Synthese (2023) 202:175
https://doi.org/10.1007/s11229-023-04373-2

ORIGINAL RESEARCH

Organized representations forming a computationally useful processing structure

Nicholas Shea[1,2]

Received: 19 December 2022 / Accepted: 4 October 2023 / Published online: 21 November 2023

Abstract

Peter Godfrey-Smith recently introduced the idea of representational 'organization'. When a collection of representations form an organized family, similar representational vehicles carry similar contents. For example, where neural firing rate represents numerosity (an analogue magnitude representation), similar firing rates represent similar numbers of items. Organization has been elided with structural representation, but the two are in fact distinct. An under-appreciated merit of representational organization is the way it facilitates computational processing. Representations from different organized families can interact, for example to perform addition. Their being organized allows them to implement a useful computation. Many of the cases where organization has seemed significant, but which fall short of structural representation, are cases where representational organization underpins a computationally useful processing structure.

Keywords Representational theory of mind · Representational organization · Representational structure · Computational processing · Semantic content

1 Introduction

Representations are a fundamental explanatory tool. They provide the basis for behavioural explanation throughout the cognitive sciences. So it is important to understand their properties. One notable characteristic of some representations is that they come in organized families. Honeybees use a waggle dance to signal the location of nectar: the more waggles, the greater the distance to nectar. (Similarly with angle of the dance and the direction of nectar.) There are neurons in the brain whose firing

✉ Nicholas Shea
nicholas.shea@sas.ac.uk

1 Institute of Philosophy, School of Advanced Study, University of London, Senate House, London WC1A 7HU, UK

2 Faculty of Philosophy, University of Oxford, Oxford OX2 6GG, UK

rate represents numerosity: the greater the firing rate, the larger the number of objects being represented. A visual bath thermometer varies in colour with the temperature of the water: redder is hotter, bluer is colder. In each of these cases, representations form a family. Only one representation in the family is tokened at a time. Different representations within the family are systematically related.

In all of these cases, the salient interrelation within the family is that similar vehicles carry similar contents.[1] The representations display 'organization' (following Godfrey-Smith, 2017). What is the representational significance, if any, of the fact that similarity corresponds to similarity? How does that representational significance arise? Some merits of representational organization have been noted before. The aim of this paper is to add another merit—a deeper reason why representations often come in organized families. Organized representations from different families can interact in a way that is suited to moving between contents in useful ways. If, neurally, the firing rate of two neurons is added together in a third neuron, then that transformation is suited, at the level of content, to calculating the sum of the quantities represented by the first two neurons. If used to perform addition, the pattern of transitions between representational vehicles is faithful to their contents. So it is an interaction between vehicles that is useful for performing computations—'computation' here understood broadly to cover content-faithful transitions of any kind, not just those characteristic of classical computationalism.

The purpose of this paper is to show that organized representations can enter into internal processing in a way that is relied on to perform useful computations. Possessing representations that form organized families is therefore an advantage. Their computational role is also a reason why organization is semantically significant: why it is right to capture the semantic mapping from vehicles to contents at the level of the family as a whole, in terms of a systematic function which maps similar vehicles to similar contents.

Section (2) characterises representational organization precisely and distinguishes it from structural representation. Section (3) shows how the distinction applies to mental representations. Section (4) argues that interactions between organized families of vehicles can form the basis of a computationally useful processing structure. Section (5) drills down into the difference between being organized and being symbolic (in one sense of that term). Section (6) concludes by defending the central claim—that organization can form part of a processing structure which is useful for implementing certain computations—against the charge of liberality, hence triviality. It argues that the idea of computational usability relied on here is not so liberal as to be unexplanatory.

[1] There may be other kinds of interrelation amongst representations that should count as instances of representational organization. The discussion here is confined to similarity relations since this covers the canonical examples.

2 Representational organization

Before getting into the merits of representational organization (Sect. 4) we need to characterise the phenomenon more precisely. Representations in a family display *organization* when the function mapping vehicles to contents applies systematically across all vehicles in the family, mapping similar vehicles to similar contents.[2] This section uses examples to elaborate on that definition, re-stating it in greater detail at the end of the section.

First off, we need to forestall a potential terminological confusion. When a family of representations display organization it is tempting to call the similarity relation within the family a ‘structure’. But that would be misleading. Within-family similarity is not structure in the sense that a sentence is structured or a map is structured. It is not a relation between parts of a representational vehicle. The vehicle doing the representing is not structured. The bath thermometer simply displays a patch of colour. The patch has no structure that relates to its meaning. Any representational vehicle will of course have some kind of structure—physical parts, say—but in cases like the bath thermometer the structure of the vehicle has no semantic significance. It is not like the compositional structure of a map or of a sentence—structure that systematically relates to the content of the whole. Kevin Lande uses the qualifier ‘extrinsic’ to distinguish these cases, using the term ‘extrinsic structure’ for the interrelations in an organized family (Lande, 2021), but I will avoid referring to the similarity relation between different vehicles drawn from the same organized family as a kind of ‘structure’. We are interested in a feature that characterises the family of representations—a systematic relation between them. I will reserve ‘representational structure’ for the relation between parts of a complex representation.

Peter Godfrey-Smith introduced the term ‘organization’ (Godfrey-Smith, 2017, pp. 279–280). He explains organization by contrasting it with an example of a non-organized, *nominal* representational system. Paul Revere used a system of lanterns to signal of the arrival of the British army: *one if by land, two if by sea*. The number of lanterns had no semantic significance: this was a nominal representational system. If, instead, they had used the brightness of a lantern to signal the size of the approaching army, that would have been an organized representational system. Similar vehicles (light intensity) would map to similar contents (size of approaching army) in a way that is semantically significant. The right way of giving the content of these representations would be by way of a systematic function specified at the level of the whole family—a mapping from light intensity to army size. The actual Paul Revere system was nominal (i.e. non-organized). Its content is given by a mapping from one vehicle to the content, *the British are arriving by land,* and from another vehicle to the content, *the British are arriving by sea*.

Godfrey-Smith’s helpful observation is that representations that are syntactically unstructured can still display organization. The brightness-signals-size scheme is one example. Syntactically structured representations do also display organization. Two

[2] This definition fits with the way Godfrey-Smith (2017, p. 279) introduces the term (he does not define it explicitly). Planer and Godfrey-Smith (2021, pp. 754–755) add two further requirements, applicable to sender-receiver signalling, that I do not adopt here.

complex representations (e.g. two sentences) can be similar because they share a syntactic constituent (e.g. 'Layla runs' and 'Layla sleeps'). The focus of this paper is on the kind of organization displayed by the brightness-signals-size scheme: a family of representations that display organization, but where the individual representations in the family lack semantically-significant constituent structure. Call this *mere* organization.

Godfrey-Smith's examples depend on explicit convention. Both nominal and organized representational systems are also found in nature, for example in animal signalling. Vervet monkeys have different alarm calls for eagles, snakes and leopards (Seyfarth et al., 1980). This is a nominal system. There is no systematic relation between calls that reflects what they represent (eagle, snake, leopard); nor the actions they direct (look up, look down, climb). The system relies on pointwise correlations between each call and a type of danger (Shea, 2018, p. 119). Camp (2021) nicely captures the contrast between this case and the bee dance: the vervet signalling system could equally well have one, or thirteen, different instances each implemented by a distinct mechanism. By contrast, the honeybee nectar dance is implemented in a mechanism that responds to all the dances in the family. It is an organized representational system, one in which similar numbers of waggles represent similar distances.

Some theorists have taken cases like the bee dance to be map-like. Maps, however, display a stronger property. They are a complex representation in which there are relations between parts of the representational vehicle; and that vehicle relation is itself representational, that is, it carries a representational content. That is to say, maps are an instance of structural representation:

> *Structural representation*
> A complex representation in which a relation on representational vehicles v_1, …, v_n represents a relation on the entities represented by v_1, …, v_n
> (Swoyer, 1991; Ramsey, 2007, pp. 77–92; Shagrir, 2012; Shea, 2018, p. 118).

Spatial relations between points on a map represent spatial relations between the locations represented by the points. In the nectar dance, there is indeed a similarity relation between different dances, but that relation is not a bearer of content. If two different dances are danced, the relation between them does not make a semantic contribution to the contents represented. No two dances are ever considered together. The obtaining of a relation between two token dances never makes a difference to downstream processing or behaviour (at least as the case is standardly described). It is not something that consumer bees are sensitive to. They react to the number of waggles in individual dances. So this is not a case where a relation on representational vehicles represents a relation on the entities represented (Shea, 2013a, 2014, 2018, 127)—it is not a case of structural representation.

The honeybee nectar dance is, however, an organized representational system. The semantic mapping from vehicles to contents applies systematically to the family of vehicles, mapping similar vehicles to similar contents. As noted, maps and sentences also display organization. They do so in virtue of their compositional structure. Nectar dances lack semantically-significant structure[3] but nevertheless form an organized representational system. They display mere organization.

Families of representations that display mere organization are often characterised as being analogue.[4] However, many other features are also in play in drawing the analogue–digital distinction (Peacocke, 2019b). Maybe the representational vehicle has to be continuous (Goodman, 1968, p. 160; Camp, 2007, p. 155). A related idea is that analogue representations are dense: between any two representations there is a third. An organized representational system need not be continuous or dense (Maley, 2011). The contrast between nominal representational systems and mere organization captures just one feature that theorists have taken to be distinctive of analogue representation. There is not scope here to join the wider debate as to how best to formulate the analogue–digital distinction. This paper will be relevant to that debate to the extent that being a mere organized representational system captures one feature that theorists have associated with being analogue.

To recap, we have seen examples of three types of representational system. A map is a structural representation. The honeybee nectar dance is a mere organized representational system. It is not a structural representation. And the signalling system used by Paul Revere (one lantern represents *the British are coming by land*; two lanterns represents *the British are coming by sea*) is a nominal representational system. The latter is also *symbolic*, in one sense of that term: similarities and differences between the different representations in the family have no semantic significance. We return to the property of being symbolic, in this sense, in Sect. (5).

When a family of representations display organization, similar vehicles represent similar contents. It not just that similar vehicles happen to map to similar contents, as they might in a pointwise system of the kind mentioned by Camp (where a system has thirteen different instances each implemented in a distinct mechanism). Something stronger must be true. The similarity between vehicles must have semantic significance. That is not something we can specify by just considering the mapping from vehicles to contents extensionally. It is a property of the vehicle-content mapping considered intensionally. Recall that a collection of representations form a family when instances of the different types (e.g. two waggles, three waggles, …) exclude one another. The family is a determinable. A determinate vehicle is tokened on each occasion. A family of representations display *organization* if and only if their content is specified by a systematic function, at the level of the whole family, which has the effect that similar vehicles represent similar contents. The similarity relation between

[3] Millikan (1984) argues that the bee dance does have compositional structure, with the time and place that a dance is performed being part of the semantically-significant constituent structure of the representation. For reasons that there is not space to go into here, I do not take time and place to be varying aspects of the representational vehicle which serve to map to different contents, in the way that the singular term and predicate in a sentence can be varied, or that the spatial relations represented on a map vary with where points are placed. Nor are they the basis for relations that carry semantic content. Two dances performed on different occasions are temporally related—they are separated by time t. But the temporal separation t is not something consumer bees are sensitive to or that otherwise enters into downstream computations. Similarly for the places at which two dances are performed (within the hive, or when the hive moves).

[4] Lee et al. (2023) put organization—what they call ‘analog mirroring’—at the centre of their account of what it is to be analogue.

vehicles then has semantic significance in the sense that the function-in-intension taking vehicles to contents applies to a vehicle property in a way that ensures similar vehicles represent similar contents.

3 Organized mental representations

We are building up to being able to discuss the merits of representational organization (§4). An important preliminary, though—in fact, a large part of the philosophical task—is to characterise the phenomenon precisely. For simplicity, I started with examples from overt signalling systems. This section elaborates on the phenomenon in the mental realm.

A rich source of examples derives from cases where the brain is standardly described as performing analogue computations. For instance, analogue magnitude representations of numerosity are relied on to perform many useful computations (Nieder, 2016). In these cases, internal magnitudes like neural firing rate are used to represent external magnitudes like the number of items in a set (Peacocke, 2019b). The way internal magnitudes interact realises a computation over the represented magnitudes (Peacocke, 2019a, pp. 52–59). For example, the firing rate of two neurons can be aggregated by a third neuron so that the firing rate of the third reflects the sum of the first two. That pattern of internal transitions can be used to calculate addition. The third neuron will represent the sum of the numbers of objects represented by the first two neurons.[5] These kinds of processes over neural vehicles are thought to be the way analogue magnitude representations are used to calculate sum, difference, product, etc. (McCrink & Spelke, 2010).

In our simple neural addition circuit, three families of organized representations interact to perform a computation. But it is not a case of structural representation. In a structural representation, a relation on representational vehicles carries content. It represents a relation on the entities represented by the relata. (E.g. spatial separation between points on a map represents the distance between the locations represented by the points.) In the neural circuit for addition, there are of course various relations between neuron1 and neuron2. There is a relation between their firing rates, for example. But this relation does not act as a bearer of content. It does not even figure in a vehicle-level causal description of how processing unfolds. The special science generalisations capturing the neural dynamics take the firing rates of neurons 1 and 2 as input and produce a firing rate of neuron 3 as output. It is not a case where the relation between neurons 1 and 2 is acting as a representation of the entities represented by neurons 1 and 2 (cp. the definition of structural representation, Sect. 2).

By contrast, the spatial cognitive map realised by neural structures in the medial temporal lobe is, plausibly, a case of structural representation. This case has been described extensively elsewhere (Grieves & Jeffery, 2017; Shea, 2018, pp. 113–116), so here I will just mention some key aspects. When moving around, activity of place cells in the hippocampus correlates with current location. The same system is used

[5] The vehicle to content mapping is unlikely to be completely linear, but treating it as such is a harmless simplification for present purposes.

offline to calculate potential routes. Offline activity proceeds through chains of place cells, chains that correspond to potential routes through the environment. Different chains of place cell firing can be compared. Picking the shortest chain of firing between start and end place cells is a way of picking the shortest route between the two corresponding locations.[6] It is the chain of activated place cells that constitutes a structural representation. (We are not here looking at the standing structure of synaptic connections in virtue of which place cells are wired together as they are.) Suppose place cell PC1 represents location L1 and place cell PC2 represents location L2. In performing an offline route calculation, the activation of PC2 by PC1 represents that L2 is near to L1. The co-activation relation carries content. It represents spatial proximity. PC1-activating-PC2 is a vehicle of content. Its content is *L1 is near to L2*. So this is a case of structural representation.

The activation of an analogue magnitude representation of numerosity is not a structural representation. It is the tokening of one out of a family of representations (which display organization). The content of a family of analogue magnitude representations (e.g. of a neuron that can fire at different rates) is specified by a systematic function, at the level of the whole family, such that the number of items represented equals firing rate multiplied by a constant. This semantic mapping has the effect that similar vehicles (firing rates) represent similar contents (number of items represented). The similarity relation does not carry content—it is not a vehicle of content.[7] Indeed, in the cases we have described, no two representations within the same family are tokened at once (or otherwise tokened in a relation to one another that plays a computational role). The family displays organization, but not in virtue of compositional semantic structure. It is a mere organized representational system. The place cell representation of spatial proximity is organized, of course, but it is not a mere organized representational system. It is a structural representation.

4 Merits of representational organization

Just as the bee dance differs from vervet alarm calls, analogue magnitude representations of numerosity differ from a piecemeal way of representing number (having one detector for 3 and another for 11, for example). The fact that a family of representations form an organized representational system confers a number of benefits. This section argues that a merit that has been under-appreciated is that having representations in organized families is computationally useful.

What I have been calling representational organization has been recognized as beneficial even by those who accept that mere organization falls short of structural representation (Shea, 2014, 2018, p. 128; cf. Artiga forthcoming). First, it is more efficient to implement. A piecemeal implementation, with a different mechanism for each quantity, would be unwieldy. Second, the system extends non-accidentally to new

[6] There are many different proposals for the way route-finding computations make use of the co-activation structure over place cells, including processes that take place in parallel across the whole array, e.g. Khajeh-Alijani et al. (2015).

[7] I am using 'vehicle' inclusively to cover, not just particulars, but also properties and relations that have semantic values.

cases. With the bee dance, even if nectar 400 m away happened never to have occurred when the system evolved, if now a current worker bee discovers nectar at 400 m and performs a four-waggle dance, the mechanism in consumer bees will cause them to do the right thing. That is no accident. The mechanism has evolved to effect a certain simple transformation from number of waggles to flight distance. The same is true of the ability of the analogue magnitude system to represent new quantities (within its overall range).

A third benefit of an organized representational system is that it is error-tolerant. Noise from external and internal sources will affect the firing rate. But similar rates represent similar numerosities. In many circumstances, the calculations or behaviour that result will still be accurate enough to be useful. For example, when comparing two collections in order to pick the largest (Barth et al., 2006), noise that somewhat perturbs an analogue magnitude representation will not often cause the agent to make the wrong choice. Vehicles that are nearby in the neural similarity space (nearby firing rates) represent contents that are nearby in the represented similarity space (numerosities).

These three merits of representational organization are also reasons why the semantic mapping that accurately portrays the way vehicles carry contents is one that applies at the level of the family of vehicles, consisting of a rule or systematic function that maps similar vehicles to similar contents. That is to say, the fact that the system counts as a case of representational organization is closely connected to the reason why it displays these merits. The causal dynamics of the system are described at the level of the determinable: variations in numerosity cause variations in firing rate; variations in the number of waggles cause variations in the distance flown by outgoing bees. The same is true of the way neural firing rates interact to implement addition. Non-semantically, those interactions are described in terms of a causal relation that applies in a uniform way across the families of potential neural firing rates. A theory of content will, then, plausibly ascribe contents to vehicles at the level of the determinable, in a systematic way across vehicles in the family.

To justify that claim properly would require a too-long diversion into theories of content. What makes it plausible is a commitment shared by many theories of content: the way vehicles interact causally, and the ways they are causally connected to the world at input and output, are important to fixing their content (Block, 1986; Dretske, 1981, 1986; Fodor, 1990; Millikan, 1984; Neander, 2017; Papineau, 1984; Ramsey, 2007; Shea, 2018). Since these causal relations are captured at the level of vehicle families, the semantic mapping delivered by a theory of content will typically end up applying systematically at the level of the family. It is these causal relations between vehicle families *inter se*, and between vehicle families and the world, that make the system have the three merits just mentioned (being efficient to implement, error-tolerant and extensible non-accidentally to new cases).

We have finally arrived at the central claim of the paper: that there is a further merit of representational organization which has not been recognized as such in the existing literature. It is also something that arises from the way different families interact in internal processing. My claim is that representational organization is often part of a processing structure that is made use of computationally. In our analogue addition example, representations from three families interact so as to perform addition. Neuron

3 represents the sum of the numerosities represented by neurons 1 and 2. The computational process (performing addition) is achieved by the interaction of the firing rates of the three neurons. Neuron 3 is configured so as to fire at a rate that is the sum of the rates of neuron 1 and neuron 2. That processing structure is computationally useful. It can be—and is—used to calculate addition. The way it is used trades on the fact that similar firing rates correspond to similar numerosities. That is a further reason why the semantic mappings characterising the contents carried by neurons 1, 2 and 3 are each such that similar firing rates map to similar numerosities—why they qualify as cases of representational organization.

The fundamental insight of the representational theory of mind (RTM) is that transitions between representational vehicles can be configured such that the transitions respect semantic content. In many circumstances, to behave appropriately the organism must combine and weigh various sources of information in order to work out how and when to act, given the current circumstances and environment. One good way to do that is to manipulate internal representations, vehicles that carry content and whose interactions respect those contents. The way vehicles interact forms a processing structure. In the example involving analogue addition, interactions between vehicles forms a useful processing structure because it enables the organism to add quantities together. In this way, three organized representational systems (three families representing numerosity) interact in a way that forms a computationally-useful processing structure.

The idea that vehicle transitions are configured so as to be faithful to semantic contents is a staple of representational theories of behaviour and of philosophical theories of representation. An example of the former is the architecture of convolutional neural networks doing image classification (LeCun et al., 2015). The architecture is set up so that the network learns common local transformations applied in the same way to all parts of the visual array. That processing disposition reflects the transitional invariance of images. The latter is described by Cummins (1991) and Ramsey (2007, p. 73) as the Tower Bridge picture: physical transformations from one representation to another march in step with computations on the contents they stand for. A symbolic system set up to do logical proofs fits this description. When it performs a modus ponens inference, the transition performed on representational vehicles of a certain form is faithful to their contents. The point I want to highlight is that, in families of representations displaying mere organization, the similarity relations can be a crucial part of the way transitions are configured so as to be faithful to content.

Theorists have tended to elide computational processing structure with structural representation (Ramsey, 2007, pp. 77–92; Gallistel, 1990, pp. 15–30). As we have seen, the two phenomena are in fact distinct. In a processing structure, transitions between representations mirror meaningful transitions between contents. The vehicle-level transition that aggregates firing rates mirrors the content-level transition from two quantities at input to their sum at output. In a structural representation, the structure of a complex representation itself has representational content: a relation between parts of the complex vehicle represents a relation between the entities represented. Where a similarity relation within a family of vehicles corresponds to a similarity relation between corresponding contents, vehicles from that family can readily interact with vehicles from other families so as to form processing structures. That is not a form of structural representation but the basis of a processing structure.

Ramsey distinguishes between structural representations and what he calls IO representations. Organized representations are closer to his IO representations (Ramsey, 2007, pp. 68–77). Similarly, I have argued that representational content (in subpersonal cases) can be based on one of two kinds of exploitable relation: exploitable correlational information or exploitable structural correspondence (Shea, 2018). Organized representations typically fall in the former camp. Representations based on exploitable correlations need not display organization, as we have seen in the case of the macaque signalling system, but many do. They are then based on 'exploitable correlational information carried by a range of states' (Shea, 2018, p. 78). Similarly with receptor representations (Ramsey, 2007, pp. 118–150; Artiga, 2022).[8] Receptor representations divide into organized and nominal representations, with most canonical examples showing organization. Unlike structural representations, organized receptor representations display mere organization.

So far I have argued that organized representations are computationally useful because representations from different families can be put together into a processing structure. Processing structure is a non-semantic characterisation of the causal dynamics of the system, of how states of the system unfold. These are non-semantic and not rational or content-based characterisations. They must capture an aspect of the causal dynamics. For example, the spikes emitted by two neurons may be more or less similar in terms of the exact timing at which spikes are emitted (synchrony). That kind of similarity will be causally irrelevant unless there are downstream processes that can detect and respond to synchrony (which there may well be). Vehicle similarities are only part of a processing structure if they are a real feature of the causal dynamics.

Many different kinds of computation can be implemented in suitable processing structures (Egan, 2014; Shagrir, 2001; Shea, 2013b). One much relied-on neural configuration is the neural accumulator. Often a neural accumulator comes with a bound which, when reached, triggers some categorical output or behaviour (Beck et al., 2008). For example, the probabilistic population coding scheme described by Pouget et al. (2003) is good for accumulating sensory evidence of all kinds and using it in the service of a binary decision. A different neural scheme is suitable for computing divisive normalisation (Carandini & Heeger, 2012). The emerging field of computational cognitive neuroscience is concerned with enumerating and characterising the kinds of computations that transformations over neurons (biological or artificial) are suited to implementing. The vast majority of this work makes use of (what I have been calling) organized representations. The focus is on the way organized representations can be combined into processing structures, and on the kind of computations those processing structures are suited to implementing.

I am not here arguing that the content-involving description of how processing unfolds constitutes a further level of content (e.g. neo-Fregean content). However, the contentful description of how computations unfold does capture something of what Fregean contents were supposed to give us. It tells us something about the

8 Artiga argues, *contra* Ramsey, that receptors or detectors count as representations within Ramsey's class of IO representations.

way the system is representing the states of affairs that it represents. Contrast the analogue magnitude representation of numerosity in our three-neuron addition circuit with a (hypothetical) pointwise way of representing numerosity. Suppose the latter just represents 3, 7 and 11, with a separate vehicle for each and no relevant similarity structure.[9] (Embellishing, we could suppose there are different behaviours appropriate to flowers with 3, 7 or 11 petals and the organism uses detectors for each so as to condition its behaviour accordingly.) There, the semantic mapping from vehicles to contents would be point-wise, like a look-up table. That contrasts with the way quantities are represented and computed with in the analogue magnitude system that adds firing rates to compute addition.

There is a three-way distinction here between structural representation, a mere organized representational system and a nominal representational system. In a structural representation, a relation between parts of a complex representational vehicle carries semantic content. The relation will figure in a causal description of the way processing unfolds. To put it another way, there will be a special science law or special science generalisation, that captures the system's dynamics, in which the vehicle relation will figure causally. In the mere organized representational system, the relevant vehicle property (e.g. firing rate) figures in the causal dynamics, and the special science law that captures the way processing unfolds will treat similar vehicles in similar ways (e.g. two firing rates are aggregated into a third), but the relation between vehicles in a family does not figure as such in the causal dynamics. A computation in which similar vehicles map to similar contents makes use of this processing structure. In a nominal representational system, any similarities between vehicles are irrelevant. They are not relevant to the way processing unfolds or they do not figure in the correct semantic mapping from vehicles to contents.

This section has argued that, in addition to previously-recognised merits, organization is beneficial because organized representations can be arranged into processing structures that perform useful computations. This claim faces an obvious objection: isn't the property that is supposedly useful here actually so liberal as to be trivial, to have no distinctive purchase? I answer that objection in the final section (Sect. 6), but first I illustrate the idea of computational usefulness further by contrasting the way organized representations function with representations that are 'symbolic' (in one sense of that term).

5 Symbols vs. non-symbolic representations

This section illustrates the relation between organization and processing by looking at artificial neural networks. There, a distinction is made between organized representations and what are often called 'symbolic' representations. I have said that similarities are irrelevant in a nominal representational system, but that is too quick. A degenerate form of similarity and difference is relevant. Two token representations can be of the same type or of a different type. This kind of similarity and difference is relevant to

[9] The mantis shrimp plausibly has a different detector for each of a dozen different colours but no common similarity system: Thoen et al. (2014), quoted by Lau (2022, p. 207).

whether we have an instance of the same representation again. It individuates the bearers of content. No other kind of similarity and difference has semantic significance. There is no graded form of similarity that connects vehicles into a family.

In fact, one way of drawing the contrast between symbols and non-symbolic representations is formulated in just these terms. 'Symbolic' is used to mean several different things: that a representation is structured or compositional; that operations on the representation are rule-based rather than associative; that the system is model based, representing how things are related in the world (e.g. causally or spatially), rather than model free. One use of 'symbolic' is just what we need: representations that do not fall into semantically related families. Thus, for LeCun et al. (2015):

> 'an instance of a symbol is something for which the only property is that it is either identical or non-identical to other symbol instances. It has no internal structure that is relevant to its use' (LeCun et al., 2015, p. 441)

Nominal representations have no compositional structure and do not display organization. They are symbolic in this sense. However, nominal representations have complete contents at the level of a correctness condition or a satisfaction condition (e.g. *the British are coming by sea*). Symbols are usually taken to be sub-propositional representational constituents, representations that compose with other representations, for instance in a sentence or a thought, in order to form a complex representation which has a correctness condition or a satisfaction condition (e.g. 'Layla runs'). These constituents, too, can be symbolic in the sense of LeCun et al. (2015). Paradigm names do not come in organized families. Any similarities between them have no semantic significance. The only semantically relevant property is whether we have an instance of the same vehicle again, or a different vehicle.

Here are some names: Pam, Sam, Zadie, Zadie. The first three are tokens of different symbols. The last two are tokens of the same symbol. The first two are similar, but their similarity is not relevant to their content. The firing rate of an analogue magnitude representation represents numerosity. There, similarities in firing rate correspond to similar contents and are relevant to the way they are processed. LeCun et al.'s point is that representations in deep neural networks are usually like that. Similarities between representations (distributed patterns of activation) are computationally significant. The point is not just that the vehicles vary along one or more dimensions. That is true of letters and words. It is that those similarities and differences are part of a processing structure that makes them semantically significant. That is to say, representations in deep neural networks often display organization. Of course names causally interact with other names in reasoning (as they do with symbols of other types), but similarities and differences between the representational vehicles—for example that 'Pam' is more similar to 'Sam' than to 'Zadie'—do not make a semantically-significant difference to how they are processed (nor to what is represented).

Representations realized by patterns of activation in an artificial neural network are typically organized and non-symbolic. (That is the genesis of the somewhat unhelpful symbolic / subsymbolic distinction.) However, there may in fact be cases where patterns of activation in a layer of a neural network count as symbols by LeCun et al.'s definition. Transitions between layers can be non-linear. Suppose downstream processing serves to divide up state space into ten discrete regions (as in Fig. 1). The network

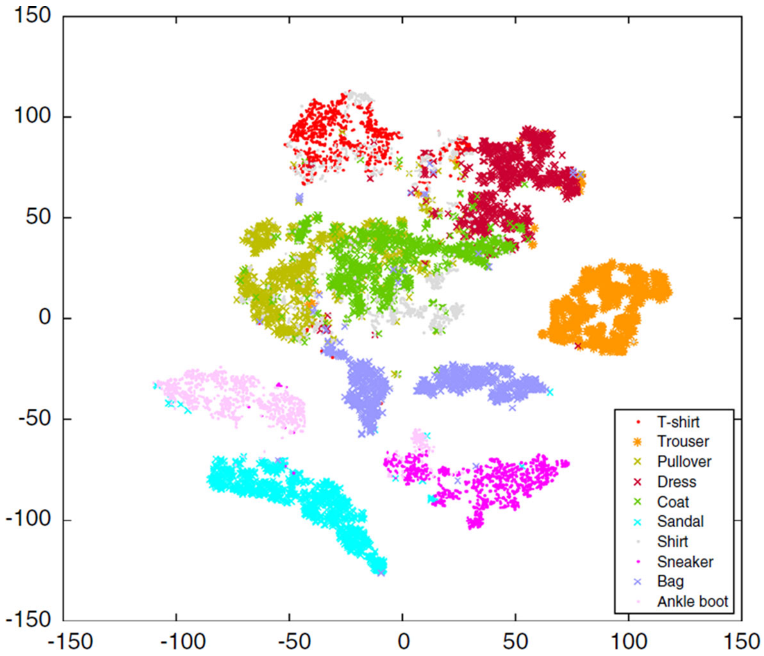

Fig. 1 The state space of the third hidden layer of a deep neural network trained to classify greyscale images of items of clothing, projected into a two dimensional visualisation using the technique t-SNE. From Qiao et al. (2020)

may do something different in response to each region, with variations within a region making no difference to what happens downstream. Then downstream processing has effectively carved boundaries in the state space. If so, the regions or clusters will be symbols (Azhar, 2016; Shea, 2007). Although different tokens in a cluster are more or less similar to one another, that is irrelevant. The only property that matters is being an instance of the same symbol or of a different symbol. Two token patterns of activation either fall in the same cluster or different clusters.

That form of processing would make for fully symbolic representations in a neural network layer. Intermediate cases are also possible. Suppose operations on state space lead to discontinuous and unrelated results across boundaries in state space, as before. Which output points in Cluster1 should be mapped to is irrelevant to the question of which output points in Cluster2 should be mapped to (Shea, 2023). It could be that, nevertheless, relations between points within a cluster are semantically significant. Perhaps clusters distinguish between different categories; but the arrangement of points within a cluster is informative about how large a particular individual is. If downstream processing relies on this dimension of variation, the placement of a point along this dimension within a cluster would represent that the individual has a certain size. Similarities between points within a cluster would have semantic significance. At the same time, the distinction between different clusters, with different clusters representing different categories, would be symbolic.

Another kind of intermediate case is where relations between regions have semantic significance. Maybe there is a dimension in the overall state space that varies with perceived dominance (Olivola et al., 2014). Discrete regions representing different individuals could be ordered along this dimension. Individuals at one end are represented as more dominant, those at the other end more subordinate. If this dimension were relied on computationally, then points in state space would come to represent dominance properties, as well as representing the individuals bearing those properties. With respect to which individual is represented, the only vehicle question is same/different (symbolic). With respect to which property is represented (highly dominant, moderately dominant, etc.) the arrangement of points along the dominance dimension is semantically significant.

In short, we can indeed separate semantically-significant vehicle properties into two classes: those for which same-different is the only relevant dimension and those for which some dimension of variation across a family of vehicles has semantic significance. Intermediate cases arise where, for the same representational token, some vehicle properties fall in one class and some in the other.

6 What makes a processing structure computationally useable?

Section 4 argued that organized representations can be put together into processing structures that are computationally useful. That claim, however, faces an immediate worry about liberality. Can't a pattern of transformations between vehicles perform more or less any transformation on contents, provided the way contents map to vehicles is suitably arranged? Is there really any bite to the claim that transformations between vehicles are suited to performing some restricted class of computations (rather than just any way arbitrary of transforming contents)? This section addresses that worry. There is not scope here to give a full treatment of liberality objections to theories of content. What I do want to show, however, is that resources that a theory of content can rely on to cut down on problematic liberality are applicable to representational organization. Thus, the claim in Sect. 4 about useful processing structures is not trivial.

I proceed by considering what it takes, in general, for a processing structure to be computationally useful. I divide that question in two: what does it take for a processing structure to be computationally useable; and to be computationally used? For vehicles of any kind—nominal representations, structural representations, symbols—the way vehicles are divided into types that are recognised by a computational system must of course be useful in its workings in some way. This paper is focusing on a special case of that, where vehicle types display similarity and that similarity is relied on. The point of this section is to show that the claim that similarities between vehicle types are used or relied on is not trivial.

A processing structure is a set of dispositions for making causal transitions between potential vehicles of content. We are focusing on the case where those transitions take place between vehicles displaying organization. There is surely an extremely large number of computational transitions that could be implemented by any given collection of vehicles, provided contents are assigned to vehicles appropriately. The idea that transitions can be used to implement computational steps thus faces a liberality objection that threatens to rob it of any explanatory purchase. As a way in, recall the liberality objection levelled against reliance on structural correspondence in a theory of content. That is a different objection, but close enough to be instructive. The objection was that the notion of correspondence or homomorphism is too liberal to determine the content of a structural representation (Cummins, 1996; Ramsey, 2007; Shea, 2013a). Given enough representational vehicles, there is always some relation-preserving mapping between vehicles and referents that preserves any relation you like on the referents. (And where there is one, there will be many.)

The same kind of problem arises in respect of processing steps. Recall the Tower Bridge picture (Cummins, 1991; Ramsey, 2007, p. 73): there is a non-semantic, vehicle-level process in which tokening of a vehicle or vehicles causes the tokening of other vehicles, and that process marches in step with a content-level computational transition on the contents the vehicles represent. The worry is that the same internal process will correspond to very many different situations in the world and transitions between contents (Fodor, 1981, pp. 207–208). Theories that appeal to an isomorphism between internal computational steps and structures in the world (e.g. Gallistel's 'functioning isomorphism' (1990, pp. 15–33)) need to say why one of the very many isomorphisms that exist is privileged and plays the content-constituting role. Indeed, theorists who fail to distinguish between structural representation and processing organisation treat the two forms of liberality as the same problem (Ramsey, 2007, p. 93). If there is no restriction on how vehicle types map to contents, then any mechanism making transitions between putative vehicles is suited to calculating an extremely large number of functions. Of course, most of these mappings are useless for practical computational purposes. They would assign contents to vehicles in an implausible, gerrymandered fashion. The problem is to say why these are ruled out.

A triviality problem also arises in the theory of physical computation. The worry there is that any sufficiently complex system—a wall, a bucket of water—could implement any arbitrary computation (Searle, 1990; Chalmers, 1996; Godfrey-Smith, 2009; Sprevak, 2010; Coelho Mollo 2019; Shagrir, 2020). The issues may play out somewhat differently when the task is to define physical computation, rather than with our question about the nature of representation (unless computation itself must be individuated semantically Shagrir (2020)). However, the structure of the problem is the same. Without substantial constraints, liberality threatens. That problem can be addressed, in theories of physical computation, by appealing to constraints based on causation, function and/or mechanistic explanation (Piccinini, 2015; Coelho Mollo 2019, 2021a, 2021b; Shagrir, 2022). The particular challenge for the claim made in the present paper is to show that the notion of organization—use of similarity across a family of vehicles—does not face its own, more insurmountable liberality objection. Liberality is problematic, from our point of view, because it drains the substance out of the idea that some processing structures are computationally useful—i.e. useable for performing some particular set of computations.

The idea that a form of processing is useable computationally is commonplace in cognitive scientific practice. We have seen some examples already: adding or subtracting magnitudes; aggregating probabilistic evidence and calculating a maximum likelihood estimate (Pouget et al., 2003); computing divisive normalisation (Carandini & Heeger, 2012). Indeed, one way of reconciling different neuroimaging findings about the functions of different neural areas is to categorise them not by the domains where they happen to be deployed (e.g. face processing, inanimate objects, etc.) but in terms of the computations that different brain areas are suited to performing (Anderson, 2016; Poldrack, 2010). Computational cognitive neuroscience is actively engaged in describing the different kinds of operations that can readily be performed in neural codes, as well as studying how these computational circuits are deployed to perform a diverse range of tasks. For example, recent neuroscientific evidence suggests that a hexagonal coding scheme, first described in the medial temporal lobe in connection with spatial navigation, is in fact widely deployed in representing more abstract properties (Behrens et al., 2018; Bellmund et al., 2016; Constantinescu et al., 2016; Doeller et al., 2010; Whittington et al., 2020). This rich seam of empirical work strongly suggests that there is something substantial to the idea that certain patterns of neural processing are computationally useful. The liberality objection does not appear to impede scientific practice.

I won't be able to give a full account here, but I do want to lay out the ingredients that I think will go into a theory of what makes a processing structure computationally useable. The first constraint we have seen already. The processing structure has to be a real feature of the causal dynamics of the system. With representational vehicles in organized families, that gave us the requirement that the way vehicles in an organized family interact causally with other vehicles is described by a determinable property such that similar values have similar causes and effects. (Contrast symbolic representations, where similarity between representations does not play a role in the content-relevant causal dynamics.)

Where the first ingredient concerns the nature of the internal transitions, the second ingredient concerns vehicle-world relations. I have pointed to two kinds of exploitable relations between vehicles and the world that are relevant to content-determination: correlational information and structural correspondence (Shea, 2018). (There may be others.) Exploitable correlational information can be carried in a systematic way by a family of vehicles (Shea, 2018, p. 78). For example, the firing rate of a neuron may correlate with the numerosity of an array of objects. Exploitable structural correspondence by its nature concerns a family of interrelated representations. Both notions define a privileged relation or relations on a family of vehicles: those that figure on the vehicle end of the exploitable relation. Where firing rate correlates with numerosity, vehicles are organised with respect to similarity and difference in firing rate.

Those exploitable relations give us a collection of candidate vehicle properties (there may be many) on both ends of an internal transition. In our toy example with two input neurons and one output neuron we home in on similarities in firing rate at input and output. Firing rates are transformed by simple summation. The third ingredient is to ascertain what functions those transitions are capable of instantiating, given potential contents at input and output. Putative contents are given by the other end of the respective exploitable relations. If Neuron1 and Neuron2 correlate with the numerosity of two arrays of objects, then the transition to Neuron3 is suited to calculating their sum. Implementing that transition explains why Neuron3 comes to correlate with the sum of the two arrays of objects. If Neuron1 and Neuron2 also correlate with the surface area of arrays of objects presented at input, then the transition to Neuron3 is also suited to calculating the total surface area of all the objects presented.

And so on, for each of the exploitable relations which the three vehicle families enter into.

These considerations will not take us to unique content assignments. Far from it. We will still need to appeal to other constraints. Importantly, very substantial constraints derive from the task the system is called on to perform (Egan, 2014; Piccinini, 2022; Shea, 2018). But we are no longer in the realm of complete liberality. We ask: given exploitable ways that putative contents attach to states, what functions is a given physical transition capable of implementing? Only some similarity relations on families of vehicles are candidates to figure in the transitions, and only some potential contents at input and output are candidates to figure in the semantic transformation implemented by that transition. A processing structure is collection of transitions between potential representational vehicles. It is useable for implementing only computations between conditions in which those vehicles stand in exploitable relations.

This is not yet a full theory of what it is for processing organisation to be computationally useful. It just highlights the factors that will be relevant. But it does show how the liberality objection for representational organization can be addressed. Computational usability is a substantial enough notion to do significant explanatory work. What it is for a useable processing structure to actually be used is another deep question. That will ultimately depend on the final true theory of representational content (which is beyond the scope of this paper). However, a central commitment of RTM is that transitions between representations are faithful to representational content (in some circumstances). So it is very likely to turn out that the correct theory of content must select amongst the computationally useable transitions that a given processing structure exhibits.

7 Conclusion

A system of representations that displays mere organization may not amount to a structural representation, as some have thought, but representational organization is important nonetheless. Representations from different organized families can interact to perform computations that are useful to the organism in order to achieve some outcome or perform some task. In such cases, the organization (the mapping from similar vehicles to similar contents) is made use of to perform the computation. A general feature of the representational theory of mind is that processing over vehicles is configured in such a way as to implement computations that are useful to the organism. An under-appreciated merit of organization—which at the same time is a reason that families of representations are rightly characterised as displaying organization—is that organized representations can form part of a computationally useful processing structure.

Acknowledgements The author would like to thank Marc Artiga, Patrick Butlin, Sam Clarke, Dimitri Coelho Mollo, Will Davies, Peter Godfrey-Smith, David Papineau, Samir Okasha and Christopher Peacocke for comments on previous drafts; and audiences at the IP lab meeting, Oxford Mind work-in-progress group and Philosophy of Science Association for constructive comments.

Funding This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 681422 (MetCogCon).

Declarations

Conflict of interest The author has no competing interests to declare that are relevant to the content of this article.

References

Anderson, M. L. (2016). Précis of after phrenology: Neural reuse and the interactive brain. *Behavioral and Brain Sciences, 39*, e120.

Artiga, M. (2022). Strong liberal representationalism. *Phenomenology and the Cognitive Sciences, 21*(3), 645–667.

Artiga, M. forthcoming. Understanding structural representations. *British Journal for the Philosophy of Science*.

Azhar, F. (2016). Polytopes as vehicles of informational content in feedforward neural networks. *Philosophical Psychology*. https://doi.org/10.1080/09515089.2016.1142070

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98*(3), 199–222. https://doi.org/10.1016/j.cognition.2004.09.011. https://www.ncbi.nlm.nih.gov/pubmed/15876429.

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., & Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron, 60*(6), 1142–1152.

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron, 100*(2), 490–509.

Bellmund, J. L. S., Deuker, L., Navarro Schröder, T., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *eLife, 5*, e17089.

Block, N. (1986). Advertisement for a semantics for psychology. In P. A. French, T. Uehling, & H. Wettstein (Eds.), *Midwest studies in philosophy, vol. 10: Studies in the philosophy of mind* (pp. 615–678). Minneapolis: University of Minnesota Press.

Camp, E. (2007). Thinking with maps. *Philosophical Perspectives, 21*(1), 145–182.

Camp, E. (2021). Representation in cognitive science by Nicholas Shea: Organization and structure in the service of systematicity and productivity. *Studies in History and Philosophy of Science*.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51–62.

Chalmers, D. J. (1996). Does a rock implement every finite-state automaton? *Synthese, 108*(3), 309–333.

Coelho Mollo, D. (2019). Are there teleological functions to compute? *Philosophy of Science, 86*(3), 431–452.

Coelho Mollo, D. (2021a). Against computational perspectivalism. *The British Journal for the Philosophy of Science, 72*(4), 1129–1153.

Coelho Mollo, D. (2021b). Why go for a computation-based approach to cognitive representation. *Synthese, 199*(3–4), 6875–6895.

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science, 352*(6292), 1464–1468.
Cummins, R. (1991). The role of representation in connectionist explanations of cognitive capacities. In W. Ramsey, S. Stich, & D. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 91–114). Lawrence Erlbaum.
Cummins, R. (1996). *Representations, targets, and attitudes*. MIT Press.
Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature, 463*(7281), 657.
Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press.
Dretske, F. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief: Form, content and function.* OUP.
Egan, F. (2014). How to think about mental content. *Philosophical Studies, 170*(1), 115–135.
Fodor, J. A. (1981). *RePresentations: Philosophical essays on the foundations of cognitive science*. MIT Press.
Fodor, J. A. (1990). *A theory of content and other essays*. MIT Press.
Gallistel, C. R. (1990). *The organization of learning*. MIT Press.
Godfrey-Smith, P. (2009). Triviality arguments against functionalism. *Philosophical Studies, 145*(2), 273–295.
Godfrey-Smith, P. (2017). Senders, receivers, and symbolic artifacts. *Biological Theory, 12*(4), 275–286.
Goodman, N. (1968). *Languages of art: An approach to a theory of symbols*. Hackett Publishing.
Grieves, R. M., & Jeffery, K. J. (2017). The representation of space in the brain. *Behavioural Processes, 135*, 113–131.
Khajeh-Alijani, A., Urbanczik, R., & Senn, W. (2015). Scale-free navigational planning by neuronal traveling waves. *PLoS ONE, 10*(7), e0127269.
Lande, K. J. (2021). Mental structures. *Noûs, 55*(3), 649–677. https://doi.org/10.1111/nous.12324
Lau, H. (2022). *In consciousness we trust*. OUP.
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.
Lee, A. Y., Myers, J., & Rabin, G. O. (2023). The structure of analogue representation. *Nous, 57*(1), 209–237.
Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies, 155*, 117–131. https://doi.org/10.1007/s11098-010-9562-8
McCrink, K., & Spelke, E. S. (2010). Core multiplication in childhood. *Cognition, 116*(2), 204–216.
Millikan, R. G. (1984). *Language, thought and other biological categories*. MIT Press.
Neander, K. (2017). *A mark of the mental: In defense of informational teleosemantics*. MIT Press.
Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience, 17*(6), 366–382.
Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18*(11), 566–570.
Papineau, D. (1984). Representation and explanation. *Philosophy of Science, 51*, 550–572.
Peacocke, C. (2019a). *The primacy of metaphysics*. OUP.
Peacocke, C. (2019b). Spatial perception, magnitudes, and analogue representation. In T. Cheng, O. Deroy, & C. Spence (Eds.), *Spatial senses: Philosophy of perception in an age of science* (pp. 49–69). Routledge.
Piccinini, G. (2015). *Physical computation: A mechanistic account*. OUP.
Piccinini, G. (2022). Situated neural representations: Solving the problems of content. *Frontiers in Neurorobotics, 16*, 846979.
Planer, R. J., & Godfrey-Smith, P. (2021). Communication and representation understood as sender–receiver coordination. *Mind & Language*, *36*(5), 750–770.
Poldrack, R. A. (2010). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science, 5*(6), 753–761.
Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annual Review of Neuroscience, 26*(1), 381–410.
Qiao, C., Gao, B., & Shi, Y. (2020). SRS-DNN: A deep neural network with strengthening response sparsity. *Neural Computing and Applications, 32*(12), 8127–8142.
Ramsey, W. (2007). *Representation Reconsidered*. Cambridge University Press.
Searle, J. R. (1990). Is the brain a digital computer? *Proceedings and addresses of the American Philosophical Association*.
Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*(4), 1070–1094.
Shagrir, O. (2001). Content, computation and externalism. *Mind, 110*(438), 369–400.

Shagrir, O. (2012). Structural representations and the brain. *British Journal for the Philosophy of Science, 63*(3), 519–545.
Shagrir, O. (2020). In defense of the semantic view of computation. *Synthese, 197*, 4083–4108.
Shagrir, O. (2022). *The nature of physical computation*. Oxford University Press.
Shea, N. (2007). Content and its vehicles in connectionist systems. *Mind & Language, 22*(3), 246–269.
Shea, N. (2013a). Millikan's isomorphism requirement. In D. Ryder, J. Kingsbury, & K. Williford (Eds.), *Millikan and her critics* (pp. 63–80). Wiley-Blackwell.
Shea, N. (2013b). Naturalising representational content. *Philosophy Compass, 8*(5), 496–509.
Shea, N. (2014). Exploited isomorphism and structural representation. *Proceedings of the Aristotelian Society, 64*(2), 123–144.
Shea, N. (2018). *Representation in cognitive science*. Oxford University Press.
Shea, N. (2023). Moving beyond content-specific computation in artificial neural networks. *Mind & Language, 38*, 156–177. https://doi.org/10.1111/mila.12387
Sprevak, M. (2010). Computation, individuation, and the received view on representation. *Studies in History and Philosophy of Science, 41*, 260–270.
Swoyer, C. (1991). Structural representation and surrogative reasoning. *Synthese, 87*(3), 449–508.
Thoen, H. H., How, M. J., Chiou, T.-H., & Marshall, J. (2014). A different form of color vision in mantis shrimp. *Science, 343*(6169), 411–413.
Whittington, J. C. R., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. J. (2020). The Tolman-Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell, 183*(5), 1249–1263.