Received: 3 June 2019 | Accepted: 17 July 2019

DOI: 10.1111/mila.12276

CRITICAL NOTICE

WILEY

Content is pragmatic: Comments on Nicholas Shea's *Representation in cognitive science*

Frances Egan

Department of Philosophy, Rutgers University, New Brunswick, New Jersey

Correspondence
Frances Egan, Department of Philosophy, Rutgers University, New Brunswick, NJ 08901.
Email: fegan@philosophy.rutgers.edu

Nicholas Shea offers what he takes to be a naturalistic account of representational content in cognitive science. I argue that the account secures determinate content only by appeal to pragmatic considerations, and so it fails to respect naturalism. But that is fine, because representational content is not, strictly speaking, necessary for explanation in cognitive science. Even in Shea's own account, content serves only a variety of heuristic functions.

KEYWORDS

cognitive science, content, explanation, heuristic, naturalism, representation

1 | INTRODUCTION

In his excellent new book, Nicholas Shea offers what he takes to be a naturalistic account of representational content. The view, a version of teleosemantics, has much to commend it. One of its virtues—reflected in its name, "varitel semantics"—is that it is *pluralistic* about the grounds for representation, recognizing that a variety of processes can give rise to content. I argue that the account secures determinate content only by appeal to pragmatic considerations, and so it fails to respect naturalism. But that is fine, because representational content is not, strictly speaking, necessary for explanation in cognitive science. Even in Shea's own account, content serves only a variety of *heuristic* functions.

Varitel semantics is offered as an account of representation for cognitive science. The book presents detailed discussion of a wide range of case studies in which content is attributed to subpersonal states. Shea gives little weight to intuitions we may have about content, which are often driven by thinking about personal-level, consciously accessible contents, in particular, the contents of propositional attitudes. This is another of the book's virtues.

I will set out, very briefly, the main features of the account, and then indicate where I think the pragmatic elements come in. Representational content arises, according to Shea, in virtue of a link between two variable elements: (a) Distally characterized *task functions* performed by a

system; and (b) *exploitable relations* that internal states of the system bear to relevant features of the environment, and which allow the system to perform those functions.

With respect to (a), at least three different kinds of processes can give rise to task functions in biological creatures: Evolution by natural selection, behavioral learning from feedback, and contribution to the continued existence of an individual organism. With respect to (b), there are at least two different kinds of exploitable relations: (1) *Correlation* between internal elements and external conditions; and (2) *structural correspondence*—the sort of relation that holds between a map and what it is about—between internal elements and external conditions. An example of (1) is the correlation between the activation of specific cells in the frog's visual system and the presence of prey in its visual field, which frogs exploit to perform the function of acquiring food. This complex of conditions enables the cells' firing to *represent* the food source.

An example of (2) is the structural correspondence between the activation of place cells in the rat hippocampus and locations in the rat's immediate environment. Coactivation of place cells corresponds to spatial proximity of locations and is exploited by the rat in calculating routes for navigation. Coactivation of place cells *represents* proximity of location.

The idea of a *task function* is a technical notion in Shea's account, so it will be helpful for following my argument later if I set out explicitly how it is to be understood:

> *Robust outcome function*
> An output F from a system S is a robust outcome function of S.
> iff
> (i) S produces F in response to a range of different inputs; and
> (ii) S produces F in a range of different relevant external conditions. (Shea, 2018, p. 55)

> *Stabilized function*
> An output F from a system S is a stabilized function of S
> iff producing F has been systematically stabilized:
> (i) by contributing directly to the evolutionary success of systems S producing F;
> or
> (ii) by contributing through learning to S's disposition to produce F;
> or
> (iii) where S is an organism, by contributing directly to the persistence of S. (p. 64)

These two notions figure in the explanation of a *task function*:

> *Task function*
> An output F from a system S is a task function of S iff
> (a) F is a robust outcome function of S;
> and
> (b) (i) F is a stabilized function of S, or
> (ii) S has been intentionally designed to produce F. (p. 65)

Task functions do not always give rise to representations. Output F is a representation (i.e., a representational vehicle with a content or correctness condition) when two further

conditions are satisfied: (a) F stands in an exploitable relation (correlation or structural correspondence) to features of the environment relevant to performing the function, and (b) F is processed according to an algorithm, and this processing is crucial to S's performing the function. This cluster of elements makes an internal structure a representation.

It is worth noting that the crucial notion of a *stabilized function* makes reference to historical facts. This is a feature that varitel semantics shares with (standard) teleosemantic accounts. Famously, the feature has consequences for "swamp creatures": Only systems with the relevant history can have representational mental states.[1] But varitel semantics has an advantage over (standard) teleosemantic accounts since evolutionary history is only one type of historical process that can give rise to appropriate task functions and hence to representation. Learning and processes that contribute to survival (and, in artifacts, *intentional design*)—processes that depend only on the organism's recent history—are others. When Swampman begins interacting with his environment—his behavioral dispositions and internal processing are of course identical to his human counterpart's—representational explanation can start to get a grip on Shea's account. This is a welcome result.

2 | IS VARITEL SEMANTICS NATURALISTIC?

I turn now to the crucial question for any naturalistic account of mental representation: Does it provide a nonintentional and nonsemantic foundation for the attribution of determinate content? I will set up the problem by reference to Lettvin, Matturana, McCulloch and Pitts' (1959) frog/fly example. The internal state responsible for engaging a frog's tongue-snapping behavior is typically triggered by flies, which are nutritious for frogs. But does the state represents *fly*,[2] *frog food*,[3] *small dark moving thing*,[4] or something else? Tokenings of the state correlate, in the frog's normal environment, with all three distal properties, but the three content candidates have different correctness conditions. If the state is indeed a representation, then there must be a definite answer to this question.

Shea attempts to secure content determinacy by appealing to a notion—*unmediated explanatory information*—that allows him to privilege some correlations over others:

> *Unmediated explanatory information*
> The UE information carried by a set of components R_i in a system S with task functions F_j is the exploitable correlational information carried by the R_i which plays an unmediated role in explaining, through the R_i implementing an algorithm, S's performance of task functions F_j. (Shea, 2018, p. 84)

[1] Davidson (1987) coined the term "swampman" to refer to a creature that coalesces spontaneously out of subatomic particles when lightning strikes a swamp. According to standard teleosemantic accounts, for example, Millikan (1984), even if swampman was a physical duplicate of a person it would have no intentional mental states, since it lacks the selection history that fixes intentional content.

[2] See Sterelny (1990).

[3] Millikan (1991).

[4] Neander (1995, 2006).

He then appeals to UE information to secure determinacy of content:

> *Condition for content based on correlational information*
> If component R of a system S with task function or functions F_j carries UE information about condition C, then R represents C. (p. 85)

Here is how Shea justifies the crucial move:

> The idea that some correlations play an unmediated role in an explanation calls for clarification. In the classic example of the frog's fly catching mechanism, the correlation of retinal ganglion cell firing (R) with little black things figures in an explanation of how the system was stabilized by evolution, but that explanation also mentions the fact that being a little black thing (condition C) correlates with being a nutritious flying object (condition C′). Without that background correlation, it would be *opaque* [emphasis added] how the correlation between R and C enabled frogs to achieve an evolutionarily beneficial outcome. So the role of the R-C correlation in that explanation is mediated. There is another explanation of stabilization that adverts directly to the correlation between R and nutritious flying objects (C′). The role of the R-C′ correlation in that explanation is unmediated. A correlation between an item R and condition C plays a mediated role in an explanation if its role depends on the explanation adverting to a further correlation between C and some further condition C′; otherwise it plays an unmediated role. (Shea, 2018, p. 84)

The account appeals to a distinction between mediated and unmediated explanation. Unmediated explanations are to be preferred; they are simpler, they require reference to fewer facts, and they are *transparent* as opposed to *opaque*. The distinction between mediated and unmediated explanation aligns with a distinction between *background correlations*—such as the correlation between R-tokenings and little (moving) black things—and (to coin a term) *primary correlations*—the correlations that figure in unmediated explanations of the system's performing its function, in this case, the correlation between R-tokenings and nutritious flying objects. Only primary correlations are content-determining. R therefore represents *nutritious flying object*.

The account appeals to a cluster of inter-related distinctions: Mediated versus unmediated explanations, background versus primary correlations, opaque versus transparent explanation. The crucial question is whether a naturalistic account of content can appeal to these distinctions to privilege one content attribution over its competitors. The distinctions appear to be *pragmatic*—in particular, to depend on preferences we, as theorists, have for simple, transparent explanations. If they do depend on pragmatic considerations, then they cannot play a content-determining role in a naturalistic theory of representation.

Let us consider the central notion of a *background correlation*. Given the specification of a task function—say, getting nutrients—what I am calling a *primary correlation* is a correlation between internal state R and the very property—being nutritious—that is specified as the task function whose performance is to be explained. Background correlations, such as the correlation between R and small dark moving things in the frog's normal environment, only explain the performance of the specified task function—getting nutrients—by appeal to an extra bit of information: That small dark moving things are correlated in the frog's normal environment with nutritious flying objects. An explanation that appeals to a background condition is

therefore more complex—more *opaque*, as Shea puts it—than one which does not need to cite the background correlation to address explicitly the task function to be explained.

Shea denies that such considerations—appeal to primary correlations and unmediated, transparent explanation—are pragmatic or interest-relative:

> The definition of unmediated explanatory information (UE information) places heavy reliance on the concept of explanation (obviously). It bases content on causal-explanatory connections. I am assuming a realist account of explanation according to which the causal-explanatory relations that figure in explanations are objective metaphysical dependence relations. This is not special pleading. Varitel semantics is making use of a resource here which other sciences also take for granted. It is not the task of a theory of content to give a theory of why causal-explanatory relations are objective. (Shea, 2018, p. 88)

I am not denying that the correlations that Shea privileges are objective. The problem is that there are too many objective correlations that might be appealed to in framing causal-explanatory generalizations of the frog's performing the task function. Consider:

> The presence of a fly causes R to fire.
>
> The presence of a small dark moving thing causes R to fire.
>
> The presence of a nutritious flying object causes R to fire. [Shea's preferred candidate]

All of these explanations appeal to objective dependence relations. The relation between R and flying nutritious things is no more *direct* (or unmediated) than the others. The unmediated (direct) versus mediated distinction does not pick out a difference in the world. Given that in the frog's normal environment nutritious flying objects are nonaccidentally correlated with small dark moving things (and with flies), the system exploits the correlation between R-tokenings and small dark moving things (and the correlation with flies) just as much as it exploits the correlation between R-tokenings and nutritious flying objects. As I noted above, an explanation that appeals to the correlation between R and nutritious flying objects is more transparent, more perspicuous *given the prior specification of the task function as getting nutrients*. But a naturalistic theory of content cannot legitimately appeal to theorists' preferences for transparent explanations to privilege one putatively content-determining correlation over the other candidates.

According to varitel semantics, representational content arises in part in virtue of exploitable relations that internal states of the system bear to relevant features of the environment. I have argued that Shea's account of the role that *correlation* plays in content determination is riddled with pragmatic elements. Appeal to the notion of unmediated explanation (UE) plays a similar role in his account of the second kind of exploitable relation—*structural correspondence*—privileging some structural correspondences over others in a way that naturalists must reject.

As will become clear in the final section, I think that Shea is right to point out that determinate content attributions in cognitive theories often depend on explanatory considerations of the sort he adduces. The dispute is about whether such considerations are pragmatic and hence out-of-bounds for naturalistic semantic accounts.

3 | IS REPRESENTATION NECESSARY?

Varitel semantics appeals to a cluster of properties to explain how organisms achieve task functions: Internal mechanisms produce stabilized and robust outcomes by exploiting correlation and structural correspondence relations to distal elements of the environment. Shea does an excellent job characterizing this cluster, its detailed structure, and its role in the explanation of an organism's behavioral success. But the cluster itself makes no reference to *representation*, and up to the point where (I have argued) he appeals to pragmatic constraints on explanation to underwrite determinate content, the account is naturalistic. But at precisely this point, where Shea tries to secure determinate content, choosing one description of the naturalistic relation over the others, the account goes pragmatic. So it would be nice for the account's naturalistic prospects if the appeal to representation were dispensable. I think that it is.

Shea argues that appeal to representation is necessary to capture generalizations that range over multiple input and output conditions, a feature central to his notion of a *robust outcome function* (Figure 1):

> Task functions are robust outcomes, so the same outcome is produced in response to a range of different proximal inputs. That means that *vehicles of content* will enter into generalizations that "bridge" across multiple proximal conditions and involve distal states of affairs (see Shea, 2018, p. 202; emphasis added).

Note the reference to *vehicles* of content entering into "bridging" generalizations. I will come back to this point below. For the moment let us focus on the notion of bridging generalizations themselves, which range over multiple input and output conditions. Shea goes on to say more about them:

> This bridging means that there are real patterns in organism-world relations—in relations of internal states to distal causes and outcomes—that are treated disjunctively in the factorized explanation. The effect of past processes of stabilization has been to key the organism into the world so that generalizations do not just concern

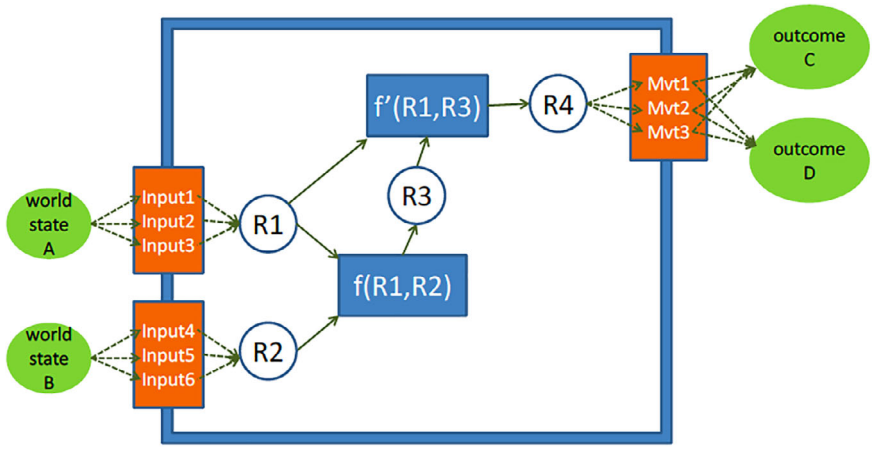

FIGURE 1 A schematic depiction of bridging at input and output (Shea, 2018, p. 202) [Color figure can be viewed at wileyonlinelibrary.com]

> how proximal causes affect the organism and how the organism affects its immediate proximal environment. A purely factorized explanation would miss the distal patterns. (Shea, 2018, p. 202)

Shea is right that a factorized explanation—appealing to disjunctions of proximal causes (Inputs 1–6 in Figure 1) and effects (Mvt 1–3)—would miss the relevant distal patterns, but the distal patterns can be characterized in *non*-representational terms. An adequate explanation will cite the correlation between world state A and internal structure R1, and between internal structure R4 and outcome C or D. In fact, a full explanation *must* cite this correlation. But nothing essential is added by saying that R1 *represents* world state A. The representational "gloss" is a convenience, drawing attention to the internal structure's correlation with relevant distal conditions (i.e., world state A) by naming the structure an *A-representation*, but the content attribution is not essential to the explanation of the organism's success. The essential function of *bridging*—"group[ing] together things that would otherwise be classified as different" (p. 207)—can be served by nonrepresentational means, in other words, by reference to the vehicles of content—R_i—precisely the structures which Shea points out (in the above quote) figure in the bridging generalizations.

Shea explicitly considers the objection that all the explanatory work can be done directly in terms of the various elements of the cluster without adverting to representation or content. I will quote his response at length:

> This...supposed challenge is not really a challenge at all, because it concedes everything we need, leaving only a dispute about the appropriateness of the label 'representation'. A distinctive style of representation-based explanation that our theories of content need to explain is that correct representation explains successful behavior and misrepresentation explains failure; that is a form of explanation where the obtaining or otherwise of facts that are distal to the organism or system make a difference to explaining its behavior. Explanations like that are part of what made representational content puzzling, even mysterious. Recognizing that the clusters I point to are real and important features of the natural world, and accepting that they underpin world-involving explanations of this kind, just is to accept that properties of the kind I have characterized do exist, and do explain behavior in the way I have claimed. (Shea, 2018, p. 205)

But Shea's response does not address the objection, which claims that the cluster can explain success (and occasional failure) without appealing to representational *content*. This is not a mere verbal dispute, as he suggests. Here is what he says of representation-based explanation:

> The content-based generalizations in psychological theories link representation with representation, often of specific types, representation with neural substrate, and representation with world. (Shea, 2018, p. 204)

But I have argued that appeal to content does not play an essential role in these explanations. When Shea says that "psychological theories link *representation* with representation ... *representation* with neural substrate, and *representation* with world" (p. 209; my emphasis), so as not to beg the question against the objection, "representation" must refer to the internal *structures* posited in these theories (i.e., the R_i), structures that are individuated

nonsemantically by their roles in processing, structures that *if they were ascribed representational content* would be called "representational vehicles."

I have argued in other work (Egan, 2014, 2018) that representational content plays various heuristic roles in theories of cognition. As I suggest above, it is convenient to gloss R_i as representing distal conditions. In my view, construing content as a heuristic makes the best sense of cognitive theorists' representational talk. To say that an internal structure represents *edge* is "shorthand" for facts about robust correlations between tokenings of the structure and distal property instantiations (typically, object boundaries) under normal environmental conditions, in other words, for some of the facts that make up Shea's cluster. I call my view a *deflationary* account of representation because it explicitly recognizes the role of pragmatic considerations in the ascription of content—in privileging one correlation over the others, as Shea's account in fact does—and preserves naturalism by confining content to a heuristic gloss. We can be straightforwardly realist about the structures (R_i) that are essential to the explanation; we can even call them *representational vehicles* (or just *representations*) because they are ascribed content and hence have correctness conditions (albeit in a gloss). These vehicles play essential roles in explanations of an organism's behavioral success (and occasional failure). To insist that they are not really representational vehicles because they are only interpreted in a heuristic gloss *would* be to engage in a verbal dispute.

I will conclude by returning briefly to the frog/fly case, changing the case slightly from frogs to toads. Neander (2017) argues that the content of a toad's internal state when it snaps at prey is *moving wormlike stimulus at location x*, supporting her argument with a detailed account of neuroscientific work on the anuran visual system. This content ascription characterizes the response profile of the relevant neural structures—T5-2 cells in the optic tectum—better than *fly* or *toad food*, since if flies and toad food do not present as moving wormlike stimuli they will not engage the toad's prey catching mechanism (moreover, a moving wormlike stimulus will activate T5-2 cells and so engage the prey catching mechanism even if it is neither a fly nor toad food). Construing the activation of T5-2 cells as *representations* of *moving wormlike stimuli* is thus a convenient way of characterizing their response profiles. But philosophers of cognition should be careful not to read too much into cognitive theorists' use of representational talk.

Personal-level thought processes are typically characterized, in commonsense, in terms of their contents. Ascribing content to the subpersonal processes described by cognitive scientists provides a common way of thinking about the two types of processes. Moreover, determinate correctness conditions support attributions of *correctness* and *mistake* (i.e., *misrepresentation*),[5] allowing us to bring subpersonal processes within the purview of normative epistemology. These may be worthwhile philosophical projects, but the cognitive scientist has no such motivation. Everything she might want to say about toad prey-recognition can be expressed more directly in nonrepresentational, causal terms. Talk of a cell's activation *representing* its distal stimulus conditions is a *gloss* that serves various pragmatic purposes but adds nothing of theoretical significance.

ORCID

Frances Egan https://orcid.org/0000-0002-5073-8340

[5] For example, where T5-2 cells are activated by something other than a moving worm-like stimulus.

REFERENCES

Davidson, D. (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, *60*, 441–458.

Egan, F. (2014). How to think about mental content. *Philosophical Studies*, *170*, 115–135.

Egan, F. (2018). The nature and function of content in computational models. In M. Colombo & M. Sprevak (Eds.), *The Routledge handbook of the computational mind* (pp. 247–258). London: Routledge.

Lettvin, J. Y., Matturana, H. R., McCulloch, W. S. & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the Radio Engineers*, *47*, 1940–1951.

Millikan, R. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

Millikan, R. (1991). Speaking up for Darwin. In B. Loewer & G. Rey (Eds.), *Meaning and mind: Fodor and his critics* (pp. 151–165). Cambridge, MA: Blackwell.

Neander, K. (1995). Misrepresenting and malfunctioning. *Philosophical Studies*, *79*, 109–141.

Neander, K. (2006). Content for cognitive science. In G. McDonald & D. Papineau (Eds.), *Teleosemantics* (pp. 167–194). Oxford: Clarendon Press.

Neander, K. (2017). *A mark of the mental: In defense of informational teleosemantics*. Cambridge, MA: MIT Press.

Shea, N. (2018). *Representation in cognitive science*. Oxford: Oxford University Press.

Sterelny, K. (1990). *The representational theory of mind: An introduction*. Cambridge, MA: Blackwell.

How to cite this article: Egan F. Content is pragmatic: Comments on Nicholas Shea's *Representation in cognitive science*. *Mind & Language*. 2020;1–9. https://doi.org/10.1111/mila.12276