

A Teleosemantic Response to Burge's Attack on Semantic Reductionism

Sérgio Farias de Souza Filho

Federal Rural University of Pernambuco

Federal University of Pernambuco

Submitted manuscript – Please cite or quote only the published version

The Version of Record of this article is published in **Erkenntnis**, and is available online at

<https://link.springer.com/article/10.1007/s10670-023-00782-4>

[DOI 10.1007/s10670-023-00782-4]

1. Introduction

Intentionality is the capacity of the mind to represent reality, that is, to be about or to stand for things in the world. Mental representations are mental states with representational powers. For instance, my belief that Brockwell Park is in South London or some dog's memory of where it buried its bone. Theories from cognitive science and other sciences of mind posit mental representations to explain the behavioural outputs of cognitive systems and their relations with external environments. Mental representations are posited to explain the behaviour of rats, vervet monkeys, honeybees, etc. (O'Keefe & Nadel 1978; Seyfarth et al. 1980; Menzel & Greggers 2015). The job of the naturalistically inclined philosopher is to give an account of the nature of mental representations in light of successful psychological explanations established by the sciences of mind.

This is one of the goals that Tyler Burge tries to achieve in his book *Origins of Objectivity* (2010). His focus is on psychological explanations which posit representational states in perceptual systems, especially visual systems. Burge claims that genuine representational states are the ones that are posited by successful psychological explanations which are not in any way replaceable by non-representational explanations. That is, there is an *explanatory gap* between representational and non-representational psychological explanations of behaviour. The positing of the representational state is required to fully explain behaviour.

Burge's main thesis is that representational states constitute a distinctive kind of natural state, that they play a unique explanatory role in psychological explanations which cannot be assimilated to any role played by any other natural state. He proposes *primitivist naturalism*, according to which representational states are primitive natural states, i.e., are irreducible to any other state. Primitivist naturalism contrasts with the orthodoxy of *semantic reductionism* (which Burge labels "the deflationary tradition" (2010, p. 294)), according to which representational states are reducible to some more fundamental natural state (Dretske 1981; Millikan 1984; Fodor 1987). While primitivist naturalism holds that the explanatory role of representational states in psychological explanations cannot be assimilated to the role played by any other natural state, semantic reductionism denies that and claims that such an explanatory role is assimilable to the role played by some *more primitive* natural state. In order to establish primitivist naturalism, Burge developed the most compelling and influential attack on semantic reductionism from a primitivist naturalist point of view (2010, pp. 291-315). The goal of this paper is to *defend* semantic reductionism from Burge's attack.

In the next section, I assess Burge's attack on the motivations for a naturalistic reduction of representational states. He claims that the motivations actually go in the direction of primitivist naturalism. The objection is that semantic reductionism is based on a misconception of the theoretical status of representational states in psychological explanations, such that it is *out of synch* with scientific knowledge and practice, while primitivism is effectively motivated by them. In the third section, I assess Burge's attack on the viability of a reductionist approach that he takes to be the most promising one, *teleosemantics*. It tries to reduce representational states in terms of the notion of biological function by identifying *truth conditions* and *proper functioning conditions*. Burge's attack on teleosemantics is that there is a fundamental *mismatch* in the teleosemantic core thesis that identifies truth conditions with proper functioning conditions. My conclusion will be that both attacks are flawed.¹

¹ Burge's third attack is that teleosemantics and other standard reductionist theories have a conception of representation that is too liberal – they consider several states that clearly are not representational as representational states (2010, pp. 303-304). However, it is beyond the scope of this paper to assess this third attack. For a sympathetic assessment of Burge's third attack on teleosemantics, see Rescorla 2013; for a critical one, see Artiga 2016. I think that this third attack ultimately fails because teleosemantics in particular is fully compatible with the minimal conditions for intentionality, that is, the minimal conditions that a given

2. The Attack on the Motivations for Semantic Reductionism

The reduction of representational states to natural states consists in the definition of representational states in purely naturalistic terms, without appealing to any notion that involves or presupposes the notion of intentionality like representation, aboutness, content, etc. Here I will assume that natural states are the ones that are recognised by the natural sciences (i.e., physics, biology, chemistry, etc.). So, the goal of semantic reductionism is to reduce representational states to some state recognised by the natural sciences. Burge's assessment of the motivations for semantic reductionism (2010, pp. 296-298) starts with his claim that representation constitutes a distinctive state in psychological explanations. He argues that the motivations in favour of semantic reductionism are misplaced because they are *out of sync* with scientific knowledge and practice: semantic reductionists mistakenly assume that representation is a mysterious and not scientifically respected notion because it is entrenched only in folk psychology. So, it must be made scientifically respectable by reducing it to some familiar and well-established notion from natural science. However, Burge claims that the notion of representation *is* entrenched in psychological explanation, it cannot be taken to be *prima facie* defective or in need of supplementation because it has long earned its *explanatory keep* by figuring in successful psychological explanations. This fact shows that the notion of representation is *scientifically respectable* and, so, that there is no urgency of reducing it to any other scientific notion. Psychological explanations are committed to entities needed to make their explanatory claims true, among which are representational states. There is no reason to hold that a naturalist reduction of representation is required for psychological explanations to be successful. Their success is independent of any reduction.

Burge argues that our ontology should be dictated by our successful scientific explanations.² Since our successful psychological explanations posit representational states in order to explain psychological phenomena, it follows that the motivations go in the direction of accepting these representational states in our ontology. After all, Burge argues, we cannot have a better reason to rely on a notion than that it figures centrally in a successful science:

sensory state should satisfy in order to constitute a genuine representation (Souza Filho 2022). Peter Schulte (2015) develops a similar response to Burge's third attack.

² Burge develops this general thesis in his paper "Mind-body causation and explanatory practice" (1993).

Notions like representation earn their keep in science [...] by figuring in successful explanation. Successful explanation is marked in the usual ways by yielding agreement, opening new questions, making questions testable and precise, engendering progressive improvement in theory and experimentation. Mainstream work in perceptual psychology displays these features. [...] One could hardly have better epistemic ground to rely on a notion than that it figures centrally in a successful science. (Burge 2010, p. 298).

Successful explanations are the one which yield agreement in the scientific community, open new questions, engender improvement in theory and experimentation, achieve pragmatic results, etc. Burge appeals to mainstream theories in perceptual psychology (especially visual science) as manifest examples of successful psychological explanations in which representational states play a central role. So, they should be accepted in our ontology.³

But if representations are irreducible primitive states, what is left for the naturalist philosopher of mental representation to do? According to Burge, their job is to determine the place of representations in the wider natural order. That is, the philosopher should clarify, explore and connect representational states with other states in the wider natural order.

In what follows, I will show why Burge's objection is flawed. My response is divided into two fronts. On the positive one, I show that there are strong motivations for appealing to a naturalistic reduction of mental representation even assuming, like Burge, that representation is a respectable scientific notion. On the negative front, I show that there is a fundamental failure in Burge's objection when he tries to establish that there is a motivation for adopting primitivism, not reductionism.

On the positive front, let's follow Burge and assume that representation is a respectable scientific notion in psychology and hence that there is no need to reduce it to a respectable scientific notion in order to make it scientifically respectable. That is, let's assume that the motivation for reductionism is not to make the notion of representation scientifically respectable. So, should we give up reductionism? Not at all, there are still a couple of strong reasons for carrying on with the reductionist enterprise.

The first reason is that the motivation for the reductionist enterprise is not only the

³ "Is reduction of the sort expected by the Deflationary Tradition [i.e., semantic reductionism] possible? Reductions are a legitimate type of explanatory unification. Occasionally reductions succeed. In principle, representation might be somehow reducible to other notions. I believe, however, that trying to reduce representation and veridicality to something more 'naturalistically acceptable' is probably pointless and hopeless. [...] the notions of veridicality and representations – and notions like perceptual state, belief, propositional inference – are scientific primitives." (Burge 2010, p. 298).

need to make it scientifically respectable, but also by the well-known *illuminating character* of theoretical reductions. A reduction would reveal the *ultimate nature* of representational states, just like the reduction of water reveals the ultimate nature of water – H₂O. So, the need for making the notion of representation scientifically respectable is just one of the possible motivations behind the reductionist enterprise. In fact, revealing the ultimate nature of a kind of state is a motivation for a theoretical reduction *even more* strong than making it scientifically respectable.

Furthermore, Burge's own description of the philosopher of mental representation's job is *fully compatible* with the reductionist enterprise. Burge is right in claiming that the philosopher should determine the place of representational states in the wider natural order by finding systematic connections involving them. Progress can be made by clarifying, exploring and connecting representational states with the wider natural order. But that is precisely what a theoretical reduction of representational states will do: it clarifies, explores and connects them with the wider natural order; and the development of a theoretical reduction establishes a systematic connection between the representational state and the rest of nature. After all, it provides a reduction of the notion of representation in more basic natural notions which refer to entities that by their turns are more widespread in nature; and it also establishes systematic connections between the representation, the natural states that constitute the basis of the reduction and the other natural states which are connected with them in multiple ways.⁴

Let's illustrate this first reason. Suppose that representational states are reducible to biological states. This reduction would show that semantic properties are biological properties, and would connect them with the biological order and in the end with the rest of nature. A representational state is *prima facie* very different from familiar states recognised by natural sciences, but a biological reduction would reveal its ultimate nature as a familiar and well-known natural state – a biological state. So, this reduction would systematically connect the representational state with the wider nature and would reveal its ultimate biological nature.

The other reason⁴ for carrying on with the reductionist enterprise is methodological.

⁴ Note that this further argument doesn't imply that Burge's semantic primitivism is *incoherent or incompatible* with his own description of the philosopher of mental representation's job. Rather, what I am arguing here is that in light of this job there is a *motivation* for reductionism, since this job is fully compatible with the reductionist enterprise, not only with the primitivist one.

We should appeal to a supposed naturalist reduction of representation since it would be ontologically more parsimonious to treat representations as reducible to a natural state rather than as primitive. After all, the treatment of representation as reducible does not enrich the list of fundamental states in our ontology.⁵

In sum, even assuming Burge's criterion that the notion of representation should be accepted by a naturalist philosopher because it earns its keep in science by figuring in successful psychological explanations (and so that there is no necessity to make it scientifically respectable), there is still a strong motivation to carry on with the reductionist enterprise.⁶ This is the positive front of my response to Burge's objection; let's move on to the negative one.

There is a fundamental problem in Burge's defence of the view that there is a motivation for accepting representation as a primitive notion. It does not follow from successful scientific explanations regarding a notion as primitive that there is a *strong motivation* for accepting this notion as primitive. That is, the fact that a given successful scientific explanation uses a certain notion as primitive *does not imply* that there is a strong motivation for regarding this notion as irreducible. Even if sometimes this is the case, this motivation is still not sound. Notice that what Burge is claiming is that the fact that a successful scientific explanation regards a notion as primitive constitutes a motivation for accepting it as primitive, but not that this is a sufficient reason for a definitive conclusion for its primitiveness. Burge holds that since successful psychological explanations appeal to representation as a primitive notion, there is a motivation to reject the reductionist enterprise. Accordingly, the prospects are not good for a viable reduction of mental representation. In what follows, I will argue contra Burge that this fact does not constitute a motivation for primitivism about representational states.

There is a lively debate in biology and philosophy of biology whether classical genetics is reducible to molecular genetics or not. This debate is alive regardless of the fact that classical genetics provides successful explanations precisely because this fact is not a

⁵ Here I take for granted a variation of the ontological parsimony principle according to which we should not multiply fundamental entities in our ontology without necessity. Nonetheless, it is beyond the scope of this paper to defend it. For a presentation and vindication of this principle, see Schaffer 2015.

⁶ Notice that this is a debate on the motivations for reductionist naturalism, not on the reasons that purport to demonstrate that representational states are reducible to more primitive natural states. That is, the validity of these motivations implies only that we should engage into the reductionist enterprise. Whether it will succeed in reducing representational states is a further matter that will be decided in the course of this very enterprise.

sufficient motivation for treating specific notions of classic genetics as irreducible. That is, if this fact were indeed a strong motivation for primitivism, there would be no such a lively debate. The antireductionist's core claim is that any number of different molecular arrangements could correspond to a single notion in classic genetics – *gene*, *locus*, *allele*, etc. Hence, supposed bridge laws for these notions would relate each of these kinds to many molecular kinds and, so, would not be genuine bridge laws. Notice that this debate is not only about the issue of the incorporation or integration of a reduced theory with a reducing theory, but is also about the ontological issue of whether or not the entities posited by classic genetics are reducible to the entities posited by molecular genetics. If classic genetics is integrated to molecular genetics, then the entities referred to by the specific notions of classic genetics are reducible to the entities of molecular genetics; and, so, the ontological reduction is fulfilled. Finally, this debate is not the only counter-example in biology to Burge's objection. There are also other debates – e.g., whether evolutionary biology is an autonomous discipline or is reducible to molecular biology.⁷

Finally, there is also a counter-example to Burge's objection from the history of thermodynamics. There was a vibrant debate among physicists about the nature of heat between the 18th century and the beginning of the 20th century. A number of different theories of heat were developed in this period, but the most prominent is the caloric theory which explained heat in terms of the flow of a hypothetical weightless fluid called caloric and the kinetic theory according to which heat should be explained in terms of kinetic energy transfer. In the early 1850's, the laws of thermodynamics were established by R. Clausius, W. Thomson and W. Rankine which appeal to the notion of heat. Even after the establishment of the successful laws of thermodynamics, the debate on whether or not heat is reducible to some other notion persisted. In the last decades of the nineteenth century, there was a lively debate in which on one side E. Mach, P. Duhem and other physicists were objecting to the kinetic theory while L. Boltzmann and others were defending it. In the end, the defenders of the kinetic theory won this debate and it is now established that heat consists in the transference of kinetic energy from one body to another via molecular motion. That is, it is

⁷ See Sterelny & Griffiths (1999) for the reductionist debate on classical genetics and Rosenberg (2006) for the debate on evolutionary biology. What if there are indeed biological notions that, after a thorough investigation, will turn out to be *irreducible*? This is a real possibility. In this case we should treat them as valid and irreducible notions. So, reduction is not *necessary* for validating scientific notions. Nevertheless, my point (contra Burge) is that this is definitely not the case regarding mental representations.

now established that heat is reducible to molecular motion, not that it consists in some irreducible substance like caloric. So, the establishment of the successful laws of thermodynamics in which the notion of heat plays a central role happened long before the establishment of heat as reducible to molecular motion. This counter-example shows that in the history of physics, the establishment of the laws of thermodynamics was, for some thermodynamicists, a motivation for treating heat as a primitive notion, while others denied that. There was a lively debate on this issue that only ended with the ultimate conclusion that heat is kinetic energy transference via molecular motion.⁸

What is the lesson to be drawn from these counter-examples? In order for Burge to defend his objection to reductionist naturalism, he has to show why it is the case that in psychological explanations, contrary to what happens in other branches of science, the fact that a successful scientific explanation appeals to a notion as primitive constitutes a strong motivation for regarding it as primitive. That is, in the case of psychological explanations that regard the notion of representation as primitive, Burge has to show that this constitutes a strong motivation for regarding representational states as irreducible.

Burge could reply, in reaction to the above counter-examples from biology, that there is an internal debate among biologists about the viability of the reduction of notions in classical genetics and evolutionary biology. So, these counter-examples are not genuine – they do not constitute real counter-examples to his criterion that philosophers should accept that an *established* scientific notion is primitive provided that this notion appears in a successful scientific explanation as primitive. That is, if there is still a lively debate among scientists about the reducibility of a given notion, then this is not an established scientific notion and, so, philosophers should remain neutral on this issue, waiting for a scientific consensus about it.

But how decisive is the scientist's word about the reducibility of a given scientific notion? Is it the final word? This is not always the case. Because of a divergence of theoretical interests, philosophers may worry about the reducibility of a given scientific notion, while scientists are neutral about this matter simply because they happen not to be interested in it and may have never even thought about it. Regarding the notion of representation, scientists usually are not interested in the ontological problem of the

⁸ This brief description of the historical debate on the nature of heat is based on Stephen G. Brush's account of the history of kinetic theory of gases (1976).

reducibility of representational states. They are not ontologically committed; they appeal to representation for explanatory purposes without wondering about its ontological status. The second problem with the above reply is that it implicitly assumes that there is a difference of kind between the philosophical and scientific activities, not a difference of degree as assumed by methodological naturalists. Accordingly, there is an unbridgeable gap between the natures of science and philosophy. It is not my goal here to take a position in this debate, but the onus of argument lies on those who hold the thesis that there is a difference of kind between the philosophical and scientific activities, not to just assume or even be neutral about this thesis.

3. The Attack on Teleosemantics

Burge takes teleosemantics to be the most promising reductionist account of mental representation. The teleosemantic proposal is to develop a naturalistic reduction of representational states in terms of the notion of biological function. However, Burge attacks the teleosemantic core thesis that identifies *truth conditions* with *proper functioning conditions*: there is a fundamental *mismatch* in the teleosemantic identification of proper functioning conditions and truth conditions (Burge 2010, pp. 300-303). In order to show that there is such a mismatch, Burge argues that a true representation has no fitness value *in itself* and proposes a counter-example against the teleosemantic core thesis. But before assessing the mismatch objection, it is worth briefly introducing teleosemantics.

3.1 The teleosemantic core thesis

According to teleosemantics, the truth conditions of a representational state is determined in terms of its biological function (Millikan 1984; Papineau 1984, 2017). Representational states have biological functions just like more familiar biological traits (e.g., the function of the heart is to pump blood). The paradigmatic conception of biological function is the aetiological conception (also called “the selected effects conception”), according to which the function of a given biological trait is *the effect* for which the trait was *selected* (Wright 1973, Millikan 1989, Neander 1991). The *selectional history* of the trait determines the function of the trait. Thus, a biological trait has a given function in virtue of the fact that it was selected by some selection process precisely to have this effect. The

standard selection process is evolutionary selection (also called “natural selection”). This is an intergenerational inheritance process in which selected traits are the ones which historically have had an effect which increased the *biological fitness* (i.e., survival and reproduction) of the species. Accordingly, the function of the heart is to pump blood because the effect that was adaptive for ancestral hearts was to pump blood.

Teleosemantics is a theory of representational content that fixes content by identifying a mental representation’s *truth conditions* with its *proper functioning conditions*. Proper functioning circumstances are those in which the representational state performs its biological function. Malfunctioning circumstances are those in which the representational state fails to perform its biological function. For instance, consider the frog’s visual system that detects the presence of flies in the surrounding environment by producing a representational state when there is a fly around. As a result, the frog’s motor-digestive system snaps the tongue in the direction of the fly, catches and digests it. But what is the content of the frog’s representation?

There is a variety of theories that develop the teleosemantic approach in different ways. In what follows, I present the most famous and influential teleosemantic theory, *consumer-based teleosemantics* (Millikan 1984; Papineau 2017). According to consumer-based teleosemantics, the content of the representation is *food* (or *frog food*). That is, the frog represents the presence of food.⁹ Let us see why.

Consumer-based teleosemanticists appeal to a distinction between the *producer* and *consumer* systems of the representational state (Millikan 1984; Papineau 2017). The producer is the system that produces or tokens the representation – e.g., the frog’s visual system – while the consumer is the system that uses or consumes the representation to fulfil its biological function – e.g., the frog’s motor-digestive system. Notice that the consumer system uses the representation to catch and digest the fly in order to ultimately increase fitness. But what is the function of these systems? The function of the visual system is to detect *food* and it does that via the production of the representational state, while the function of the motor-digestive system is to catch and digest *food*. After all, ancestral visual systems were selected

⁹ Here I assume that the frog’s representational state is a nonconceptual representation, i.e., it has a nonconceptual content. Since beliefs (and other propositional attitudes) have conceptual content, it follows that the frog does not have a belief that there is food around. There is a whole debate on whether there are indeed nonconceptual contents, but it is not the goal of this paper to assess it. For an overview, see Bermúdez & Cohen 2020.

to detect food because that detection was required for ancestral motor-digestive systems to catch and digest food and, so, to get nutrients. As a result, the representational state has the function of detecting food. What about the content of the frog's representation?

In light of consumer-based teleosemantics, it is the *function of the consumer system* that determines content. Suppose that the consumer system has the biological function E and that it responds to the tokening of representation R with behaviour B in order to fulfil E. Now the producer system has the function of producing R when external condition C obtains. C is precisely the external condition that should be the case for B to fulfil E. So, the truth conditions of R are the external condition C that should be the case for B to fulfil the function of the consumer system, namely, E (Papineau 2017, p. 99). In light of this framework, it is clear that the content of the frog's representation is *food*. The function of the consumer is to catch and digest food, but the behavioural output only succeeds in doing so provided that the producer tokens the representation when a certain condition obtains – the presence of food. So, the truth conditions of the frog's representation are *there is food*.¹⁰

The teleosemantic core thesis is the identification of a mental representation's *truth conditions* and its *proper functioning conditions*. The truth conditions of the frog's representation are identical with its proper functioning conditions. Since the function of the representation is to detect food, the presence of food is required for its proper functioning. So, the truth conditions of the frog's representational state is the presence of food.¹¹

Finally, a qualification is necessary here. It is not that all and only true representations trigger behaviours that promote fitness. Such an identification holds only under those *historical conditions* into which the producer and consumer systems were selected. On one hand, a true representation may not promote fitness. The truth of the representation promotes fitness provided that it triggers the *appropriate behaviour* – i.e., that behaviour which

¹⁰ But why isn't the content of the frog's representation *fly* or *small-dark-moving-thing*? Why isn't the function of the consumer system to catch and digest flies or small-dark-moving things, given that in the historical environment, the frog usually obtained food when it caught flies or small-dark-moving things? This is a case of *functional indeterminacy* that threatens not only the viability of consumer-based teleosemantics, but also other teleosemantic theories. However, it is beyond the scope of this paper to assess functional indeterminacy problems. For an overview, see Neander & Schulte 2022.

¹¹ Evidently, this is a very rough simplification of consumer-based teleosemantics. For instance, the presence of food is not the *only* external condition required for the proper functioning of the frog's motor-digestive system. Other conditions are also required, such as the presence of oxygen. In light of it, Millikan claims that the content of the representation is *food*, not *oxygen*, because the producer system was selected to produce representations that *correspond* to food, not oxygen. For a full development of consumer-based teleosemantics, see Millikan, 1984, 2004; Papineau 2017.

achieves the selected effect and, so, increases fitness (e.g., the frog's prey-catching behaviour). On the other hand, a false representation may trigger a behaviour which happens to promote fitness in virtue of a *lucky coincidence*. The requirement that, in the specification of content, the relevant situation fulfils such historical conditions rules out inappropriate behaviours and lucky coincidences.

3.2 The Mismatch Objection

So far, so good. But Burge claims that teleosemantics is doomed to fail precisely in virtue of its core thesis. The mismatch objection maintains that there is a fundamental mismatch in the identification of truth conditions and proper functioning conditions – the truth conditions of a representational state are not identical with the fulfilment conditions of its biological function. The performance of a biological function by a given system historically contributes to its fitness and, so, has fitness value.¹² However, Burge argues that a true representation has no fitness value *in itself*. While it is guaranteed that the performance of a biological function has fitness value, there is no guarantee that a true representation has fitness value. Therefore, representational states are not reducible to biological functions and, so, the teleosemantic enterprise is flawed. In Burge's own words:

There is, however, a root mismatch between representational error and failure of biological function. The key deflationist [i.e., teleosemantic] idea in explaining error is to associate veridicality and error with success and failure, respectively, in fulfilling biological function. Biological functions are functions that have ultimately to do with contributing to fitness for evolutionary success. Fitness is very clearly a practical value. It is a state that is ultimately grounded in benefit of its effects for survival for reproduction. [...] But accuracy is not in itself a practical value. Explanations that appeal to accuracy and inaccuracy — such as those in perceptual psychology—are not explanations of practical value, or of contributions to some practical end.” (2010, p. 301).

In light of the mismatch objection, Burge considers the teleosemanticist reply that the biological function of a representational state is to detect the presence of a given distal condition and that the detection of this is *in itself* a contribution to fitness, while a failure of

¹² Notice that the performance of the biological function by a system contributes to fitness from a *global* point of view; a particular performance may not contribute, but the performance of the function contributes to fitness when globally (not locally) considered. That is the reason that the relevant system was favoured by natural selection throughout its evolutionary history.

detecting this is in itself a failure to contribute to fitness. However, Burge objects that the detection in itself *is not* a contribution to fitness. No biological function resides *strictly* in the detection of anything. Rather, it is the *causal properties* of the representational state which initiate or trigger the *response* to the detected distal condition that actually contributes in itself to fitness. That is, the representational state contributes to fitness by triggering the response (usually a behaviour) towards the distal condition, not by detecting it. So, the biological function of the representation is to trigger the organism's response to the distal condition. It is this initiation, not the detection *per se*, that contributes to fitness: "*in itself* detection does literally *nothing* to contribute to fitness. It is the causal properties of the detecting state in affecting responses that contribute" (Burge 2010, p. 301). This thesis is well illustrated by the case of the frog: the function of the frog's representation is *not to detect* food, but rather to *trigger* the appropriate response towards it – to catch and digest food.

However, this argument is flawed. The thesis that the biological function of the representation is only to trigger the appropriate response towards the detected distal condition is untenable. First of all, it is not true that the detection by itself does not contribute to fitness, while the triggering of the response by the detecting state does. Rather, no effect in the chain of effects which increase fitness *by itself* contributes to fitness. Consider the case of the frog again. The snap of the frog's tongue does not contribute by itself to fitness either; rather, what contributes by itself is the *whole chain of effects* – the detection of food, its capture, its digestion, the transport of the resulting nutrients in the bloodstream, etc. Notice that the capture of food does not contribute to fitness without the digestion of food, the transport of the nutrients in the bloodstream, etc. No single effect by itself contributes to fitness, only the whole chain of effects contributes by itself. The detection is an effect in this chain that contributes to fitness precisely because it is an indispensable element of this chain, just like the capture of food or the digestion. A given effect in the chain does not contribute to fitness by itself, but does so in virtue of being part of a chain of effects which, as a whole, contributes to fitness, provided that the absence of this effect would disrupt the contribution to fitness.

Burge's thesis is that the detection does not contribute to fitness by itself. The truth of a representational state is not in itself adaptive: "Evolution does not care about veridicality [i.e., truth conditions]. It does not select for veridicality *per se*" (2010, p. 303). The representational state, however, has an adaptive effect since it triggers the appropriate

response and hence increases fitness. Therefore, the frog's representation contributes to fitness by triggering the catching behaviour, not by detecting the presence of food. In order to show why Burge's thesis is flawed, it is necessary to show in detail how the detection of a distal condition is an effect of the representational state.

First, to say that the biological function of a given system is to detect a certain distal condition is a *façon de parler*. What is being said is that the function of the producer system is to token the representational state in reaction to the presence of the distal condition (even if such a reaction is not perfectly reliable). The detection of the distal condition by the producer system is the production of the representational state in reaction to the distal condition. Another way of stating this is to say that the biological function of the producer system is to produce true representations (after all, the representation is true if and only if the condition that it represents obtains). Thus, the first effect of the producer system that is triggered by the presence of the distal condition is the detection of the distal condition, i.e., the production of the representational state. The second effect is the initiation of the response directed to the distal condition (usually a behaviour). In the case of the frog, the first effect is the production of the representational state which consists in a neural firing in the brain, the second effect is in the triggering of the behaviour to capture of food, etc.

In light of this, why take the triggering of the response to the distal condition as an effect of the producer system that contributes to fitness, but not the detection of the distal condition that consists in the production of the representational state? This is *plainly arbitrary*. There is no principled way of maintaining that the contribution of the producer system to fitness resides only in the triggering of the response by the state, but not also in the production of the very state. Instead, this is the right criterion to establish which effects of the producer system contribute to fitness: every effect whose absence would *prevent* the response to increase fitness. That is, the neural firing which constitutes the detection of the distal condition, the triggering of the response, etc. Notice that in case of absence of the detection itself, there would not be any adaptive response at all. The production of the representation which consists in the detection of the distal condition is a necessary element in the chain of effects, without it there would be no triggering of the response and hence no adaptation. In the case of the frog staring at food, the absence of the neural firing implies that there would be no trigger of the catching behaviour. In sum, the detection of food is an *indispensable element* in the chain of effects. Burge could reply that the production of the representation is

adaptive only inasmuch as it triggers the response, not because it detects anything. But notice that in the absence of the detected distal condition, the response would not be adaptive – the state would trigger a non-adaptive response.

Burge recognises that there is a strong coincidence or correlation between the detection of the distal condition and the triggering of the response directed to the distal condition. However, he notes that even this strong coincidence or correlation is not identity: “[b]eing fitted to successful evolution is a matter of functioning well enough to contribute to survival and reproduction. Well enough often coincides with veridicality. But even coincidence is not identity” (2010, p. 303). Indeed, Burge claims that there are cases in which the detection of the distal condition and the triggering of the response *come apart*. In order to demonstrate that there is such a mismatch, Burge appeals to cases in which the producer system fails to detect the distal condition, but nevertheless triggers an adaptive response. That is, cases in which the truth conditions and proper functioning conditions *do not coincide*. There is no sense in saying that in these cases the producer system fails to perform its biological function – after all, there is a contribution to fitness. So, Burge concludes that producer systems were selected not because of their accuracy in detecting distal conditions, but because they trigger the appropriate responses to the distal conditions.

Burge claims that there are a plenty of cases in which the producer system contributes to fitness by triggering a behaviour which increases *strength and agility*, and, so, ultimately is adaptive, even though it *misrepresents* the presence of nearby predators. In his own words,

For example, suppose that the avoidance mechanism functioned to increase strength and agility — in avoiding the predator — even in cases in which the animal engaged in avoidance behavior, because of a misrepresentation as of a predator, when no predator was present. Suppose that in each case, whether or not the predator is present, the avoidance mechanism contributes to the animal’s fitness for avoiding predators. Then, although the ultimate *raison d’être* for the mechanism might be absent in a given case, there would be *no* biological sense in which the mechanism failed to fulfill a biological function when it effected avoidance behavior in cases where the distal condition was not present. The biological function is to contribute to a fit response to the predator — which entails contributing to avoiding predators. Failure of accuracy need not be failure to realize any biological function. Functioning in interacting successfully with respect to a beneficial or detrimental distal condition is not the same as accurately detecting the condition. (Burge 2010, p. 302)

That is, the producer system contributes to fitness no matter whether it accurately represents or not the presence of nearby predators because it increases strength and agility. So, in this case the producer system performs its biological function and, thus, there is no biological malfunction. This conclusion, however, is flawed.

The reason is that a given system may have *more than one* biological function. Evolutionary selection is plainly compatible with *distinct and parallel* functions. So, in Burge's example, the producer system may have two distinct parallel biological functions – to detect predators and to increase strength and agility. Assuming that the producer system has two parallel biological functions, it is perfectly possible for it to perform its function of increasing strength and agility but not its function of detecting predators and vice versa. It is plainly compatible that the system fulfils one function, while failing to fulfil the other function. Therefore, the producer system contributes to fitness in light of the adaptive effect of increasing strength and agility, but fails to contribute in light of the adaptive effect of avoiding predators. Generalizing this result, in Burge's example the producer system has the biological function of detecting a given distal condition that results in the avoidance of predators and also the function of triggering a behaviour which results in the increase of strength and agility.

My thesis is that the detection of the distal condition and the triggering of the response directed to the distal condition *are both* adaptive effects of the producer system. On one hand, in the majority of cases (e.g., in the case of the frog) they will constitute only one biological function – the detection of the distal condition causes the triggering of the response to the distal condition, while the failure of such a detection does not lead to an increase in fitness. Notice that both detection and triggering of the response are indispensable and that the absence of one of them breaks the chain of effects which ultimately increases fitness. On the other hand, in a minority of cases like Burge's example, the producer system has two parallel biological functions respectively constituted by the adaptive effect of detecting the distal condition to avoid predators and the adaptive effect of triggering avoidance behaviour to increase strength and agility.¹³ In a minority of cases, the system has the function of detecting the distal condition to avoid predators no matter whether this behaviour will increase strength or not; and it also has the function of triggering the avoidance behaviour to

¹³ Or even a third biological function constituted by a third effect as long as it is adaptive. There is no pre-established limit of the number of parallel functions that a given biological system may have.

increase strength and agility no matter whether this behaviour will avoid predators or not. Both functions are parallel and fully compatible and, as a matter of fact, nature is rife with biological systems that were selected to have more than one adaptive effect and, thus, more than one biological function (e.g., some feathers have the parallel functions of aiding flight and regulating internal temperature).

Finally, notice that only the detecting function has a semantic nature, the function of increasing strength and agility has no semantic nature at all – the latter function has nothing to do with the representation of anything. The function of increasing strength and agility is constituted by the effect of generating a certain behaviour that increases strength and agility, but the triggering of this behaviour is independent of the representation of any distal condition. In fact, there are several examples of systems that also increase strength and agility but which do not represent anything at all. It is just a coincidence in these minority cases that the system that contributes to the increase of strength and agility by triggering this kind of behaviour is also a *representational* producer system – it might not have been.

It is useful to compare the above response to Burge's example with David Papineau's response to another problem facing teleosemantics (2017). The problem is that, since there are cases of representations which really serve a biological function in virtue of *being false*, truth conditions cannot be identical with proper functioning conditions and thus teleosemantics is flawed. For instance, there are cases of *depressive realism* in which psychologically healthy people tend to have inflated beliefs about their own social status, contrasted to depressed people who tend to have accurate beliefs about their own social status.¹⁴ Following Papineau (2017, p. 103), let us assume that these false beliefs among healthy people have the biological function of encouraging them to be enterprising by increasing self-esteem. Thus, contrary to teleosemantics, these beliefs are adaptive in virtue of being false, not true.

Papineau's response consists in stating that this objection only arises because it mistakenly assumes that these false beliefs have only one biological function, when in reality they have more than one function (Papineau 2017, p. 104). They have the function of accurately representing reality and also the function of encouraging enterprising behaviour.

¹⁴ The depressive realism hypothesis was first proposed by L. B. Alloy and L. Y. Abramson (1979, 1988). It is not the goal of this paper to defend this hypothesis, for a critical overview of the evidences for and against it, see Ackermann & DeRubeis 1991.

So, these false beliefs fail to perform the function of accurately representing reality since they are false, but they fulfil the function of encouraging enterprise.

In fact, I was inspired by Papineau's reply in explaining why we should not worry about Burge's example. However, by doing this I am not committed to this reply to the depressive realism case. These are different and independent cases and indeed I think that this two-functions strategy is more promising in order to deal with Burge's example than to deal with the depressive realism case. The reason is that in Burge's case it is a matter of coincidence that the same producer system has the function of producing accurate representations and the function of increasing strength and agility, while in the depressive realism case it is not a coincidence that the same producer system has the function of producing true beliefs and the function of producing false beliefs in order to encourage enterprising behaviour. The crucial difference between these cases is that in Burge's example it is a coincidence that the production of a false representation increases strength and agility, while in the depressive realism case the belief is adaptive *in virtue of being false* – only a false belief can have this consequence. After all, only a *representational* producer system can produce false beliefs in order to encourage enterprise, but other systems may increase strength and agility by non-representational means.

In the depressive realism case, there is the problem of explaining how a producer system can be adaptive in virtue of producing sometimes true beliefs, sometimes false beliefs. Furthermore, how can they constitute distinct functions given that their functional statuses are not completely independent?¹⁵ Note that in depressive realism cases, it is impossible for the producer system to simultaneously fulfil both functions of accurately representing and encouraging enterprising behaviour. By contrast, these problems do not arise in Burge's example since there it is just a coincidence that sometimes the production of a false representation is adaptive; more importantly, there these functions are completely independent – it is plainly possible for them to be simultaneously fulfilled or not fulfilled, or

¹⁵ What is the criterion to determine for any given trait when it has not just one, but two or more parallel biological functions? Why are they different functions? This is a general problem in philosophy of biological function that is not my goal to assess here. However, for the present purpose it is sufficient to show that in Burge's example there are two different functions because (i) there are situations in which one function is fulfilled but not the other; and (ii) there are two different selection processes favouring different effects – to detect the distal condition and to increase strength and agility. Together (i) and (ii) are sufficient to show that in Burge's example the producer system has two different biological functions.

one fulfilled but not the other and vice-versa.

3.3 A Comparison With Other Proposals

In this last subsection, I assess and compare my proposal with two other proposals developed to respond to Burge's mismatch objection: Agustín Vicente's (2012) and Peter Graham's (2014) proposals.¹⁶ Just like my proposal, Vicente has proposed that in Burge's example the producer system has the biological function of detecting predators and also the biological function of increasing strength and agility. However, Vicente's proposal is fundamentally different from mine. He claims that these are biological functions of different kinds: the function of increasing strength and agility is an *aetiological* biological function, but the function of detecting predators is a *non-aetiological* biological function. That is, increasing strength and agility is an *adaptive effect* of the producer system, but detecting predators is not. Like Burge, Vicente maintains that evolution does not care about the accuracy of the producer system – it does not care whether or not the system detects the distal condition. Rather, the producer system has a non-aetiological biological function of detecting the distal condition, it is a *special kind* of biological function. Vicente claims that the producer system was not selected to accurately detect the distal condition. Since the aetiological function of a given system is the *selected effect* of this system, it follows that the producer system has no aetiological function of detecting anything. Nevertheless, Vicente claims that the detection is a *biological* function, since it arises from the action of natural selection – it is “the result of a process of natural selection” (2012, p. 132). That is, the detection is a function that the organism acquires as a consequence of this selection process. This is a *non-aetiological* function because it is not constituted by a selected effect. So, Vicente's conclusion is that we should “reconsider the selected effects account of functions [i.e., the aetiological conception of biological functions]” (2012, p. 132).

Burge and Vicente agree that the producer system's effect of detecting the distal condition is not adaptive and, as a result, it was not selected by evolutionary selection. Thus,

¹⁶ Justin Garson (2019, pp. 187-212) has developed a new kind of teleosemantics based on his generalised selected effects theory of biological function. It contrasts with traditional teleosemantic theories that are solely based on the standard aetiological theory of biological function. Garson suggests that his new version of teleosemantics avoids the mismatch objection (2019, p. 211). However, since he does not develop his response to this objection, I will not assess it here. Finally, I wasn't aware of Graham's paper before submitting this paper. I thank an anonymous reviewer for bringing Graham's paper to my attention.

it is not an aetiological function of the system. However, while Burge claims that accurate detection is not a biological function of the producer system *at all*, Vicente claims that it is a biological function, though not an *aetiological* one, since it is not constituted by a selected effect. This move leads Vicente to propose the revision of the aetiological conception of biological function. By contrast, what I have been arguing here is that the function of detecting the distal condition *is* an *aetiological* biological function of the producer system. Therefore, my fundamental disagreement with both Burge and Vicente is that while they claim that detecting the distal condition is not an adaptive effect of the producer system, what I am defending is that it is in fact an adaptive effect.

The problem with Vicente's proposal is that it is unclear what the nature of this *special kind* of biological function is. Notice that Vicente's approach to the detecting function is incompatible not only with the aetiological conception, it is also incompatible with other available conceptions of biological function, such as the systemic and fitness-contribution conceptions (for an overview, see Garson 2016). Thus, it seems that Vicente would have to propose an alternative conception of biological function to accommodate his thesis that the producer system's accurate detection of the distal condition is a function of a *distinct kind*. There is nothing problematic with alternative conceptions, and I am indeed open to pluralism about biological functions. However, since it is not clear what the nature of this new conception of biological function is, it is not possible to assess its viability. So, until the nature of this new conception of biological function is clarified, the thesis that accurate detection is not an adaptive effect of the producer system, but nevertheless constitutes a special kind of biological function, remains nebulous.

Peter Graham (2014), just like me, also claims that detecting the distal condition is an *adaptive effect* of the producer system. Graham argues that the producer system "contributes to fitness by accurately representing the environment, and so have accurately representing the environment as a function" (2014, p. 19). He also proposes, regarding Burge's example, that the producer system has the distinct and parallel functions of detecting predators and increasing strength and agility (2014, p. 25). So, there are two fundamental similarities between my response and Graham's response to the mismatch objection. However, there are also fundamental differences.

First of all, my proposal is a *general defence* of semantic reductionism in general and teleosemantics in particular, while Graham does not defend semantic reductionism, he even

suggests that he rejects the reductionist naturalist enterprise (2014, p. 29). Second, on one hand my proposal assumes the *aetiological conception* of biological function, in contrast with Graham's proposal (2014, p. 14). On the other hand, Graham assumes the Cummins-style functional analysis (Cummins 1975) in order to determine the contribution to fitness of a given biological trait (Graham 2014, pp. 14, 19), while my proposal is *solely based* on the aetiological conception. Third, Graham assumes that there are effects that contribute *by itself* to fitness and effects that don't (2014, pp. 20-21), while one of my main arguments against the mismatch objection is that *no single effect* of any trait contribute by itself to fitness, *only the whole chain* of effects contributes by itself to fitness. Finally, I argue that the right criterion to determine whether an effect of a given trait contributes to fitness is to verify whether its *absence* would *prevent* the relevant chain of effects to increase fitness, while Graham's criterion to determine whether an effect contributes to fitness is based on the Cummins-style functional analysis. In sum, even though my proposal and Graham's proposal share two important theses in our responses to Burge's mismatch objection, they have quite different goals and they develop fundamentally different arguments and theses in order to rebut the mismatch objection.

4. Conclusion

In this paper, I assessed Burge's attack on semantic reductionism. He first attacks the motivations behind this reductionist enterprise and I argued that it is flawed for two reasons. First, there are strong reasons for appealing to a naturalist reduction of mental representations even assuming, following Burge, that representation is an established and respected scientific notion. Second, there is a fundamental problem with Burge's thesis that there is a strong motivation for adopting primitivist naturalism – the fact that a successful scientific explanation appeals to a notion as primitive does not constitute a strong motivation for regarding it as primitive.

Burge's second attack is specifically directed to teleosemantics which assumes that truth conditions are identical with proper functioning conditions. He proposed the mismatch objection, according to which there is a fundamental mismatch between these conditions. In order to show it, Burge argued that a true representation has no fitness value in itself and specified an example in which truth conditions and proper functioning conditions do not

coincide. I argued that the mismatch objection is based on a misconception about the nature of the chain of effects that ultimately increase fitness. In reality, no effect of any system contributes by itself to fitness – it only contributes in virtue of being an element in such a chain of effects. Furthermore, the detection of the distal condition contributes to fitness since its absence would prevent such a chain of effects to increase fitness. Finally, I argued that in Burge’s specified example the producer system has two parallel biological functions – to detect the distal condition and to increase strength and ability.

Evidently, there are still several problems that challenge the viability of semantic reductionism and teleosemantics. Are mental representations ultimately reducible? Is it possible to fully determine content by identifying truth conditions and proper functioning conditions? How to give an account of cases in which representational states serve a biological function in virtue of being false (e.g., depressive realism)? These are urgent questions for anyone interested in assessing the prospects of naturalist theories in general, and teleosemantics in particular, of successfully naturalising intentionality.

6. References

Ackermann, R., & DeRubeis, R. J. (1991). Is depressive realism real?. *Clinical Psychology Review*, 11(5), 565-584.

Alloy, L. B., & Abramson, L. Y. (1979). Judgement of contingency in depressed and non-depressed students: sadder but wiser?. *Journal of Experimental Psychology: General*, 108(4), 441–485.

— (1988). Depressive realism: four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive processes in depression* (pp. 223-265). Guilford Press.

Artiga, M. (2016). Liberal representationalism. *Dialectica*, 70(3), 407-430.

Bermúdez, J. & Cahen, A. (2020). Nonconceptual mental content. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Summer 2020 Edition). Retrieved August 17, 2023, <https://plato.stanford.edu/archives/sum2020/entries/content-nonconceptual/>

Brush, S. (1976). *The kind of motion we call heat: a history of the kinetic theory of gases in the 19th century, Book 1 – Physics and the atomists*. North-Holland Publishing Company.

Burge, T. (1993). Mind-body causation and explanatory practice. In J. Heil & A. Mele (Eds.), *Mental causation* (pp. 97-120). Clarendon Press.

- (2010). *Origins of objectivity*. Oxford University Press.
- Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, 72(20), 741–765.
- Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge, MA: MIT Press.
- (1989). In defense of proper functions. *Philosophy of Science*, 56(2), 288-302.
- (2004). *Varieties of meaning*. Cambridge, MA: MIT Press.
- Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press.
- Fodor, J. (1987). *Psychosemantics*. MIT Press.
- Garson, J. (2016). *A Critical overview of biological functions*. Springer.
- (2019). *What biological functions are and why they matter*. Cambridge University Press.
- Graham, P. (2014). The Function of perception. In Abrol Fairweather (Ed.), *Virtue scientia: Bridges between virtue epistemology and philosophy of science* (pp. 13-31). Synthese Library.
- Menzel, R., & Greggers, U. (2015). The Memory structure of navigation in honeybees. *Journal of Comparative Physiology A*, 201, 547–561.
- Neander, K. (1991). The Teleological notion of function. *Australasian Journal of Philosophy*, 69(4), 458-64.
- Neander, K. & Schulte, P. (2022). Teleological theories of mental content. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Summer 2022 Edition). Retrieved January 25, 2023, from <https://plato.stanford.edu/archives/sum2022/entries/content-teleological/>.
- O’Keefe, J., & Nadel, L. (1978). *The Hippocampus as a cognitive map*. Oxford University Press.
- Papineau, D. (1984). Representation and explanation. *Philosophy of Science*, 51, 550–572.
- (2017). Teleosemantics. In D. L. Smith (Ed.), *How biology shapes philosophy: new foundations for naturalism* (pp. 95-120). Cambridge University Press.
- Rescorla, M. (2013). Millikan on honeybee navigation and communication. In D. Ryder, J. Kingsbury and K. Williford (Eds.), *Millikan and her critics* (pp. 87-106). Wiley-Blackwell.
- Rosenberg, A. (2006). *Darwinian reductionism: or, how to stop worrying and love molecular biology*. University of Chicago Press.

Schaffer, J. (2015). What not to multiply without necessity. *Australasian Journal of Philosophy*, 93(4), 644-664.

Schulte, P. (2015). Perceptual representations: a teleosemantic answer to the breadth-of-application problem. *Biology and Philosophy*, 30, 119-136.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, 210(4471), 801-803. <https://doi.org/10.1126/science.7433999>

Souza Filho, S. (2022). A Dual proposal of minimal conditions for intentionality. *Synthese* 200(115). <https://doi.org/10.1007/s11229-022-03483-7>

Sterelny, K., & Griffiths, P. E. (1999). *Sex and death: an introduction to philosophy of biology*. University of Chicago Press.

Vicente, A. (2012). Burge on representation and biological function. *Thought*, 1(2), 125-133.

Wright, L. (1973). Functions. *The Philosophical Review*, 82, 139-168.