

Robustness, exploitable relations and history: assessing varitel semantics as a hybrid theory of representation

Nicolás Sebastián Sánchez

Forthcoming in *Crítica* vol. 56, nº 168

Abstract

A constitutive theory of representation must address two challenges. The content determination challenge requires specifying why a particular state has a given content. The job description challenge requires spelling out the explanatory role that representational notions play in that theory. Recently, Nicholas Shea has advanced *varitel semantics* as a hybrid approach to representation to answer those challenges, supplementing teleosemantics with non-historical features –namely, exploitable relations and robustness. In this paper, I critically assess the hybrid theory’s answers to both challenges, arguing that their hybrid nature undermines their merits. In each case, I will show that it is hard to establish how the alleged complementariness of the hybrid account components works. I will conclude that internal problems beset Shea’s theory of representation.

1. Introduction

Representation is a central notion in cognitive science to which philosophers have paid much attention. Philosophers tend to consider that a constitutive theory of representation faces two challenges. First, there is a **content-determination challenge** -to specify why a given state has a given content- and second, a **job description challenge** (Ramsey 2007) -to spell out the explanatory role that representational notions play in that theory.

In *Representation in Cognitive Science* (Shea 2018, hereafter RCS), Nick Shea advances *varitel semantics* as a naturalistic theory of representation that faces these challenges. According to this view, content is determined partially by historically based and fruitful interactions between a system and its environment on the one hand and, on the other, by its relations to some of the circumstances present when the representation tokens -what Shea calls *exploitable relations*. Through this hybrid theory, Shea shows how a given system state can have determinate contents that refer to distal features of its environment. Shea also

addresses the job description challenge with a hybrid theory. In this case, the system to which the varitel semantics ascribes representational notions must meet two conditions. First, the behavioral output of the system must be a token of a behavioral type that results from historically fruitful interactions between the system and its environment. Second, the system must have mechanisms to track distal environmental features involved in that behavior, what Shea calls *robustness*.

In this paper, I will critically assess those answers. The structure of the paper will be the following. In **section 2**, I will present the two challenges a constitutive theory of representation faces, focusing on how they present themselves to naturalistic theories. In **section 3**, I will characterize Shea's answers to those challenges. Finally, in **section 4**, I will present my objections to those answers.

2. Two challenges for naturalist theories of representation

One of the main problems in the philosophy of mind is to elucidate in virtue of what there are such things as representations with intentional content (see Hattiangadi 2018: 1040). A theory that provides such elucidation may be called a foundational or *constitutive theory of representation*, one that identifies the underlying features that constitute representational phenomena so that they cannot fail to have those features and still be those phenomena (see Burge 2010: xvi). In the following subsections, I will describe the two challenges usually considered relevant for such a theory: *the content-determination challenge* and *the job description challenge*.

2.1 The content-determination challenge

A constitutive theory of representation calls for a content-determination account. Namely, an account regarding what kind of relationship between a given internal state and something else -other internal states or external circumstances- gives rise to representational content. Accordingly, constitutive theories posit a thesis like the following:

- S has E if and only if the tokening of S is brought about by the tokening of R.

Where S stands for an internal state of the system, E stands for intentional content, and R stands for some relation between S and elements internal or external to the system. Naturalistic constitutive theories auto-impose specific conceptual requirements: they want to explain how semantic properties result from physicalistic (thus non-semantic) properties. In that sense, naturalists regard R as acceptable only if it is a causal relation. Thus, expressed crudely, tokened states have DUCK as their content because of being caused by ducks¹.

¹For ease of exposition, I will describe naturalist constitutive theories as involving only causal relations between internal states and external circumstances.

A proper account of content determination requires accommodating misrepresentation, the intuitive idea that internal states may token in inappropriate circumstances. Naturalistic theories struggle to accommodate misrepresentation: if what determines the content of an internal state are the circumstances in which it tokens, the idea of inappropriate circumstances is unintelligible. The inability of a theory to distinguish content-determining circumstances from those that are not has been called the disjunction problem (see Fodor 1990). That is, if the content of an internal state depends on all external circumstances that cause its tokening, then the content of the state is disjunctive -i.e., not DUCKS but DUCKS OR GEESE. In this sense, the disjunction problem diagnoses a failure to satisfy a pretheoretical constraint of the content-determination challenge.

Even if naturalists can accommodate misrepresentation by picking just one of the external circumstances that token the internal state as content-determining, two further problems remain. First, a subtler version of the disjunction problem persists since multiple descriptions may be appropriate for that circumstance. Taking a classic example from the philosophical literature about content determination, what does a frog represent when it darts out its tongue? In this case, we can identify the darting of the tongue as an indication of S's tokening and the objects in the visual field as the ones that token R. We may describe the objects in the visual field of the organism as FLIES, but also as NUTRITIOUS FLYING OBJECTS. Furthermore, that circumstance may be described as indicating BLACK DOTS, considering that there is a reliable statistical correlation between black dots and flies in the frog's environment. These three descriptions refer to the same event, so they are, it is said, coextensional. However, intentional explanations are supposed to capture the system's perspective of the relevant circumstance, and perspectives are sensitive to description. Thus, if the best result of a theory of content-determination is a disjunction of coextensional descriptions, it is unsatisfactory.

The second source of worry, labeled 'distality attribution' problem (Artiga and Sebastián 2020: 620-622), is that the kind of content-determination causal relations can provide may not show that the system's representational states are about something distal to it. In the case of the frog, the tokening of S in the presence of flies can be accounted for by fly-shaped retinal stimulations in the frog's eyes. So, which of these two candidates should be the content of the frog's representations, FLIES or FLY-SHAPED STIMULATIONS? A proper answer to the content-determination challenge must show that the frog -and representational systems in general- is in some cognitive relation with its environment and not with mere appearances.

2.2. The job description challenge

The second challenge is to specify a job description for representational notions, to identify a property predicated upon representational explanations that account for their difference-making features. William Ramsey, to whom we owe one of the most recent formulations of this challenge (for similar views, see also Sterelny

1995: 254), states that “there needs to be some sort of account of just how [the] possession of intentional content is [...] relevant to what it does in the cognitive system” (2007: 27). In this sense, a proper answer to the job description challenge needs to identify a property that gives relevance to representational explanations and how a particular theory shows that it ascribes that property to these explanations.

Furthermore, theorizing about such a property is independent of endorsing a particular view on content. While a content-determination thesis specifies what relation a state must have with something else to qualify as having a given content, the job description challenge focuses on the specific properties that representational explanations have. This independence between the two challenges steers us towards its chief philosophical usage: to provide a constraint against too liberal use of intentional concepts. In particular, the main target of theorists that give relevance to this latter challenge is deflationist theories -like naturalistic theories: constitutive theories that consider we should explain the behavior of simple systems in representational terms.

Having presented the two challenges for a constitutive naturalist theory of representation, I analyze how Shea’s approach faces them in the following section.

3. Varitel semantics

In RCS, Shea (2018) advances varitel semantics, a theory of content determination that has as a *desideratum* meeting the job description challenge. Focusing on the explanatory practices of cognitive scientists, Shea aims to supplement them with a metaphysical foundation that renders them systematic rather than arbitrary. Thus, in speaking about these explanatory practices, he says that “[o]ur question is a meta-level question about these [scientific] theories: in virtue of what do those representations have those contents (if indeed they do)?” (10)². The exclusive focus of varitel semantics is subpersonal representations, states that have a crucial role in explaining behavior -especially in scientific contexts- without being conscious or structured like sentences in a natural language (see p. 26)³.

In the following exposition, I will focus on how Shea takes up the challenges presented in section 2.

3.1 Varitel semantics: a hybrid answer to the content-determination challenge

To present varitel semantics and how it answers to the content-determination challenge, we can pay attention to its name: ‘varitel’. The last part (‘tel’)

²All quotes with this format refer to Shea’s RCS.

³This limitation in scope does not diminish its ambition as a constitutive theory. A theory that provides a foundational account of linguistic or conceptual representations without having much to say about non-linguistic or non-conceptual ones is still a constitutive theory of a particular kind of representation.

emphasizes the influence of teleosemantics, “the closest precursor” (15) of his constitutive theory. The first part (‘vari’) comes from variety, pointing out that content determination also depends on exploitable relations -relationships between the system and its environment that are useful to the system. I will explain these two influences in turn, starting with teleosemantics.

For this latter naturalistic theory, what determines content is the evolutionary function that the mechanisms that use the representation have. The content of a representation is, then, “the circumstance that enables it to fulfill [its] function” (Macdonald and Papineau 2006: 5; see Millikan 1995). Furthermore, it is through behavioral outputs that a representation fulfills its function, and this is why the locus for content determination in teleosemantics is behavior. Varitel semantics “shares with teleosemantics a reliance on teleofunctions” (76) to establish what the system is doing and what it represents. The role of teleofunctions in varitel semantics is to individuate the behaviors that are candidates for representational explanations or, as Shea calls them, task functions.

Moreover, a behavior is a task function if it results from *causal stabilization process*, interactions between system and environment that help to cement some behavioral outcomes over others through the benefits they produce⁴. Three causal stabilization processes are relevant for varitel semantics: natural selection, survival or persistence of the organism, or feedback-based learning. Natural selection explains the presence of a behavioral outcome in an organism of a given species because that type of outcome has an adaptive advantage for the members of that species. The second stabilizing process concerns how a given behavioral outcome contributes to the persistence or survival of a particular individual. Finally, feedback-based learning explains the presence of a given behavioral outcome because the same outcome was beneficial for the individual in the past. For varitel semantics, as for teleosemantics, the notion of benefit is explained etiologically: a given outcome is present because “outcomes of the same type have been produced in the past” (48). By broadening the processes that stabilize behavioral outcomes towards targets in the environment, varitel semantics offers a contrast with standard teleosemantics by providing “a way for specifying the task being performed by a system [...] that does not depend on deep evolutionary history” (48).

The crucial difference between varitel semantics and teleosemantics is that historically constituted stabilized functions are not the only sources of content determination. The other source is exploitable relations, “relations between internal states and the world that are useful to the system” (75). In this way, Shea advocates for a pluralistic or, as I will call it, *hybrid theory* of content that has “2 (exploitable relations) x 4 [stabilization processes]⁵ content-determining conditions” (42)) where the two components are necessary and both jointly

⁴This is one of the conditions for an outcome to be a task function. I will explain the second condition in the following subsection.

⁵Shea considers 4 stabilization processes because he considers deliberate design as one of them. I will ignore this subtlety in what follows.

sufficient for content determination. One of those exploitable relations Shea is interested in is the correlational information⁶ that an internal state carries about external circumstances. An event or state carries correlational information about another if, by knowing something about the first, we also know something about the second. For the theory, the crucial point is that “correlations between internal elements and distal features of the environment show how a system’s internal organization⁷ is keyed into the world” (83).

Two motivations are behind the impulse to incorporate exploitable relations as sources of content determination. On a conceptual side, Shea’s view is that he must supplement the exclusively historical and output-based approach to content determination of teleosemantics with an input condition (see the title of his Shea 2007). For traditional teleosemantics, successful representation is just an instance of a historical pattern: the pattern of contributing to the reproduction of that kind of system. On the contrary, Shea emphasizes that behavioral success should be “explained synchronically, by internal components and exploitable relations” (70). Another motivation is that correlational information is crucial for neuroscientists and their scientific practice when they attempt to establish what part of the brain tracks a given environmental feature.

One issue with correlational information as a source of content determination is that an internal state of the system may correlate with many properties instantiated in a given external circumstance, which leads to issues such as the subtle version of the disjunction problem presented in section 2.1. Shea admits that his “definition of exploitable correlational information is extremely liberal” (79). As an additional constraint, Shea focuses on the explanatory relevance of some correlations over and above others. In this sense, “correlations that are content-constituting should be those which explain how the system achieves its task functions” (83). This content-constituting correlation affords *unmediated explanations* of behavioral outcomes, also called UE information. This conceptual tool allows satisfactory results in cases such as the frog’s darting out of the tongue. By considering the explanatory relevance of each of these contents, we see that FLY or BLACK DOT are explanatorily mediated by being nutritious because that property contributed to selection. Thus, NUTRITIOUS FLYING OBJECT is UE information in that case.

To illustrate how Shea applies varitel semantics, let’s consider one of his case studies regarding the analog magnitude system. This system, located in the parietal cortex, is used by organisms of different species to track numerosity. Its activation correlates to “the number of items in the array or sequence presented, be they visual objects, flashes, tones, etc.” (98). Given this feature of the system,

⁶A second exploitable relation Shea considers is structural correspondence, map-like representations of space. Due to space constraints, I will only consider correlational information as the relevant exploitable relation, although parallel points apply to structural representations.

⁷In RCS, the set of organized internal elements that carry the relevant correlational information is called an algorithm (see p. 34). My exposition and further critical assessment concern mainly the contents carried by the algorithm more than their bearers, so I will not focus further on the notion.

Shea asks whether it may represent “number of objects in some contexts, number of tones in others” (98) or if it “represents something common - numerosity- for all the uses to which it is put” (98). To complete the description of the case, Shea supplements it with a task function: the disposition to “report the number of items [the participants] have just been presented with” (98). That disposition became a task function because of feedback-based learning, where participants received a monetary reward for their efforts. What does, then, the analog magnitude system represent? For Shea, the correlation that unmediatedly explains the behavioral outcome, in this case, is NUMEROSITY rather than NUMBER OF OBJECTS or NUMBER OF TONES, because of the functional specialization of the parietal cortex and because the activation of the same structure in different contexts “will push in the direction of its having a common content, one which abstracts away from particular sensory features of particular situations” (99-100).

Varitel semantics, then, is a hybrid account that relies on exploitable relations and causal stabilization processes to show how systems get to represent distal objects in a determined manner.

3.2 Varitel semantics and the job description challenge

Varitel semantics is not only intended as a theory of content determination but also as an attempt to provide “a framework for content-determination specifically designed to elucidate the explanatory role of content” (23). Thus, it needs to comply with a further constraint, a *desideratum* stating that “recognizing the representational properties of [representational] systems enable better explanations that would be available otherwise” (29). The first clarification Shea provides regarding this *desideratum* is what would count as failing to meet it. Representational explanations do not meet the *desideratum* if the explanation can be ‘factorized’ into

the way the environment causes changes to intrinsic physical properties of inputs to the system; the way those inputs cause changes to other internal states of the system [...]; and the way [...] movements of the system cause changes to its distal environment (29, see also pp. 201-203, 213).

In other words, if we can factorize a representational explanation, then the theory applies representational notions too liberally and “robs content of its distinctive explanatory purchase” (42). Shea claims that this is one of the problems of “[m]ost theories of content [including teleosemantics:] while telling us how content is determined, [they] have relatively little to say about why content determined in that way has a special explanatory role” (23). By contrast, his promise with varitel semantics is that, although it may “apply quite widely” (174), it is not “unduly liberal [because its explanations] have explanatory purchase” (174).

How does varitel semantics offer a better account than a factorized explanation? The general idea is that “contents come into view when we target a

different *explanandum* [that involves][...] distal effects [and][...] distal objects and properties in the environment” (32). Representational explanations are distinctive, then, because “[they] ‘bridge’ across multiple proximal connections and involve distal states of affairs (202). In this sense, distal connections would render a purely mechanical account of that behavior inadequate. Moreover, how does varitel semantics determine that the proper *explanandum* is what makes contents come into view, how does varitel semantics determines that the proper *explanandum* is present? That is what the notion of a task function captures. As I described earlier, in the context of the content determination challenge, what makes a behavior a task function is having been subject to a causal stabilization process. However, although causal stabilization processes are the part of task functions that determine content, Shea considers another condition to determine task functions. For Shea, outcomes subjected to causal stabilization must also be *robust*, they must show that the system persists in searching for a goal. Robust outcomes, in this sense, are “behaviors that we humans are inclined to perceive as goal-directed” (52). More precisely, an organism’s output is robust if that organism “produces [the output] in response to a range of different inputs [and] [...] in a range of different relevant external conditions” (55). A system that meets these two criteria has task functions, the *explananda* of representational explanations.

Furthermore, the role of this additional constraint seems to be to meet the job description challenge since the “robust outcome aspect of task functions contributes to the proprietary explanatory purchase of representational content [...], [r]obust outcome functions ‘bridge’ to common outcomes across the range of different proximal conditions (71). This additional constraint implies a further difference with teleosemantics given that this latter theory”do[es] not require robust outcome functions. [...] Millikan’s definition of function does not include a condition that functions should be outcomes that are robustly produced” (72-73). Shea thus considers that, for teleosemantics, robustness is an accidental by-product of stabilization, while for his theory, robustness has a specific role. That role involves excluding the behaviors of some organisms from representational ascription. Shea does grant that, for varitel semantics, many outcomes subjected to stabilized processes are robust -because they are nomologically connected. However, when Shea asks if all living systems have robust outcome functions, he answers with a sound ‘no’: “[T]here are principled reasons why many cases are excluded” (213), and these reasons concern robustness.

As we can see, there are two senses in which varitel semantics is a hybrid theory: as a theory of content determination and as an answer to the job description challenge. In other words, we can say that it both constitutes the *explananda* and the *explanans* in a hybrid way. It constitutes hybrid *explanans* through stabilization processes and exploitable relations and hybrid *explananda* through robustness and stabilization processes.

4. Assessing varitel semantics

Having described the challenges for a constitutive theory of representation and Shea’s answers, I will advance objections to both answers in this section.

4.1 Varitel semantics and its answer to the job description challenge

My first criticism concerns how varitel semantics answers the job description challenge. As I argued in the previous section, what allowed the theory to yield more restrictive results than teleosemantics was the constraint that outcomes should be both stabilized *and* robust⁸. In this subsection, I will argue that the definition of robust outcome admits two readings and that none of them is satisfactory for Shea’s purposes.

For varitel semantics, the conceptual tool for meeting the job description challenge is the robustness of outcomes, behavior brought about ‘in response to different inputs’ and ‘to a range of different external conditions’. To assess which behaviors are robust and which are not, we need to specify further this functional definition since *input*, *external conditions* and *different* may be cashed out in multiple ways. However, at this point, we do not get a clear answer. Regarding different inputs, Shea says that cases of stimulus generalization “would not count as different input[s]” (56) and that we must “look at the facts of a particular case to assess what counts as a different input”. Shea does not tell, however, under which criteria we should assess those particular cases nor why stimulus generalization is not a case of dealing with different inputs. Concerning which external conditions count as different, he says that that notion “also needs careful handling” (56) and warns us that some external conditions are irrelevant to the outcome. Again, it is not clear under which criteria to determine which external conditions are relevant. Regardless of these comments, and in what we can consider the lower border of representational systems or behaviors, Shea states that *e. coli*’s behavior of “mov[ing] away from harmful chemicals” (58) deserves a representational explanation. It is “a distal outcome of the bacterium’s behavior” (58) that presents “robustness of [...] outcome (safe location) in the face of variation in external and internal biochemical parameters” (58-59).

This unconstrained definition regarding what outcomes count as robust enables *weak reading of robustness* in which every behavior stabilized by natural selection is a robust outcome, given that a behavior selected for doing something is the most basic form of goal-directedness in living systems. Shea lends support to this reading in at least two places in RCS. First, he considers that the two processes are “linked in a natural cluster” (54-55, see also p. 65 fn. 13) and that they tend to go together. Second, when introducing the notion of robustness, the only example of a non-robust outcome comes from a ball that shakes itself, a system that “would reach the bottom of a rough shallow crater from many different

⁸There is another sense in which varitel semantics is less restrictive than teleosemantics, given that natural selection is not the only process of causal stabilization that constitutes task functions.

initial positions” (55). However, that output is not robust for Shea since it “is not in any way adapting its behaviour to its circumstances” (55). Given that the only example of a non-robust outcome is a non-stabilized one, that counts in favor of taking these features as going hand in hand.

This weak reading of robustness is unsatisfactory for varitel semantics theoretical needs, namely, to show that it yielded more restrictive results than standard teleosemantics. For that purpose, we need to see how robustness operates as an additional constraint besides stabilization, and the crucial test for this is a stabilized but not robust outcome. However, the definition of robustness Shea offers when he presents the notion does not contribute to showing that there are such cases. In that sense, the weak reading of robustness is not acceptable for varitel semantics theoretical purposes.

In later moments of RCS, however, Shea focuses on which cases would be considered undeserving of representational explanation because of not being robust. In this context, Shea presents an example of a stabilized but not robust outcome:

Consider a plant that opens its flowers in the day and closes them at night. Suppose it just relies on changes in temperature, which alter internal biochemical processes. The opening and closing behaviour is produced in response to *only one input* [my emphasis], and so would not be a task function. It would be an evolutionary function of the plant, but would lack the robustness to be a task function. Now supplement the case slightly, making it more biologically realistic, so that the plant is also sensitive to light levels, giving it a second way of detecting that evening has arrived. Then the plant has two ways of detecting that it is evening, and so the flower-closing behaviour would be a (very simple) task function of the plant (Shea 2018: 213-214).

Through this case, we see how stabilization and robustness come apart. Shea provides a way of constraining what counts as *different inputs* through a commitment regarding the functional complexity of the system. This constraint requires not only that we perceive the system as goal-directed but also that it pursues its goals in a particular way by having two informational channels that track a distal feature (see Ganson 2019: 287 for a similar interpretation).

This constraint enables a *strong reading of robustness* that shows varitel semantics as having what it needs to answer the job description challenge: an additional criterion restricting representational ascription. However, what Shea needs from robustness is that it *complements* stabilization, and at this point problems arise. In what sense robustness complements stabilization? In this context, the idea should be that to determine whether some outcome is a task function, two sets of facts are necessary: facts about robustness and about causal stabilization. However, under the strong reading of robustness, facts about functional architecture are sufficient to establish that the system deserves an intentional explanation. Furthermore, Shea’s description of the plant works

under the assumption that facts about robustness are not dependent on facts about causal stabilization since we can speak about the functional architecture of the system without making recourse to historical considerations. If this is the case and facts about functional architecture are all we need to determine if a system is robust⁹, it is unclear how those facts complement facts about causal stabilization. Thus, the only reading of robustness that could answer the job description challenge turns out to be too strong since it would make Shea abandon causal stabilization processes as criteria for assessing if a behavioral outcome deserves an intentional explanation.

Regarding these critical remarks, Shea claims (personal communication) that stabilization and robustness are conceptually independent but nomologically connected, which explains why they tend to come together. In addition, he claims that the differences between robust and non-robust outcomes are a matter of degree and that there is no bright line difference between them. Since these are possible answers to my criticisms, I would like to comment on them. First, consider the idea that robustness comes in degrees. I do not see how this consideration helps Shea's prospects. If one is interested -as Shea is- in differentiating between cases for which representational vocabulary does not yield *any* explanatory benefits -because of the factorization of the explanation- and cases where representational explanation provides *some* explanatory benefits, then one needs an account that can distinguish between the presence or absence of robustness. Furthermore, at least under the strong reading, this is what Shea himself does.

On the other hand, consider Shea's comment regarding the nomological connection and its bearing on the objections to both readings of robustness. Regarding the weak reading, as we mentioned, this consideration is not only not helpful but counts against the conceptual independence of these features. Regarding the strong reading, however, it might be an answer to my worry in the following sense. As I argued, facts about robustness seem sufficient to determine if a behavior deserves an intentional explanation. However, Shea may argue that nomological necessity connects both sets of facts and that we cannot consider them independently. Thus, both would be necessary and jointly sufficient to determine that an outcome is a task function. About this suggestion, I do not see how the alleged nomological dependence implies that both sets of facts are necessary to *determine whether the behavior of a system is a task function*. One may grant that all robust systems are systems designed by natural selection without granting that the facts that determine whether a system is robust need to include facts about how natural selection designed the system. For example, one could consider that our best scientific theories regarding cognitive systems establish facts about functional architecture, a plausible suggestion since it would help explain why Shea can provide facts about what makes the hypothetical

⁹I would argue that robustness, under the strong reading, counts as a criterion for content determination. In the example, the plant has two channels that detect *that evening has arrived*, which is a determinate content. To pursue this argument further would require more space than the one I have available.

plant robust. To determine those facts regarding the inner workings of brains through our best scientific knowledge, however, there is no need to establish the particular facts about the process of natural selection that gave rise to brains with that configuration. In this sense, then, I do not see that nomological dependence undermines my objection to the strong reading of robustness: the idea that facts about functional architecture are sufficient to establish if the behavior of a system deserves an intentional explanation.

4.2 Varitel semantics and its answer to the content determination challenge

As mentioned in section 3, varitel semantics is a hybrid theory that attempts to meet the content determination challenge by combining two kinds of elements: exploitable relations and causal stabilization processes. I believe varitel semantics does not succeed in providing an answer to the content-determination challenge in the way that a hybrid theory should. My argument will proceed as follows. To begin with, I will formulate an intuitive constraint for a hybrid theory of content determination. Second, in the following two subsections, I will argue that varitel semantics does not satisfy the constraint. Finally, in the third subsection, I will assess the consequences of accepting that varitel semantics is not a hybrid theory of content determination.

Let us restate what a hybrid theory of content determination is. These kinds of theories posit that at least two factors play a crucial role in determining the content of an item. That means that each is necessary, and together, they are sufficient for content determination¹⁰. From this definition, we can extract a constraint that a hybrid theory should meet:

- **(HT)** A hybrid theory must show that given the semantic contribution of one factor: (HT1) said contribution is insufficient for answering the content determination challenge and that (HT2) adding the semantic contribution of the second component allows to answer the content determination challenge.

In other words, (HT) states that a hybrid theory fails to fulfill its promises in cases in which one of its components suffices for answering the content-determination challenge -thus failing to meet (HT1)- or cases in which the semantic contribution of the second component cannot solve the problems bequeathed by the first -thus failing to meet (HT2).

We should consider specific cases to assess whether the theory meets the constraint¹¹. Furthermore, I will focus on cases where Shea gives priority to one

¹⁰I will assume as uncontroversial that, for a hybrid theory, each factor makes the same kind of contribution for every item it applies.

¹¹Varitel semantics leaves open the possibility of some indetermination of content not relevant to the explanatory benefits of the contents assigned. In the subsequent discussion of Shea's approach, I will only focus on the kind of determination considered relevant to answering the content determination challenge, acknowledging that it is not required for the theory to provide an exhaustively precise account of content determination.

factor to check if its semantic contribution is sufficient or if a second factor should provide some further determination. In this context, giving priority to one of the factors means presenting it as what contributes in the first place to a semantic description of the case. I believe we can find that Shea prioritizes the role of each factor regarding content determination at different parts of RCS.

4.2.1 Emphasizing the role of exploitable relations in content determination

Shea prioritizes exploitable relations in content determination in his case studies and while discussing swamp systems. Let us analyze each of these in turn.

4.2.1.1 Correlational information in case studies As we saw in the previous section, Shea emphasized the role of correlational information in cases like the analog magnitude system. That structure evidenced a correlation between its activation and the number of items in the environment, and Shea wondered whether the content of that correlation was NUMBER OF OBJECTS (or NUMBER OF TONES) or just NUMEROSITY. So, the first approximation to content is a disjunction: the analog magnitude system represents NUMEROSITY OR NUMBER OF OBJECTS (OR TONES). To provide nondisjunctive content, Shea considers the functional specialization of the activated structure and experimental results regarding its activation in different contexts. These considerations favor NUMEROSITY as the content of what the analog magnitude system represents. What role do causal stabilization processes play in this case? To assess it, consider Shea’s description:

suppose people have been trained [...] to report the number of items they have just been presented with. A visual array should be reported by pressing a button a corresponding number of times, and a sequence of tones is reported by moving a graduated slider on a screen (98).

We see here a general task function like “report number of items” that comprises two more specific task functions that we could rephrase as “report number of tones” and “report number of visual stimuli”. However, it is unclear that Shea can interpret that task function as “report number of items” if he has not previously settled whether the system represents NUMEROSITY or NUMBER OF OBJECTS (OR TONES). For him to treat the system as performing the *same* task function in cases of visual and auditive stimuli and not doing different things, he must have already settled on what the system represents. By formulating in this way the task function, Shea assumes that the system represents NUMEROSITY.

Thus, establishing what the system is doing not only does not contribute to saying what it represents but rather the other way around: the task function of the system depends on determining the correlational information that the system exploits. This order of determination, from correlational information

to task functions, suggests that correlational information does all the work of determining content. In the same vein, to disambiguate between NUMEROSITY and NUMBER OF OBJECTS (OR TONES), he cites facts about functional specialization, experimental evidence of activation across contexts, and general knowledge about how “[p]erceptual representations [...] generally work” (100), facts that do not depend on specifying a task function for the analog magnitude system. The same dynamic is present in other of his case studies. First, Shea describes them taking into account exploitable relations so that what they represent gets specified, and only after that (see 80-82) does he ‘give’ a task function to the system that does not modify the contents ascribed in the initial description. In none of these cases, adding a task function contributes to the semantic description of the case. In such cases, we can say that varitel semantics fails to meet (HT1): insofar as there is content determination, and as long as causal stabilization processes and correlational information are the only two factors of the theory, correlation is sufficient.

Against this approach, one may say that to describe the system in any semantically determined way, some task function has to play at least some implicit role, for example, by arguing that unmediated explanations require causal stabilization processes. In the case of the frog analyzed in section 3.1, to single out UE information for the case, we must have the process of natural selection in view. So, in applying the UE framework in the case of the analog magnitude system, task functions should play, at least, an implicit role. Regarding this suggestion, two comments are relevant. First, Shea says that in this case, he does not apply the UE framework in this way -see the line of p. 100 quoted on section 4.3. Second, see footnote 14 for a full answer to this suggestion.

4.2.1.2 Correlational information in swamp systems Swamp systems are a molecule-by-molecule duplicate of a given biological organism but with a different history: an accidental history like having been stricken by lightning¹². For Shea, the role of this thought experiment is to force “us to reflect on whether there are good reasons for representational content to be based on history” (23). The question in these cases is whether they represent something, and -in the case that they do not- what kind of history would suffice for such systems to count as representing. With this in mind, Shea imagines “a swamp system that is an intrinsic duplicate” (167) of the neuroscience cases he is interested:

The swamp system would have the same behavioural dispositions, so would have robust outcome functions. For example, a swamp [system] would have a disposition robustly to catch [...] [an] object [...]. It would do so making use of a structure of internal processing, where those internal elements stand in appropriate exploitable relations to distal features of the environment. Since there are robust outcomes involving distal objects and properties, which proceed via a multitude of different proximal routes, there will be distal-involving real patterns

¹²A thought experiment due to Davidson (Davidson 1987: 443).

in the way the object would interact with its environment, patterns that do not depend on history (167).

Up to this point, Shea says that the system does not possess content because it has no history. His idea is that content determination would come into view with a short stabilization process like feedback-based learning. Without this history, “there are no other ingredients to draw on to make it the case that some consequences should count as successes and others not” (167). Hence, to establish to what end a system is exercising its representational capacities, it is indispensable to have it interact with its environment by, for example, cementing its dispositions with a reward.

To assess this case, I will first analyze the description of the swamp system in more detail. What Shea needs with this case, as an example of hybrid content determination, is something halfway between saying that the system has full-blown determined content -because then causal stabilization is unnecessary and (HT) is not satisfied- and that it does not represent anything -because then causal stabilization would do all the work of content determination and, again, (HT) would not be satisfied. The crucial element for a semantic description of the case is that the swamp system stands in exploitable relations to distal features of the environment. Shea also adds that the distal objects are processed “via a multitude of [...] proximal routes”, seemingly adopting the strong reading of robustness. The crucial step, however, to say that exploitable relations provide *some* determination of content but not everything needed in semantic terms is to describe the case as standing in exploitable relations with distal objects without specifying *which* objects the system is connected to.

Regarding this strategy, one could question the assumption that one can determine that a swamp system has a distal relation with something without specifying what that something is. The swamp system, as the hypothetical plant described in section 4.1, tracks distal objects via a multitude of proximal routes. In the case of the plant, however, the idea was that there were two proximal routes -light levels and temperature- that tracked the *same* thing (that evening had arrived). It seems dubious to assert that the system has a *distal* relation with the environment without specifying which object the two proximal routes are tracking. For Shea to follow this line of thinking would imply that robustness is sufficient to say to which distal features the system is related, and, thus, varitel semantics would fail to meet (HT1).

Even without considering that difficulty, adding causal stabilization processes to the semantic description of the swamp system does not seem to play the role a hybrid theory expects of them. Since said semantic description requires ascription of, at least, disjunctive contents, let us say that the above swamp system represents -through exploitable relations- RED SPHERICAL OBJECT OR APPLE. According to Shea, a feedback-based learning process would help to say which instances of behavior count as successes and failures. Reinforcing the disposition to catch objects will make the system a catcher while reinforcing the disposition to miss will make it a misser. However, a catcher (or misser) *of*

what? Since, by hypothesis, we typed the content of what the system represents as RED SPHERICAL OBJECT OR APPLE, the reinforced disposition to catch (or miss) will make the system a catcher (or misser) of red spherical objects or apples. Thus, it seems that the introduction of causal stabilization processes does not modify the partially determined content established by the initial semantic description but puts representational explanations to work in particular cases. However, this interaction between the system and specific objects and properties is just the tokening of already typed representational relations. In that sense, bringing in causal stabilization does not help to pick one of the disjuncts but rather contributes to the individuation of outcomes of the system. But what was needed to satisfy (HT) was to show that exploitable relations were insufficient to determine content and that causal stabilization would solve that problem. In this sense, Shea fails to satisfy (HT2).

4.2.2 Emphasizing the role of causal stabilization in content determination

Shea emphasizes the role of causal stabilization processes by pointing out that they are responsible for bringing in a different *explanandum* (see p. 32), an *explanandum* that refers to the environment. That constrains the terms available for the *explanans*. In addition, as I mentioned in Section 3, causal stabilization is the process by which the system connects to its environment. These considerations and the connections of varitel semantics to teleosemantics might explain why Shea, at some points of RCS, describes his cases starting from causal stabilization processes and tries to complement that description with exploitable relations to reach content-determination. In the case of the frog, for example, Shea claims that:

Some indeterminacies left open by considering task functions and causal explanations of stabilization are resolved when we ask *how a collection of correlations carried by a collection of components explains how task functions are performed* [emphasis added]. [...] [C]ausal explanations of stabilization might not choose between *fly at (x,y,z)* and *object worth eating at (x,y,z)*. But the frog's fly-capture tongue-dart mechanism is just one of the ways it gets prey. Other internal states correlate with other types of object worth eating to allow the frogs to ingest those. Saying they all just represent *object worth eating* would not capture relevant differences. So, the correlation with flies offers a more perspicuous explanation of how the whole organism achieves its suite of task functions[...]. (152)

This description stipulates that natural selection cannot disambiguate between FLY and OBJECT WORTH EATING. For example, a worm is an object worth eating for a frog. Thus, natural selection would not be precise enough to differentiate worms from flies, and we should supplement the case with correlational information. This description of the case would thus (HT).

Two comments are crucial regarding this strategy for vindicating varitel semantics as a hybrid theory. On the one hand, we have textual evidence that shows Shea endorsing a more determinate role for natural selection. When Shea discusses the case of the frog in other parts of RCS, for example, he says that natural selection disambiguates between FLIES and NURTITIOUS FLYING OBJECTS (on p. 84 and Shea’s answer to Frances Egan’s remarks in Shea 2020: 3). On the other hand, this description of the case has conceptual problems regarding the explanatory role of natural selection. I will focus on this second problem.

Why would natural selection give a disjunctive property as a result? It seems to be a condition of natural selection as an explanatory mechanism to single out the *unique* feature that was causally efficacious in contributing to the differential reproduction of the trait: this is the idea of something selected *for* something else. It is true, however, that a disjunction may ensue as a result. For critics of teleosemantics, natural selection may not be determinate enough to pick out a single description of that causal contribution, and that is why teleosemantics will fall prey to the subtle version of the disjunction problem described in section 2. Surprisingly, in his description, Shea grants a less determinate role for natural selection than critics of teleosemantics since the resulting two disjuncts are not coextensional. Other objects are worth eating, not only flies¹³. However, this disjunction cannot result from natural selection. If the mechanism was selected for catching flies, it cannot be selected for catching objects worth eating (sometimes flies, sometimes worms). If what mattered for selection was that there were flies present, then the mechanism was selected for catching flies. If it was irrelevant to selection whether the environment presented flies or worms, then the mechanism was selected for catching objects worth eating. I want to point out that these are two different selective histories and, thus, two different -and not disjunctive- selective results.

Considering the above remarks and describing the case correctly, if the frog represents disjunctively, the content of its representation is FLY OR FLYING NUTRITIOUS OBJECT, as these are coextensional disjuncts. But what is the role of correlational information in this case? It does not seem clear that correlational information can disambiguate between these two disjuncts. Recall that Shea’s account of correlational information was liberal and that applying the UE framework solved the problems of indetermination. But to apply the UE framework, stabilization processes are needed, and this leaves correlational information with no autonomous role in determining content. However, if correlational information cannot determine further the initial indetermination left by natural selection, Shea’s conclusion that ‘the correlation with flies offers a more perspicuous explanation’ does not hold. To summarize this argument, once we describe the case taking the proper explanatory role of natural selection, we

¹³No less surprisingly, he answers critics of teleosemantics by stating that the causal process of natural selection does pick out a privileged description over the set of coextensional ones: he picks NUTRITIOUS FLYING OBJECTS as the content of the frog’s representation since “it is because something nutritious was captured that the behavioral disposition was selected” (150).

see that if the first component of the theory does not provide sufficient content determination, the second will not do it. Thus, in this case, too, varitel semantics fails to meet (HT2)¹⁴.

A different case where Shea gives priority to stabilization processes is another discussion of the analog magnitude system. In that case, we see that exploitable relations do not determine content once the role of causal stabilization processes is specified:

Consider the analogue magnitude system. It is deployed in situations where behaviour is conditioned on the relative numerosity of collections of objects, doing so by using internal correlates of numerosity and comparing them. [...]. [Thus], the analogue magnitude system is an intermediate in achieving the task function of selecting the more numerous collection of objects. That goes a considerable way to making numerosity, rather than other related properties, figure in the contents represented (151).

The *explanandum* in this case is *selecting the more numerous collection of objects* while the *explanans* is *numerosity*. To the question, then, of ‘why does the system select the more numerous collection of objects?’ The answer is, trivially, ‘because it tracks numerosity’. Thus, Shea takes causal stabilization as the process specifying the *explananda* of representational explanations. Through that process, the *explanans* becomes automatically determined, and there is no complementary role that exploitable relations play. Thus, in this reading, varitel semantics does not satisfy (HT1).

4.2.3 What if varitel semantics is not a hybrid theory?

The next step is to ask if Shea can accept that his account is not hybrid after all but rather rely only on one of the factors to determine content. On the one hand, relying only on causal stabilization would be unacceptable for Shea

¹⁴The discussion of this case leads us to assess another of Shea’s attempts of vindicating varitel semantics as a hybrid theory, using the description of the case that I presented in section 3. There, Shea told us that correlational information gave us candidates for content determination and that causal stabilization gives us UE information: what unmediately explains what the frog represents is NUTRITIOUS FLYING OBJECTS because it gave the frog crucial benefits. In this sense, correlational information needs stabilization to be determinate. That is problematic for the theory because, on the one hand, why would the same theory, applied to the same case, rely sometimes on stabilization but at other times on exploitable information to determine contents? Second, Shea presents the case as if correlational information were bringing in partial determination and then further determination when supplemented by stabilization. However, if stabilization is sufficiently precise to select NUTRITIOUS FLYING OBJECT from a disjunctive ascription, it seems to be doing all the determination work. Why should we grant that correlational information gives us a disjunction that stabilization solves? It seems more straightforward to say that stabilization gives us sufficient content determination by applying the UE framework to the case. But if this is the case, stabilization is enough for content determination, and Shea’s theory fails again to meet (HT1). This description of the case seems to show that if stabilization must be in view when considering correlational information, it is not clear that stabilization is not doing all the determination work.

since it would make varitel semantics just a version of teleosemantics, losing its theoretical originality. More interesting is the theoretical option of basing content determination on exploitable relations. I think it is also not suitable for his purposes.

In all the cases in which Shea prioritizes the role of correlational information, he draws on scientific knowledge to determine content constituting correlations. Thus, in the analog magnitude case, when applying the UE framework, he draws on scientific knowledge about perceptual systems and the functional specialization of the parietal cortex. He recognizes that these considerations are not “built into the framework” (100) but uses them anyway to disambiguate between NUMEROSITY and NUMBER OF OBJECTS (OR TONES). In another case study, Shea asks if the system represents COLOR -distal content- through the prefrontal cortex or a more proximal content mediated by “activity in the organism’s primary sensory cortex” (102-103). He claims that “to find the UE information, we need to know what worldly conditions have to obtain for the monkeys to get [the] reward” (102). How do we determine those? The crucial step is the claim that “this has been set up in this case in distal terms” (102), so we can determine that the system represents the distal content COLOR and not proximal content. The question is: who has set up the case in distal terms? And the answer: neuroscientists. The implicit suggestion here is that we should take at face value scientific knowledge regarding how they describe their studies. Finally, in the case of the swamp system, to describe it as having exploitable relations, he draws on the strong reading of robustness that requires a description in contentful terms of the functional architecture of the system -what distal feature proximal routes track. And science seems to be the source of that information.

Why is it this problematic? Because it would make varitel semantics lose its naturalistic component. RCS starts with the idea that we need a foundational theory that allows a metaphysical justification for scientific discourse. Shea makes explicit his naturalistic ambitions: varitel semantics intends to be an account that relies on “non-semantic, non-mental” (10) factors. But then, if to get determined content from exploitable relations, he helps himself with the representational descriptions that scientists already give those systems, he is failing in his promise for a naturalistic approach. As an example, when he motivates why a theory of content that takes correlational information into account is descriptively adequate, he describes that “the correlation being probed [in experimental settings] are very often with distal features of the environment” (80) as a feature of the practices of neuroscientists. However, his foundational idea was to assess whether that vocabulary was justified, not to take it at face value. This move makes varitel semantics become a pragmatic content-determination account (see Egan 2020 for a different version of this argument). Thus, if varitel semantics relies only on exploitable relations, it becomes unsatisfactory as a naturalistic theory of content determination. Thus, if varitel semantics relies only on exploitable relations, it becomes unsatisfactory as a naturalistic theory of content determination.

4.3 Anticipating an objection

As a final comment, I would like to address a possible objection regarding the analytical framework I use for assessing Shea's account. This objection could come from the following claim made by Shea:

It is popular to distinguish the question of what makes a state a representation from the question of what determines its content [...]. I don't make that distinction. To understand representational content, we need an answer to both questions. Accordingly, the accounts I put forward say what makes it the case, both that some state is a representation, and that it is a representation with a certain content (10).

Thus, someone may be inclined to say that the approach I use to assess Shea's account is a way of putting things that Shea does not accept. However, when Shea says that he does not 'make that distinction', I do not take him as claiming that this kind of analytical framework is misguided. As the following sentences in the quote make clear, both answers to the challenges are necessary to understand representational content -or to give a constitutive account of representation¹⁵, as I call it. In my view, Shea claims in the above passage that the philosophical projects that only account for one of the challenges are insufficient and lack resources to 'understand representational content'. Furthermore, his strategy is to present a theory of content-determination "designed to elucidate the explanatory role of content (23)" or, in other words, that answers the job description challenge, and that's why he does not distinguish between two kinds of questions. Moreover, he addresses the issues about determinacy of content and the question of what cases are excluded from representational ascription separately. In that sense, my analytical approach is not unfair to Shea's project.

Conclusion

In this paper, I focused on Shea's constitutive account as an answer to two challenges. After describing his answers as hybrid accounts, I assessed each of them.

¹⁵A different objection regarding my analytical framework may be posed by saying that Shea is not providing a constitutive theory in the sense of positing necessary and sufficient conditions for content determination or for something to be a representation but rather, only sufficient conditions or those phenomena. In that sense, for example, he says that "there is no need to find a single set of necessary and sufficient conditions that covers all possible cases" (41) of representation. However, I believe that Shea is trying to provide necessary and sufficient conditions for something to be a representation and to have determinate content. As evidence for this, we have the idea that his project addresses a "meta-level question" (10) that attempts to provide a metaphysical justification to scientific discourse. Furthermore, in discussing task functions, he claims that they are "a necessary part of some sufficient conditions for content" (65). What to make of the first quote, then? I believe that the key expression there is 'single set'. The idea Shea is putting forward is that his theory, by being pluralist, offers a disjunctive set -and not a single set- of necessary and sufficient conditions. In fact, he later tells us that "the result of pluralism is that I am not offering a single overarching set of necessary and sufficient conditions" (42). I thank an anonymous reviewer for pressing this point.

Regarding the content-determination challenge, I argued that varitel semantics could not be seen as a hybrid theory because we could not establish how exploitable relations and causal stabilization processes worked together to determine content. Then, I argued that none of the components of the hybrid theory was successful for Shea on its own. By choosing exploitable relations as the sole source of content determination, he would end up with a pragmatic theory of content. On the other hand, by choosing causal stabilization processes as the source of content determination, varitel semantics would be nothing more than another version of teleosemantics.

Regarding the job description challenge, I argued that his hybrid account also failed to answer it. In this case, the key to his answer to this challenge was the notion of robustness. This notion admits two readings in RCS. I argued that none of those readings was satisfactory. In one case, robustness and stabilization did not come apart, showing that all stabilized outcomes were robust. Under a second reading, robustness becomes an architectural feature that does not need causal stabilization to be determined. In this case, the hybrid account does not show either how its two components work together. To conclude, Shea's answers to the job description challenge and the content-determination challenge are unsuccessful, according to his theoretical aims. In other words, internal problems beset Shea's constitutive theory of representation.

Although Shea's proposal has been discussed sympathetically, many internal problems that I have pointed out in this paper have not –to my knowledge– been addressed.

References

- Artiga, Marc, and Miguel Ángel Sebastián. 2020. "Informational Theories of Content and Mental Representation." *Review of Philosophy and Psychology* 11 (3): 613–27. <https://doi.org/10.1007/s13164-018-0408-1>.
- Burge, Tyler. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.
- Davidson, Donald. 1987. "Knowing One's Own Mind." *Proceedings and Addresses of the American Philosophical Association* 60 (3): 441–58. <https://doi.org/10.2307/3131782>.
- Egan, Frances. 2020. "Content Is Pragmatic: Comments on Nicholas Shea's *Representation in Cognitive Science*." *Mind & Language* 35 (3): 368–76. <https://doi.org/10.1111/mila.12276>.
- Fodor, Jerry A. 1990. "A Theory of Content I: The Problem." In *A Theory of Content and Other Essays*, by Jerry A. Fodor, 51–87. Representation and Mind Series. Cambridge, MA, USA: A Bradford Book.
- Ganson, Todd. 2019. "Representation in Cognitive Science, by Nicholas Shea. Oxford: Oxford University Press, 2018. Pp. 292." *Mind* 130 (517): 281–90. <https://doi.org/10.1093/mind/fzz066>.
- Hattiangadi, Anandi. 2018. "Normativity and Intentionality." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 1040–64. <https://doi.org/10.1093/oxfordhb/9780199657889.013.45>.

- Macdonald, Graham, and David Papineau, eds. 2006. *Teleosemantics: New Philosophical Essays*. Oxford ; New York: Clarendon Press.
- Millikan, Ruth Garrett. 1995. "Biosemantics." In *White Queen Psychology and Other Essays for Alice*, Reprint edition, 83–102. Cambridge, Mass: A Bradford Book.
- Ramsey, William M. 2007. *Representation Reconsidered*. Cambridge ; New York: Cambridge University Press.
- Shea, Nicholas. 2007. "Consumers Need Information: Supplementing Teleosemantics with an Input Condition." *Philosophy and Phenomenological Research* 75 (2): 404–35. <https://doi.org/10.1111/j.1933-1592.2007.00082.x>.
- . 2018. *Representation in Cognitive Science*. Oxford University Press. <https://doi.org/10.1093/oso/9780198812883.001.0001>.
- . 2020. "Representation in Cognitive Science: Replies." *Mind & Language* 35 (3): 402–12. <https://doi.org/10.1111/mila.12285>.
- Sterelny, Kim. 1995. "Basic Minds." *Philosophical Perspectives* 9: 251–70. <https://doi.org/10.2307/2214221>.