *Language: A Biological Model*. Oxford: Oxford University Press.

Morgan, A., & Piccinini, G. (2017). Towards a cognitive neuroscience of intentionality. *Minds and Machines, 28*, 119–139. https://doi.org/10.1007/s11023-017-9437-2.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4*, 135–183.

Newell, A., & Simon, H. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the Association for Computing Machinery, 19*, 113–126.

O'Brien, G., & Opie, J. (2015). Intentionality lite or analog content? *Philosophia, 43*, 723–730.

Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.

Sellars, W. (1963). Philosophy and the scientific image of man. In *Empiricism and the philosophy of mind*, 1–40. London: Routledge & Kegan Paul Ltd.

Shea, N. (2013). Naturalising representational content. *Philosophical Compass, 8*, 496–509.

Skyrms, B. (2010). *Signals: Evolution, learning and information*. Oxford: Oxford University Press.

Slaby, J. (2017). More than a feeling: Affect as radical Situatedness. *Midwest Studies in Philosophy, 41*, 7–26.

Strawson, P. F. (1985). *Skepticism and naturalism: Some varieties*. London: Methuen & Co. Ltd..

Tonneau, F. J. (2011/2012). Metaphor and truth : A review of representation reconsidered by W. M. Ramsey. *Behavior and Philosophy (Online), 39*(40), 331–343.

van Dijk, L. (2016). Laying down a path in talking. *Philosophical Psychology, 29*, 993–1003.

van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy, 92*, 345–381.

Waskan, J. (2006). *Models and cognition*. Cambridge: MIT Press.