

Book Reviews

Nicholas Shea, Representation in Cognitive Science, Oxford: Oxford University Press, 2018, 304 pp.

Nicholas Shea's excellent book *Representation in Cognitive Science* is the most recent attempt to provide a naturalized theory of representational content, that is, an attempt to explain how representations, understood as content-baring physical particulars, acquire their content, using non-semantic, nonmental and non-normative descriptions (11). The account offered in his book is a continuation of the two most influential naturalistic approaches to mental content—teleosemantic approach (Millikan, Papineau) and informational approach (Dretske, Neander). His account relies on standard resources of these theories—most important of those being the notions of function, information and correspondence, but develops an original understanding of how these notions converge in the metaphysical determination of representational content. In doing so it also relies on the work by Peter Godfrey-Smith. Shea's book is abundant with case studies ranging from studies of simple artificial and animal systems to those more complex but sufficiently understood by cognitive neuroscience, such as the spatial-navigation system in the rat's hippocampus. He uses them to develop his account, but also to test it in relation to standard objections made against teleosemantic approaches, such as the problem of indeterminacy of function or the infamous swampman objection.

The book is divided in three parts. The first one is introductory and offers a framework for the account developed in the rest of the book (chapters 1, 2). Part two presents his account of representational content, centering on its three main ingredients—task function, correlational information and structural correspondence (chapters 3, 4, 5). Part three answers aforementioned indeterminacy and swampman objections, offers an account of the distinction between descriptive and directive representation and concludes with several considerations concerning the explanatory role of content and content of higher personal, conscious states (chapters 6, 7, 8). Lack of space makes it impossible to present every chapter separately. It also makes it impossible to present the numerous case studies and important subdiscussions which comprise Shea's book. While admitting its insufficiency, the review will be focused on presenting several key aspects of Shea's book.

The ingenuity of Shea's account consists in its ability to reconcile resources of different, competing theories of mental content. The main idea behind his theory of representational content is that content arises out of the convergence of three elements—functions, in his account called *task*

functions, exploitable relations and *internal mechanism*—each of which is necessary for content. The core idea is that representational vehicles acquire content by bearing certain relations to the environment, relations that are exploited by an internal mechanism in order to perform a certain function. What is immediately noticeable from this, somewhat crude, definition is that on Shea's account, content is partially determined by the relational properties of vehicles, that is, externalist. This is justified by the fact that behaviour, as an explanandum, involves responding to distal environmental features in order to bring distal environmental effects. Secondly, Shea's account is committed to there being real vehicles of representation. His account is thus a version of a representational theory of mind. However, he intends his theory of content to apply first and most to subpersonal, unconscious representations, since these figure prominently in cognitive science and cognitive neuroscience. Considerations about first-person, conscious mental states are given in the final chapter of the book. Finally, there is no reliance on the notion of a *representational consumer* characteristic of teleosemantics, since Shea finds it problematic to apply a consumer-based analysis on complex, "multi-layered" and feedback-involving systems. Even though his account shares with teleosemantics the view that the content of representations is determined by their use, it is their use in downstream processing by an internal mechanism that determines content, not their being used by a dedicated consumer system.

One important point of departure between Shea's account and its predecessors is his pluralistic framework, which he calls *varitel semantics*. Content, according to Shea, can be determined by more than one sufficient condition. The source of this pluralism is a disjunctive account of functions and exploitable relations, which makes content determination different depending on the type of function and exploitable relation present in each case.

The proper way to explain functions (in his account task functions) possessed by an organism, according to Shea, is by giving a *consequence etiology*. This is a point which Shea shares with other teleosemantic approaches. Dispositions for behavioral outcomes produced by an organism are explained by the fact that the same behavioral outcomes produced certain distal effects which were beneficial for the organism in the past. Because of their beneficial consequences, the production of these distal effect became a *stabilized function* of the behavioral outcome produced by an organism. However, while teleosemantics admits of natural selection as the only process that stabilizes functions (learning being a derivative of natural selection), Shea allows for three more types of stabilizing processes. In living organisms, these are *learning with feedback* and *persistence of organisms* (contribution to survival of an individual organism); while in artificial systems this process is deliberate design. On Shea's account each of these four types of stabilizing processes tend to converge with another property which Shea finds important in describing function, and that is *robustness*. Stabilizing processes tend to stabilize those behavioural outcomes which are sufficiently robust, that is, which can perform their function in a range of different conditions. One way for an organism to possess this feature is for its internal mechanisms to be sensitive to a wide range of different inputs. Representational explanation offers a solution of how organisms manage

to achieve this robustness. By having vehicles that bear relations to varying environmental features, an internal mechanism is able to exploit these relations in order to successfully perform a function in different conditions.

This brings us to the second part of Shea's theory—*exploitable relations*. Remember the central idea that a system performs its task function by having an internal mechanism whose processing is able to exploit the relations its vehicles bear to distal features of environment. These exploitable relations can be of two types. The first type of relation is *correlational information* (Chapter 4). Correlational information is usually conceived as a relation of probability raising. Obtaining of a certain state A raises the probability that another state B is also obtaining. However, a certain state usually correlates with a number of different states (and properties) and correlates so in varying strengths. The appropriate way to identify the correlation that is *actually* exploited by an internal mechanism, Shea believes, is to see which correlations have an *unmediated role* in explaining how a system performs its function. In turn, to explain how a system performs a function is to explain how that function became stabilized and robustly produced. Focusing on the explanatory role of exploitable relations ties content-determination tightly with explanatory considerations (a fact which does not imply a dependence on an intentional observer). To use a famous example, in explaining how a frog is able to catch a fly, we can identify different correlations which exist between the frog's retinal ganglion cells and the properties attributable to fly and thus a number of candidates for content-determination. The cells correlate with there being a *little black thing, a fly, a nutritious object, a fly I saw two minutes ago* etc. This is one aspect of the notorious problem of indeterminacy of content. However, Shea argues, a correlation with *little black thing* explains why the frog is able to perform its function of catching a fly only in a *mediated* way, that is, by also correlating with *nutritious object* or *fly* and *nutritious object* or *fly* is what unmediatedly explains why the function of catching flies came to be stabilized and robustly produced. If *little black thing* had not correlated with *nutritious object* or *fly*, an explanation of how the function of catching flies came to be (historically) stabilized by the frog would miss out on an important explanatory pattern. Let us now turn to the second type of exploitable relation that figures in Shea's varietal semantics—*structural correspondence* (Chapter 5).

Shea defines structural correspondence in the following way: a structural correspondence exists between a relation V on vehicles v_m and a relation H on entities x_n iff there is a function f which maps the v_m onto the x_n and $\forall_{i,j} V(v_i, v_j) \leftrightarrow H(f(v_i), f(v_j))$ (117). It is a well-known point that structural correspondence is too liberal for fixing content. This is because a certain relation V on v_m can correspond to many different relations H_1, \dots, H_n on worldly entities x_n , given that v_m and x_n are of the same cardinality. Shea proposes to constrain this liberality by requiring that the correspondence is actually exploited by a system in order to perform a function. This involves a double restriction on the candidates for content-constituting structural correspondences. First, the mechanism performing a certain function has to be sensitive to the relation which exists on the vehicles (the relation has to be used in downstream processing) and the correspondence has to play an unmediated role in explaining how a system performs a function (it has to

be of *significance* to the system). This way the notion of structural correspondence becomes sufficiently constrained and also tractable by empirical investigations. In chapter 5 Shea gives a number of wonderful demonstrations of how his account can be applied to cases of representing spatial relations, similarity and causal structure.

Next, in chapter 7 Shea offers an account of the distinction between directive and descriptive content based on his varitel framework and gives a comparison with existing accounts (Millikan, Price, Artiga, Sterelny etc.) At first approximation, directive content is concerned with producing a certain condition, while descriptive content is concerned with reflecting a certain condition. Shea draws the distinction by relying on the resources of his varitel semantics. A vehicle R standing in an exploitable relation (correlational information or structural correspondence) with a condition C has *directive content* if the production of C by a vehicle R plays an unmediated role in explaining how a system performs a task function. On the other hand, a vehicle R standing in an exploitable relation with a condition C has *descriptive content* if C's obtaining when R is tokened plays an unmediated role in explaining the system's performance of a function, but not via R's producing C (pp. 180–181). Shea then demonstrates how his way of drawing the distinction applies to different cases presented in his book.

A final aspect we will examine is the application of Shea's account to the problems of indeterminacy of content (the problems of distality, specificity and disjunction) which is presented in Chapter 6. In presenting his solution, Shea makes several comparisons between his and other solutions present in the literature. He aligns his theory with the so-called "high church" teleosemantics which ties content to explanations of successful behaviour prompted by a representation, as opposed to "low church" teleosemantics which ties content with discriminative abilities of the organism. This is so because on Shea's account the determinacy of content is grounded in the determinacy of task functions and the exploitable relations which play a role in explaining these functions. Since explanations of stabilization and robustness of task functions are causal explanations they make a restriction on the type of properties that are adequate in explaining task functions, on both the explanandum and the explanans side. This makes the proper identification of the task function involve properties which actually led to stabilization and robustness. Returning to the frog example, this means that the correct description of its task function is *catching flies* and not, for example, *catching little black things*. On the side of the explanandum, similarly, the adequate properties will again be those that figure in a causal explanation of stabilization and robustness. On Shea's account those will include, of course, the exploitable relations of correlational information and structural correspondence. Given Shea's requirement that these relations have to be those relations that unmediately explain stabilization and robustness, they will turn out to be properties such as *fly* or *nutritious object* rather than *little black thing* or *my favourite fly*. It is noticeable that this still leaves a certain degree of indeterminacy of content. However, Shea argues that this consequence is due to the nature of simple systems such as the frog. Even in the simple cases, Shea argues, determinacy of task functions and exploitable relations provides a considerable degree of determinacy on content.

It is safe to conclude that Shea's *Representation in Cognitive Science* will become an essential reading in the literature concerning mental representation. Apart from offering an ingenious account of representational content, the book provides a clear identification of the standard problems which surround theorizing about representational content. It also makes numerous comparisons to existing accounts and provides discussions about other themes, such as the notion of biological function or explanation. In spite of its complexity and extensive use of results from cognitive science and neuroscience, which demand a more specialized reader, these facts make the book sufficiently accessible to graduate level students and other, less informed, scientific and philosophical audiences that are interested in exploring the nature of mental representation.

MARKO DELIĆ
University of Split, Split, Croatia