

# The realizers and vehicles of mental representation

## *Abstract*

The neural vehicles of mental representation play an explanatory role in cognitive psychology that their realizers do not. Cognitive psychology individuates neural structures as representational vehicles in terms of the specific causal properties to which cognitive mechanisms are sensitive. Explanations that appeal to properties of vehicles can capture generalisations which are not available at the level of their neural realizers. In this paper, I argue that the individuation of realizers as vehicles restricts the sorts of explanations in which they can participate. I illustrate this with reference to Rupert's (2011) claim that representational vehicles can play an explanatory role in psychology in virtue of their quantity or proportion. I propose that such quantity-based explanatory claims can apply only to realizers and not to vehicles, in virtue of the particular causal role that vehicles play in psychological explanations.

*Keywords:* cognition, content, mental representation, psychological explanation, realizers, vehicle

## *1. Mental representations: realizers and vehicles*

In cognitive psychology, mental representations are understood as concrete particulars with both semantic and non-semantic properties. In virtue of their semantic properties, representations can be interpreted as bearers of content, and thus account for the intentionality of mental states. In virtue of their non-semantic properties, representations can play a causal role within a mechanistic system, which is usually assumed to be computational. The *realizers* of representations are whatever physical structures satisfy the appropriate roles; in human beings, these physical structures are likely

to be neural configurations.<sup>1</sup> This framework of roles and realizers is familiar from functionalist approaches to philosophy of mind.

Traditionally, representational mental states are individuated by their intentional contents: a thought that refers to cats and a thought that refers to dogs are distinct thoughts in virtue of having different referential contents. But it is a well-known fact that two mental states with the same referential content don't always play the same explanatory role in intentional explanation. Lois Lane's SUPERMAN concept does not play the same role as her CLARK KENT concept in explaining her behaviour, for example, despite the fact that the concepts co-refer. And as Twin-Earth cases demonstrate, two mental states can sometimes play the same explanatory role despite having distinct references: it is arguable that Oscar's WATER concept plays the same explanatory role in his behaviour as Twin-Oscar's WATER concept plays in his behaviour, despite the fact that Oscar's concept refers to H<sub>2</sub>O while Twin-Oscar's concept refers to XYZ. Any approach to psychological explanation needs to account for the *cognitive significance* of a representation's referential content: the way in which the referent is represented. In philosophy, this role is often associated with Fregean sense. Fregean senses, however, are not the sorts of naturalistic entities that cognitive psychology can endorse.

Naturalistic approaches to mind attempt to account for semantic properties (e.g. meaning, reference) in terms of non-semantic (e.g. causal, functional, or teleological) relations between representations and the aspects of the world they represent. Naturalistic theories of content acquisition can give an account of the referential content of representations, but naturalistic psychological explanation still needs to explain why co-referential representations can play distinct causal roles with the cognitive system. This requires that we have a way to individuate representations non-semantically, to give a finer-grained classification of representations than by

---

<sup>1</sup> The physical structure may itself be functionally characterized, and realized by a physical structure at a lower level of abstraction. For further details of the role/realizer relation and its application to representational theories of mind, see Bermudez, *Philosophy of Psychology* (2005).

their referential contents, but in a way which doesn't appeal to non-naturalistic semantic notion like Fregean sense. This is where representational *vehicles* come in.

In cognitive psychology, representational vehicles are the bearers of content. In addition to their semantic properties, they also have non-semantic properties to which the cognitive mechanisms are sensitive. The nature of these mechanisms will determine which non-semantic properties are the relevant ones: they might be formal properties, functional properties, or physical properties.<sup>2</sup> As Shea (2007) emphasises, the non-semantic properties which individuate representational vehicles are those which are recognisable to the cognitive mechanisms in question:

"The entities that vindicate the cognitive revolution's first and defining commitment – to an internal mechanism involving vehicles of content which are individuable non-semantic – deserve to be called representations. Typing representations as vehicles of content groups together different internal entities into classes that are importantly alike for internal processing [...]: they are different realisations of the same vehicle of content." (Shea 2007, 247-248)

Individuating representations by the relevant non-semantic properties enables us to say whether two neural structures realize the same representational vehicles or distinct representational vehicles. This notion of representational vehicles allows us to explain why the same referential content can play distinct explanatory roles, but without appealing to semantic concepts like Fregean sense. Instead, cognitive psychology appeals to the fact that the same referential content can be carried by distinct vehicles of representation, where those vehicles play distinct causal roles within the mechanism. Conversely, the fact that distinct referential contents can be carried by the same vehicle of representation, where the bearers of the two contents play the same causal role within the mechanism, explains how those contents can play the same explanatory role. Examples of

---

<sup>2</sup> For discussion of these approaches to vehicle individuation, see Stich (1983) and Aydede (1999).

representational vehicles include symbols in classically computational architectures, clusters in state space in connectionist networks, and attractor basins in dynamical systems.<sup>3</sup>

Notice that representational vehicles play a casual explanatory role in cognitive psychology that their realizers do not. Cognitive psychology appeals to generalizations over neural structures in terms of the way they participate in cognitive processes: the features in virtue of which those neural structures play the same or distinct causal roles within the mechanisms responsible for our cognitive capacities. Focusing on the vehicles of representation allows us to capture these generalizations in a way that we can't capture if we focus solely on their lower-level neural properties. The distinction between neurological and psychological explanations can be grounded by the distinct explanatory roles of the realizers of representations and the vehicles of representations.<sup>4</sup>

In work by Robert Rupert (2011), this distinction between vehicles and realizers of representation plays an important role.<sup>5</sup> Rupert proposes a view of the mind as “massively representational”: he thinks not only that the brain supports more mental representations than has previously been acknowledged, but also that the quantity or proportion of mental representations plays an explanatory role in the sciences of the mind. Rupert explicitly claims that he is thinking of these mental representations as distinct vehicles, rather than merely as distinct realizers. I'll outline Rupert's claim in the next section of this paper, before going on to explore his justification for these claims. I will argue that Rupert's commitment to the explanatory role of distinct vehicles leaves him unable to establish that the mind is massively representational in the sense he suggests. Furthermore, I will suggest that the only way for Rupert to establish that quantity is playing a key explanatory role is for him to accept that the states in question are merely distinct realizers.

---

<sup>3</sup> See Shea (2007) for discussion of vehicles in different mechanisms, with a focus on connectionist networks.

<sup>4</sup> The distinction between neurology and psychology is not actually this clear cut. But this rough approximation is assumed by Rupert (2011), and will be acceptable for my purposes in this paper.

<sup>5</sup> All further references to Rupert's work are to the 2011 paper unless other specified.

## 2. Rupert and the 'massively representational' mind

Rupert proposes that the mind is “massively representational”. First, he proposes that there are far more representations than previously acknowledged: that “[t]he mind-brain contains an enormous number of basic mental representations” (102). Second, he argues that the quantity or proportion of representations can play a role in psychological explanation: that “variation in the sheer number of vehicles [...] plays a causal-explanatory role in the production of certain forms of behavior” (111).

Rupert makes clear that his proposal concerns the number of representational *vehicles*, rather than merely the number of *realizers*. Rupert explicitly rejects the “more modest claim that, for any given mental representation, the subject is likely to have many realizers of it or merely has many psychologically equivalent vehicles” (111); and he rejects any interpretation of his view on which the cognitive states that he is treating as distinct vehicles of representation “are, instead, various realizers of the same mental representation” (110).

Furthermore, Rupert assumes that in making a claim about vehicles rather than realizers of representation, he is offering a psychological explanation rather than a neurological explanation:

“Partly because the number of active vehicles does explanatory work in cognitive science, I take being a particular vehicle to be a psychological-level construct—unlike mere realizers, which appear only at some lower level than, or as part of an explanatory enterprise distinct from, psychology.” (111)

Rupert’s key claim, that the number of representational vehicles involved in a cognitive process is a causal-explanatory factor in psychology, can thus be understood as the conjunction of two claims: one claim about the *quantity* of cognitive states involved and a second claim about the *vehicular* status of these cognitive states:

**Quantity:** Some psychological phenomena are explained by the number or proportion of cognitive states contributing to their production.

**Vehicle:** The cognitive states that explain the psychological phenomena in question are distinct vehicles of representation; they are not merely distinct realizers of the same vehicle.

I will argue that Rupert cannot hold both the Quantity and Vehicle claims, and thus cannot establish his claim that the mind is massively representational. In order to establish that the cognitive states are distinct *vehicles* of representation, the states in question need to be individuated by the non-semantic properties to which the cognitive mechanisms are sensitive, in virtue of which distinct vehicles play distinct causal roles. But in order to establish that the *quantity* of a certain kind of cognitive state is explanatorily important, the states in question must share a property in virtue of which they qualify as the same kind of cognitive state. But any such shared property seems to 'screen off' their distinct causal roles, meaning that their role in explanation is not *qua* distinct vehicles. The very concept of a representational vehicle precludes vehicles from playing an explanatory role in virtue of their quantity. If Rupert wants to argue that the quantity of states involved is explanatorily important, he will have to conclude that the states in question are merely realizers of the same vehicle.

Rupert's case for both the Quantity and Vehicle claims relies on an appeal to the same set of empirical cases. In the next part of this paper, I'll describe these cases and explore the extent to which they can be interpreted as supporting either the Quantity claim or the Vehicle claim.

### *3. Rupert's argument*

#### *3.1 The empirical evidence*

Rupert claims that both his Vehicle and Quantity claims are supported by appeals to empirical cases. He covers a number of examples in the paper, but I'll focus here on the first two, before bringing in a third which seems to better make Rupert's point.

Rupert's first empirical case is taken from the experimental data concerning the visual enhancement of touch. He cites experiments where subjects are asked to make judgments about parts of their body being touched, including cases where the subject can visually perceive the body part at the same time as it is being touched. The data suggest that in the latter case, where subjects can both see and feel the stimulation, their judgments are more accurate and made more quickly than when the subject has only tactile stimulation. Rupert claims that the best explanation of the subjects' increased accuracy and reaction time is a result of interaction between visually activated representations in multisensory areas and representations in primary somatosensory cortex. He concludes that "[e]nhanced tactile sensitivity results specifically from the cooperation of multiple [...] representations" (106-107).

Rupert's second case concerns motor commands and efference copies. It is widely accepted that when the brain sends a motor command (e.g. to reach out one's hand, or to move one's eyes), it also uses a copy of that command to simulate the results. If the simulated result doesn't match the goal state, a corrective motor command can be sent: this allows the brain to correct motor errors faster than waiting for environmental feedback. Rupert claims that the subject's ability to quickly and accurately correct mistakes in their motor behaviour is best explained by the fact that the motor instruction is not operating alone: a copy of it is simulating the behaviour and identifying mistakes early. He concludes that multiple representations are responsible for producing the effect: "Whereas one might have thought that the brain-mind issues a single command to move a limb, the efference copy increases that number" (107).

Given that Rupert intends to show that quantity is playing a key explanatory role, his first two examples don't seem particularly helpful: in each case the result is achieved by the addition of only one further cognitive state. The sort of picture that Rupert has in mind is probably best illustrated by an example to which he devotes less consideration. This third case concerns facial recognition. When we consciously experience an image of a face, there is a distinctive pattern of neural activation in a

part of the cortex known as the fusiform face area (FFA). When subjects are presented with a face stimulus to only one eye, and a colour-cancelling stimulus is presented to the other eye, subjects report seeing only a homogenous coloured field instead of a face. The FFA is still activated in such cases, but to a lesser extent than when the face is consciously experienced. Rupert claims that the subject's ability to report seeing the face stimulus is a function of the number of active representations in the appropriate area of the cortex: "the more face representations active in FFA, for instance, the more likely the subject is to report verbally that she's seeing a face" (110).

### *3.2 Establishing the Vehicle claim*

I shall now explore the extent to which these experimental results establish Rupert's Vehicle claim, that there are distinct types of representational vehicle involved in the explanation, and not merely distinct realizers of the same representational vehicle.

Rupert is clear that he intends his claims about the massively representational mind to be claims about distinct vehicles of representation. He does not explicitly say why we should understand the cognitive states involved in his examples as distinct vehicles; instead he assumes that it will be obvious from the scientific data: he "take[s] various empirical results to establish a massive number of distinct representations" (110-111). But while the empirical results might give us reason to think that there are distinct states involved, whether these constitute distinct vehicles of representation will depend on how we are individuating the states in question.

Rupert doesn't explicitly say how he is individuating representational vehicles, other than to rule out individuating them by their referential content (110-111). This is unsurprising: as we have seen, such an approach would be unable to accommodate cases where co-referential states played distinct causal roles.<sup>6</sup> Rupert's interpretation of the empirical cases confirms that he takes vehicles to be

---

<sup>6</sup> It's worth noticing that Rupert's rejection of the semantic approach to vehicle-individuation is in no way controversial. As Piccinini (2008) points out, it is highly unlikely that anyone subscribes to a wholly content-



individuated by their non-semantic properties, but it is not clear whether he thinks the properties in question are formal, functional, or physical. Since all such approaches individuate vehicles by their causal properties, however, this is of little consequence.

Consider Rupert's first case, concerning the visual enhancement of tactile stimulation. He proposes that the increased sensitivity to touch results from the addition of a multisensory representation to the existing somatosensory representation. But why is Rupert assuming that these two representations are distinct vehicles? On one hand, he might be adopting a purely physical approach to vehicles individuation: claiming that these two representations are distinct vehicles in virtue of being neural states in distinct locations in the brain, one in a multisensory area of the cortex and the other in the primary somatosensory area of the cortex. On the other hand, however, Rupert might be adopting a functional approach to vehicle individuation: claiming that these two representations are distinct vehicles in virtue of having distinct functional roles, one playing a visually-activated role and the other playing a tactile role. Representations in multisensory areas of the brain and representations in somatosensory areas of the brain are likely to be available to different kinds of processes.

Rupert takes his second case, concerning motor commands and efference copies, to establish that there are two distinct vehicles involved in generating fast and accurate motor control: he proposes that increases in speed and accuracy result from the addition of an efference copy representation to the existing motor command representation. Again, there are different possible justifications for taking these representations to be distinct vehicles. We could interpret Rupert's case as resting on a purely physical approach to vehicle individuation, based on neural location: he claims that the motor command vehicle is sent by "one part of the cortex" while the efference copy vehicle "is sent somewhere else in the brain" (107). But this claim could also be interpreted as evidence that Rupert

---

based approach to vehicle individuation, on which only the semantic properties of a vehicle are relevant to its individuation.

is individuating vehicles functionally: the motor command representation plays the role of instructing behavior, while the efference copy representation plays a simulatory role.

The third case under consideration, involving facial recognition and the FFA, is less helpful in understanding Rupert's approach to vehicle individuation. He proposes that in cases where we consciously perceive and report on the presented facial stimulus, there are a greater number of representations active in the FFA than in cases where we don't consciously perceive the stimulus. But whereas the first two cases clearly involved representational states, the experimental data in this case merely states that there is increased neural activity in the relevant area of cortex. Rupert proposes that "what is sometimes treated as strength of activation in a given area of cortex might be better understood in terms of the number of [...] representations active in that area of cortex" (109), but doesn't argue for this representational interpretation. And it is harder still to justify interpreting these states as distinct representational vehicles: this would require them to be different ways of representing the referent (the facial stimulus). But in what sense do individual neurons or neuronal events constitute *different ways of representing* the stimulus? Unlike Rupert's previous examples of distinct vehicles, it's not clear that there are any representational mechanisms which are sensitive to distinct properties of these individual states: they are not in different areas of the brain, for example, nor do they have access to different cognitive functions.

In conclusion, it is difficult to say with certainty which non-semantic properties Rupert is using to individuate vehicles of representation, because he does not say what sort of cognitive mechanisms he understands to be operating over the vehicles. Without knowing which non-semantic properties the representational processes are sensitive to, we cannot state with any confidence the causal properties of the vehicles that might type-individuate them. But while it is at least plausible that there are relevant causal properties in the first two cases, the third case is more puzzling. In order to claim that the increased neural activity is to be understood in terms of vehicle quantity, there would

need to be processes which behave differently with regard to each candidate vehicle. Rupert has given us no reason to think that there is such a mechanism in the third case.

Rupert's claim about distinct vehicles is thus largely stipulated rather than argued for. But even if we grant Rupert the benefit of the doubt on this matter, his claims about the massively representational mind still require him to demonstrate that distinct vehicles do their explanatory work in virtue of their quantity.

### *3.3 Establishing the Quantity claim*

Let us allow that Rupert has established that the cognitive states doing explanatory work in the empirical cases he cites are distinct vehicles of representation. This would be to grant him the Vehicle claim, but it still leaves the status of the Quantity claim to be established. Are these supposedly distinct vehicles doing explanatory work in virtue of their quantity, or in virtue of some other property? To answer this we must turn back to the empirical cases, which Rupert takes to be supportive of the Quantity claim as well as the Vehicle claim.

In the first case, involving enhanced tactile sensitivity, Rupert claims that "activation of the additional bodily representation accounts for enhanced tactile sensitivity" (106): the addition of the visual representation has an effect that the tactile representation alone does not. But this does not entail that the visual representation has this effect in virtue of being a *second* representation, i.e. in virtue of increasing the quantity of representational vehicles present. If this were the case, then we could expect to substitute a different representation (e.g. an auditory representation) for the visual representation without changing the effect, because the number of representations would remain the same. But this seems unlikely, especially compared to the alternative explanation that the visual representation is playing a distinct causal role in generating the behavioural effect, perhaps in virtue of its being a *visual* representation. The increased sensitivity could be a result of adding a new

representation that has access to a particular set of processes to which at the original representation lacked access, for example. It looks like Rupert needs to either rule out such an explanation, or show why an explanation in terms of quantity alone would be preferable, in order to establish the Quantity claim.

In the second case, involving motor commands and efference copies, Rupert claims that motor action is more accurate when the motor command is supplemented by an additional representation: “Whereas one might have thought that the brain-mind issues a single command to move a limb, the efference copy increases that number” (107). But again, this does not show that the addition of a second representation achieves this effect in virtue of increasing the *quantity* of representations involved. If quantity were the key explanatory factor, then supplementing the motor command with a representation other than the efference copy (e.g. a memory of performing the same action) should produce the same results. Again, this seems unlikely: the standard explanation of the increased accuracy is in terms of the distinct causal properties of the efference copy and the motor command. While the motor command generates bodily movement, the efference copy generates simulations of bodily movements. Again, it looks like Rupert needs to show why such an explanation should be rejected in favour of the claim that the efference copy increases motor accuracy by adding to the *quantity* of representations involved.

In the third case, involving facial recognition and FFA activation, Rupert claims that the subject’s ability to report on the stimulus is correlated with increased activity in the appropriate area of the cortex. He treats strength of cortical activity as “a function of the number of representations active in an area densely packed with representational vehicles” (112), and proposes that “the more face representations active in FFA, [...] the more likely the subject is to report verbally that she’s seeing a face” (110). In this case, it might initially seem more plausible that the additional representational vehicles are playing an explanatory role in virtue of their quantity, because it is less obvious that there is an alternative explanation in terms of their causal properties: as described, they don’t seem

to be playing distinct causal roles from one another within a cognitive mechanism. But if this third case seems to offer greater support than the other two cases for Rupert's Quantity claim, it does so at a price. If the representations aren't playing distinct causal roles from each other within some cognitive mechanism, then Rupert hasn't given us a reason to think that they are distinct vehicles. To use this case to support the Quantity claim seems to diminish the (already weak) support it can provide for the Vehicle claim. Both Quantity and Vehicle claims are required to establish Rupert's conclusions about the massively representational mind. This poses a pressing problem for Rupert that I'll consider shortly. First, however, it's worth addressing the question of why Rupert thinks that explanations in terms of quantity are to be preferred in these three cases, and why he doesn't consider the obvious alternative explanations.

### *3.4 Diagnosing the explanatory problem*

Why does Rupert think that his empirical cases demonstrate that quantity of representational vehicles is doing the explanatory work? Importantly, Rupert thinks that representational *content* cannot be playing an explanatory role. Like other naturalist proponents of mental representation, Rupert is committed to the claim that mental content is referential content, on the grounds that such theories of content "better satisfy naturalistic scruples" (102). Rupert emphasises that in each of the empirical cases he discusses, the representations in each case have the same referential content as each other: the visual and tactile representations in the first case both represent the same body part; the motor command and efference copy in the second case both represent the same action; the FFA representations in the third case are all representations of the same face. (Rupert takes the sameness of content to have been established by the example in question, and

does not argue for this conclusion. While I take his claims about sameness of content to be controversial, I will accept them in what follows.<sup>7)</sup>

Sameness of content, according to Rupert, makes these representations “redundant”. He claims that “the brain is rife with representations that possess overlapping and redundant content” (99), and that any account of our cognitive processes “must take into account the widespread effects of a buzz of representations, many with redundant content” (100). His claims about the massively representational mind thus include a commitment to the claim that each of us possesses many distinct representational vehicles with the same content: “For any given subject *S*, for very many representational contents *C*, *S* possesses many more than one vehicle with *C* as its content” (102). Rupert thinks that the existence of distinct vehicles with the same content is sufficient to make those contents in some sense redundant.

With respect to the empirical cases, Rupert seems to think that the redundancy of the representations involved – the fact that they have the same content – rules out a certain form of explanation, making way for his alternative explanation in terms of the quantity of representations involved. And he is right that there is a sense in which representations with the same content are explanatorily redundant. If someone behaves differently towards dog and cats, for example, we might explain their behaviour in terms of the distinct animals represented by their cat-representations and their dog-representations. But if someone behaves differently towards the planet Venus in the morning (e.g. by calling it Hesperus) and in the evening (e.g. by calling it Phosphorus), we can’t explain that difference in behaviour solely in terms of the planet Venus. The fact that the same referent is picked out both times means that we can’t point to the referential content as the explanatory factor in accounting for the two distinct behaviours: in Rupert’s terms, the content is explanatorily redundant. The empirical cases that Rupert cites are similar to the Venus

---

<sup>7</sup> Notice that in each of Rupert’s cases, there are other candidates for referential content that Rupert doesn’t consider. In his first case, the content could be the state of the affairs that the arm is being stroked; the content could be the arm itself or a particular part of the arm; or the content could be a property such as the frequency or intensity of the arm-stroking. I take it that there are similar issues with respect to the other empirical cases Rupert cites.

case, as each case involves a behavioural difference that cannot be explained solely in terms of the contents involved, because each case only involves one referential content (i.e. the body part, the action, the face).

As we have already seen, however, this is an established problem with psychological explanations that rely solely on referential content. Cognitive psychologists generally deal with this problem by appealing to the non-semantic properties of a representation (e.g. its formal, physical, or functional properties) to play the sort of explanatory role traditionally according to Fregean sense or narrow content: to show how the same referent can be represented in different ways. Where the representational contents are 'redundant', the causal properties of representational vehicles can play an explanatory role in psychological explanation. And indeed this is what we find when we look at Rupert's empirical cases. In each case, there is a change in behaviour as a result of more than one state with the same content. But while the co-referential content might be explanatorily redundant, the distinct representational vehicles can play different causal roles that account for the behavioural change in question. The relevant features of the vehicles are often described as constituting the 'narrow content' of the representation. But Rupert's examples of 'redundant' representations are representations with the same referential content only: they quite clearly have distinct narrow contents.

Rupert thinks that the quantity of representations involved can play an explanatory role in psychological explanation because he takes himself to have ruled out the alternative explanation in terms of the contents of the representations involved. But this is to overlook the fact that psychological explanation in terms of representations can (and generally does) appeal to the causal properties of the vehicles of representation. So Rupert is right that sameness of referential content rules out a certain kind of representational explanation, this does not pave the way for his own explanations of the behaviour in terms of the quantity of representations involved. In order to make this case, he would have to either rule out an explanation in terms of the different causal properties

of distinct vehicles, or show why his quantity-based explanations are to be preferred. In what follows, I will argue that neither strategy is available to Rupert.

#### *4. The incompatibility of Vehicle and Quantity claims*

##### *4.1 The explanatory role of representational vehicles*

One might think that Rupert's task here is a straightforward one: he needs to engage in further interpretation of the empirical cases, in order to persuade us that there are good reasons to endorse his explanation in terms of the quantity of representations rather than in terms of their distinct causal contributions to the system. But if we look back to his third case, involving facial recognition, then it becomes apparent that the situation is more complicated.

Recall from Section 3.3 that the Quantity claim seemed most persuasive in the facial recognition case, where the large number of representations active in a part of the cortex was correlated with verbal reports of face perception. The idea that these representations are playing an explanatory role in virtue of their quantity seems more reasonable here, because there isn't an obvious difference in the causal properties of each representation that would offer an alternative explanation. But Rupert needs there to be such a difference in causal properties in order to claim that the representational states are distinct vehicles. It is generally accepted in these cases that there is increased neural activity, but Rupert thinks that a "plausible gloss of such discussions [...] treats cortex as itself being saturated with distinct representational vehicles and treats variations in strength of cortical activity as being a function of the number of representations active in an area densely packed with representational vehicles" (112). In order to make this claim about distinct vehicles, Rupert needs to hold that each state has distinct causal properties which allow it to play a different processing role from the other states. Notice that if these representations are treated alike



by the cognitive mechanisms involved, they are (by definition) not distinct vehicles of representation. They are merely realizers of the same vehicle.

I propose that a similar problem befalls the other empirical cases. The more supportive they are of the Vehicle claim, the less able they are to support the Quantity claim, and vice versa. In the first case, for example, we'd be more inclined to judge that quantity of representations was playing an explanatory role if the representations involved were in the same area of the brain, with access to the same processes, and playing similar roles within the cognitive system. But then we'd be less likely to agree with Rupert that these are distinct vehicles of representation: they'd look more like realizers of the same vehicle. And in the second case, the claim about distinctness of vehicles is a strong one because the motor command and the efference copy clearly have distinct causal properties that allow the system to treat them differently. But here there is the least motivation to say that the increased accuracy of behaviour is a function of the quantity of representations, rather than their distinct functional profiles.

Recall that Rupert's aim is to establish that the mind is massively representational, in the sense that "variation in the sheer number of vehicles [...] plays a causal-explanatory role in the production of certain forms of behavior" (111). He thus needs to establish both his Vehicle and Quantity claims to make his case. But it looks like any case in which the representations are distinct *vehicles* will, in virtue of the distinct causal roles played by those vehicles, provide an explanation in terms of the different ways that the vehicles represent their referents – removing the need for a quantity-based explanation. And any case in which the representations look to be doing explanatory work in virtue of their *quantity* will be one in which representations can be replaced without changing the outcome, as long as the overall quantity or proportion of representations is kept constant. In such cases, there is little motivation to understand these representations as distinct vehicles of representation, because they are playing similar causal roles within the system.

It might be objected that the seeming incompatibility between the Vehicle and Quantity claims is only a problem for the particular empirical cases under discussion, and that Rupert would have a better chance of establishing his claim that the mind is massively representational if he focused on different examples. But as I see it, the problem stems from the very concept of a representational vehicle. The point of individuating representations non-semantically in terms of their vehicle properties, as well as semantically in terms of their referential contents, is to allow us to give a causal-mechanistic psychological explanation – particularly in cases when purely content-based explanations are insufficient. To pick out representational vehicles is to individuate the underlying neural states in a particular way: the way in which they play distinct causal roles within the cognitive system; the way in which they constitute different *ways of representing* a referent. And Rupert seems to realise that his ‘redundant’ representations are distinct modes of presentation of the same referential content:

“In other words, for whatever property, individual, or kind a subject sometimes represents, she has lots of ways of representing it—where having “different ways of representing *C*” amounts to nothing more than having distinct, psychological-level units that represent *C*.”

(102)

This means that wherever we find distinct vehicles, there will be an explanation of their contribution to the behaviour in terms of these different causal roles that constitute different ways of representing. To replace such an explanation with one in terms of the quantity of representations, these causal differences need to be irrelevant: they need to be screened off by another property, e.g. quantity of representations. But whatever property screens off the causal differences between the vehicles effectively makes their status *qua* vehicles irrelevant. To individuate representations as vehicles is to commit to their playing explanatory roles within a system in virtue of their individuating properties; if the representations in question are no longer playing that explanatory role, as would be the case if that role was screened off by a quantity-based explanation, then it is no

longer appropriate to individuate them as distinct vehicles. Any quantity-based explanation would have to claim that quantity of representations screens off the differences between vehicles, such that the vehicles would be interchangeable within the system. But this is just to say that the vehicles would be treated the same by the system's processes, which leaves us with no justification for thinking of these representations as distinct vehicles.

This seems to leave Rupert with a dilemma: he can either choose to hold onto the Vehicle claim or the Quantity claim, but he will not be able to establish both.

#### *4.2 Rupert's options*

Since Rupert's project is to establish the claim that the mind is massively representational, it might make sense for him to focus on the Quantity claim: to try to establish that psychological explanation relies on the quantity of representations involved, even where these are not distinct vehicles. But notice that this makes Rupert's 'redundant representations' truly redundant: in addition to sharing the same content, they are now also instances of the same vehicle. But if representations can only be individuated semantically by their contents, or non-semantically by their vehicles, then on this move there would be no sense in which the cognitive states in Rupert's examples would be distinct representations. To speak of these cases as involving numbers of representations, or proportions of representations, would be misleading. The only way in which Rupert could appeal to the quantity of mental states involved in these cases would be to acknowledge that these states are merely distinct realizers of the same representational state.

If Rupert instead focuses on the Vehicle claim, and maintains that the mental states in question are indeed distinct vehicles of representation, then he is restricted in what he can claim about quantity. He can't claim that psychological explanation appeals to numbers of distinct vehicles, if he means this in the sense that the quantity itself, rather than the individual contributions of each vehicle, is

playing the key explanatory role. He could claim, however, that we may have many more vehicles with co-referential content than has been previously acknowledged. But this claim would leave the standard practise of psychological explanation untouched. For each psychological phenomenon, we will be able to explain it in terms of the distinct causal contributions of the many different vehicles. It might be an interesting feature of the cognitive architecture to note that there are many co-referential vehicles, but there is no sense in which claims about quantity play an explanatory role in the explanation of each psychological phenomenon.

Notice that on its own, there is nothing particularly distinctive about Rupert's claims about redundant representations. The idea that there can be more than one vehicle with the same content is built into the very foundations of cognitive psychology.<sup>8</sup> If Rupert wishes to make the more controversial claim that this phenomenon is larger than previously thought, then he has work on his hands. He would have to develop his empirical cases to demonstrate that these redundant representations are in fact cases of sameness of content carried by distinct vehicles – as we saw, this is something that he stipulates rather than argues for. Furthermore, the more persuasive of these cases (tactile sensitivity, motor accuracy) each only involved two distinct vehicles, rather than many. The case that looked most likely to support claims about quantity was the FFA case, but this is much harder to understand as a case of redundant representation, because it is not clear that this is a representational case at all.

#### *4.3 Rethinking psychological explanation*

At this point, it is worth taking a step back and reflecting on the importance of representational vehicles in cognitive psychology. Recall that vehicles are a way of individuating representations non-

---

<sup>88</sup> Fodor, for example, takes the view that vehicles are individuated with respect to their semantic *and non-semantic properties*: "Vehicles, like other symbols, are individuated with respect to their syntactic and semantic properties [...] two vehicle tokens are type distinct if they are syntactically different or if they express different propositions" (Fodor 1989, 67).

semantically. If representations are individuated wholly semantically, in terms of their referential contents, we can't explain the same representation can play different roles in our cognition. Our representations of Hesperus and Phosphorus, individuated semantically, are tokens of the same representation type. If cognitive psychology were to individuate representations semantically, it would not be able to explain the different ways in which we think about and behave towards the planet Venus. Notice, however, that the standard examples of such 'Frege cases' involve *conceptual* representations: the amodal states whose contents constitute the contents of our thoughts. Frege cases arise because co-referential conceptual representations do not always play the same role in the different thoughts to which they contribute. Individuating representations by their non-semantic properties (i.e. as vehicles) can thus be understood as a solution to the problem of co-referential *concepts*.

Not all representations, however, are conceptual representations. Conceptual representations contribute to our central cognitive processes, such as thinking, inferring, and decision-making. But cognitive psychology is also concerned with the nonconceptual cognitive processes that take place in cognitive subsystems, rather than centrally. These representations are subdoxastic or subpersonal, in the sense that their contents are generally not available to the cognizer.<sup>9</sup> When early visual processing builds up a three-dimensional description of the scene, for example, or when linguistic processing yields judgments concerning the grammaticality of a sentence, the representations involved in these processes are not available to our central thought processes. The representations in question are generally understood to be nonconceptual.<sup>10</sup>

Are nonconceptual representations subject to Frege cases? In other words, are there cases where two *nonconceptual* representations have the same referential content but differ in their causal role?

Does it even make sense to think of nonconceptual representations as having 'cognitive significance'?

---

<sup>9</sup> See Drayson (2012, 2014) for elaboration of the circumstances under which representations are subdoxastic or subpersonal.

<sup>10</sup> Bermúdez (1995) argues that subdoxastic computational states should be understood as having nonconceptual content.

when the causal contributions they make are not to the thoughts of the cognizer? The answers to these questions are beyond the scope of this paper, but if nonconceptual representation does not give rise to the same issues as conceptual representation, then it is not clear that we need to use the same approach to individuation for them both. Without the possibility of Frege cases, nonconceptual representations could be individuated wholly semantically, in terms of their referential content. There would be no need for the 'vehicle' talk that accompanies the non-semantic individuation of conceptual representations, and thus no need to individuate representations by their causal properties.

This is important because Rupert's examples all seem to draw on nonconceptual representations. (One does not need to have a FACE concept to have FFA activation in the presence of faces, for example.) I am suggesting that if nonconceptual representations don't generate Frege cases, then they could be individuated wholly semantically. At first, this may not look helpful to Rupert: he wants the representations in question to be distinct vehicles, which would require non-semantic individuation. I want to suggest that Rupert's focus on the vehicles of representation is misplaced. He is motivated to focus on representations as distinct vehicles primarily because he wants to avoid the states in question being classified as "mere realizers":

"I take being a particular vehicle to be a psychological-level construct—unlike mere realizers, which appear only at some lower level than, or as part of an explanatory enterprise distinct from, psychology." (111)

Thinking about representations as either vehicles or their realizers is not, however, the only option. In cases where nonconceptual representations are individuated wholly semantically, two representations with the same referential content are *tokens* of the same representation-*type*. They are neither the same vehicle nor realizers of the same vehicle. To see this, consider that realization is

a relation between a second-order functional state and the first order state that has that function.<sup>11</sup> When we individuate representations as vehicles, we are appealing to their second-order functional role, while the first order states that play that functional role are its realizers. But if nonconceptual representations don't generate Frege cases, then we don't need to invoke functionally-characterized vehicles to explain the different causal contributions of co-referential representations. And without second-order functional vehicles, there are no first-order realizers. Instead, we can allow that two representations with the same referential content are distinct tokens of the same representational type. The type/token distinction does not require a distinction between first-order and second-order states, and so there is no requirement that token representations be thought of as "mere realizers".

Explanatory concerns motivate Rupert's desire to avoid characterising the representations in his examples as realizers of the same vehicle. He thinks that vehicles appear in psychological explanation, while their realizers appear in explanations at lower (presumably neuroscientific) levels (111). This view of the relationship between psychology and neuroscience is a familiar one, in which psychological explanations are not reducible to lower level explanations precisely because psychological kinds are functional kinds which can be multiply realized at a lower level.<sup>12</sup>

Rupert's examples of multiple representations contributing to a behavioral phenomenon can, I have suggested, be understood as token representations of the same type – where the type is not a functionally-individuated vehicle, and the tokens are not its realizers. If these distinct token representations can bring about an effect in terms of their quantity, it is unclear where such an explanation would fit into Rupert's current taxonomy: it is neither the explanation in terms of vehicles that Rupert wants, nor the explanation in terms of their realizers that Rupert rejects.

Perhaps, then, Rupert should reconsider his characterization of psychology as an autonomous science of functional kinds. He would not be the first: people working in embodied cognition,

---

<sup>11</sup> For a comprehensive discussion of the realization relation and its applicability to psychological states, see Polger and Shapiro (2016).

<sup>12</sup> This picture of the autonomy of psychology is most often associated with Fodor (1974).

proponents of bottom-up methodologies in the mind sciences, and type-identity theorists in philosophy of mind have all criticised the drawing of a distinction between psychology and lower level sciences in terms of functional kinds and their realizers.<sup>13</sup> I propose that if Rupert is keen to pursue his theories of the massively representational mind, he too should dismiss the idea that psychology is a science of functional kinds (vehicles) and that neuroscience is a distinct science of their realizers.

### *5. Conclusion*

Claims about the vehicles of representation are claims about a particular way of typing the realizers of representational properties: a way of individuating neural states as representations, but in terms of their non-semantic properties rather than in terms of what they refer to or represent.

The non-semantic properties of representations that individuate distinct vehicles from each other are those properties in virtue of which the representational system treats states as similar or different for the purposes of processing the representations. This means that the individuating non-semantic properties of representational vehicles will differ according to the nature of the representational system involved. Classically computational approaches to cognition, for example, propose that vehicles are symbols in a language of thought. Shea (2007) has proposed that in connectionist approaches, the vehicles are clusters or regions in state space; Clark (2001) proposes that in dynamical systems, the vehicles are attractor basins.

Vehicles of representation, whatever their individuating features, play an important role in psychological explanations. They constitute the distinct ways of representing referential contents: they are the naturalist alternative to Fregean senses, offering distinct modes of presentation for co-referring mental states, and thus allowing co-referring mental states to play distinct causal roles in

---

<sup>13</sup> Criticisms of psychology as a science of multiply realized functional kinds can be found in Bickle (1997), Bechtel and Mundale (1999), and Shapiro (2000).



psychological explanation. But because vehicles play this explanatory role in representational explanations of behaviour, then alternative approaches to psychological explanation cannot help themselves to the notion of representational vehicles while still claiming to be *alternative* approaches to psychological explanation. Either vehicles are doing the explanatory work, in which case this is a standard explanation in cognitive psychology; or vehicles aren't doing the explanatory work, in which case the states in question are not vehicles of representation.

If representational vehicles are an essential feature of psychological explanations, then Rupert will be unable to demonstrate that there are psychological explanations which appeal to the quantity of vehicles involved. Rupert himself seems wedded to the idea that representations either appear as vehicles in psychological explanations, or as their realizers in lower level explanations, and he is keen to avoid the latter. I have argued that the best way forward for Rupert's theory of the "massively representational mind" would be to consider a different approach to thinking about representations. If the representations in which he is interested are nonconceptual, and if nonconceptual representations don't generate the same problems (e.g. Frege cases) as conceptual representations, then individuating them by their non-semantic properties as vehicles may not be necessary. Instead, we could individuate nonconceptual representations wholly semantically, by their referential contents. If Rupert is claiming that the quantity of such representations with the same content plays an explanatory role, then the representations involved could be distinct token representations of the same representational type. While this would not meet Rupert's Vehicle criterion, it fulfils his desire that the distinct representations should not be mere realizers of the same vehicle.

This leaves interesting questions about whether the resulting explanations should be understood as psychological or not. In Rupert's case, this may require him to re-examine his understanding of the relationship between psychology and lower level sciences.

## Acknowledgements

This paper was presented at the 'Realizability and the Levels of Reality' workshop at the University of Paris 1 in June 2014. A shorter version was written for a workshop at the University of Edinburgh in August 2013. I am grateful to the audience at both events for their comments, to Alexandru Manafu for his editorial work, and to two anonymous referees for their helpful feedback. Particular thanks are due to Rob Rupert for being a patient and illuminating interlocutor throughout.

## References

- Aydede, Murat (2000). On the type/token relation of mental representations. *Facta Philosophica* 2 (1):23-50.
- Bechtel, William and Mundale, Jennifer (1999). Multiple Realizability Revisited: Linking Cognitive and Neural States. *Philosophy of Science* 66 (2):175-207.
- Bermúdez, José Luis (1995). Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States. *Mind and Language* 10 (4):333-369.
- Bermúdez, José Luis (2005). *Philosophy of Psychology: A Contemporary Introduction*. Routledge.
- Bickle, John (1997). *Psychoneural Reductionism: The New Wave*. MIT Press.
- Clark, Andy (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York: Oxford University Press.
- Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese* 28 (2):97-115.
- Fodor, Jerry A. (1989). "Substitution Arguments and the Individuation of Belief" in *A Theory of Content and Other Essays*, J. Fodor, Cambridge, MA: MIT Press, 1990. (Originally appeared in *Method, Reason and Language*, G. Boolos (ed.), Cambridge, UK: The Cambridge University Press, 1989.)

- Piccinini, Gualtiero (2008). Computation without representation. *Philosophical Studies* 137 (2):205-241.
- Polger, Thomas W. & Shapiro, Lawrence A. (2016). *The Multiple Realization Book*. Oxford University Press UK.
- Rupert, Robert D. (2011). Embodiment, Consciousness, and the Massively Representational Mind. *Philosophical Topics* 39 (1):99-120.
- Shapiro, Lawrence A. (2000). Multiple realizations. *Journal of Philosophy* 97 (12):635-654.
- Shea, Nicholas (2007). Content and Its vehicles in connectionist systems. *Mind and Language* 22 (3):246–269.
- Stich, Stephen P. (1983). Some evidence against narrow causal theories of belief. In *From Folk Psychology to Cognitive Science*. MIT Press.