sciendo

Can Constancy Mechanisms Draw the Limits of Intentionality?

Why the Minimal Distance Problem Threatens the Constancy Mechanism Proposal

Sérgio Farias de Souza Filho
Federal Rural University of Pernambuco

doi: 10.2478/disp-2022-0008 BIBLID: [0873-626X (2022) 65; pp. 133–56]

Abstract

What are the minimal conditions for intentionality that a sensory state should satisfy for it to constitute a representational state? That is, what are the limits of intentionality? This is the problem of demarcation. The goal of this paper is to assess a specific demarcation proposal for the minimal conditions of intentionality—the constancy mechanism proposal. Accordingly, it is a minimal condition for the intentionality of a given state that the sensory system should employ a constancy mechanism in the production of this state. First of all, I introduce the problem of demarcation and show its relevance for the debate on the viability of naturalist theories of mental representation. After that, I present the explanatory role requirement for the positing of representational states by intentional explanations of behaviour and show how it constitutes a criterion for the assessment of demarcation proposals for the limits of intentionality. Finally, I assess the constancy mechanism proposal and show that its viability is seriously jeopardised by the minimal distance problem.

Keywords

constancy mechanism proposal; minimal distance problem; minimal intentionality; problem of demarcation

1 Introduction

Naturalist theories of mental representation frequently face objections for being *too liberal* regarding the conditions for a state to be representational. A theory of mental representation is too liberal if it treats certain states as representational when they are evidently not representational. Behind this objection is *the problem of demarcation* (also called "the problem of the breadth of application" (Burge 2010: 304)): what are the limits of intentionality? That is, what is the lower border of intentionality that distinguishes the limiting cases of representational states from non representational ones? The objection that a theory is too liberal is just the objection that it has drawn the limits of representationality too low. But how to determine which states are genuine representations and which ones are not? A useful way of approaching this issue is to ask what are the *minimal conditions* that a given state should satisfy for it to constitute a representational state. Minimal conditions for intentionality are precisely those conditions satisfied by the most primitive representations. If a certain state fails to satisfy these conditions, then the state lies outside the representational domain, no matter how it might look like a representation. My goal in this paper is to assess a specific proposal for the minimal conditions for intentionality, *the constancy mechanism proposal*. Tyler Burge and, arguably, Kim Sterelny propose that it is a minimal condition for the intentionality of a given state that the system employs a constancy mechanism on its production (Burge 2010: 315–9, 342–436, Sterelny 1995). Despite the initial plausibility of the constancy mechanism proposal, I will argue that its viability is seriously threatened by *the minimal distance problem*.

2 The problem of demarcation and the objection of liberality

Representational mental states are intentional, they represent the world in a certain way. Beliefs, desires and other propositional attitudes are paradigmatic representational mental states. But along with such complex representational states there are more primitive representational states as well. There is a hierarchy of representational mental states in which very simple and primitive representational states are at the bottom of the hierarchy, very sophisticated representations at the top and representational states with intermediate complexity at intermediate levels. That is, the hierarchy is constituted by higher-level representational states at the top, lower-level representations in the bottom and intermediate-level representations in between. At the top there are propositional attitudes, abstract and scientific representations (e.g., representations of numbers, quarks etc.); and so on. At intermediate

levels, there are less complex representations like the mental representations behind prelinguistic infant intentional gestural signals (Bates, Camaioni & Volterra 1975), great ape gestural communications (Call & Tomasello 2007) and so on. What about the primitive states at the bottom of the hierarchy? Here one may ask a number of subsidiary questions about primitive states: what kind of representational states are there? Are honeybee dances which indicate the whereabouts of nectar genuinely representational? Are amoebas and anaerobic bacteria genuine representational systems? In order to answer these questions, one must establish *minimal conditions for intentionality*, based on which one can determine which representational states lie at the bottom level of this hierarchy. This is the only way to deliver a justified response for the problem of demarcation. Otherwise, any response would be arbitrary and unfounded.

Let us assess a given set of representational states. Some states are higher-level representations whose representational statuses are uncontroversial, but other states are primitive states whose representational statuses are controversial. Suppose that there is a demarcation line distinguishing representational states from non-representational ones. Accordingly, the closer a given representational state is to this line, the more doubtful is its representational status. Furthermore, this line distinguishes primitive representations from other states that even though are not representational, their non-representational statuses are controversial too. These are *borderline cases* of primitive representations. *The problem of demarcation* consists precisely in how to demarcate which states are representational and which states are not. The role of minimal conditions for intentionality is to establish conditions to distinguish, among borderline cases of primitive representations, states that are genuinely representational from states that are not.[1]

Here I suppose that primitive representations are subpersonal *sensory representations* such that the limits of intentionality are demarcated by sensory representations. That is, lower-level representational states *are* sensory representations. The role of minimal conditions for intentionality is to distinguish representational sensory states from non-representational sensory states. The latter just *register* or *indicate* the relevant conditions, in contrast with sensory representations that *represent* them. Let us consider paramecia, unicellular ciliated organisms that are highly sensitive to the presence of light: whenever there is light in the surrounding environment, they move in the opposite direction of it. This

[1] I have further developed this approach of minimal conditions for intentionality to the problem of demarcation, see Souza Filho 2018, 2022.

phenomenon is called tropism, i.e., the turning of the organism (or parts of it) in a given direction in response to the presence of a certain external stimulus. So, if paramecia are representational systems, they represent the direction of light; if they are not, they merely register or indicate it. Non-representational sensory systems have sensory discriminations of the relevant conditions, but there is no genuine representation, in contrast with sensory representational systems that genuinely represent them.[2]

In light of the problem of demarcation, several philosophers have objected that naturalist theories of mental representation are too liberal on the minimal conditions for a given mental state to be representational—they draw the limits of intentionality too low. This is the *objection of liberality*. It has been time after time proposed specially against reductionist naturalistic theories such as teleosemantics (Papineau 1993, 2016 and Millikan 1984, 2004) and causal-informational theories (Dretske 1981, Fodor 1987). They are criticised for treating states that evidently are not representational as genuine representational states. The following cases are illustrative of this objection: magnetosome states of anaerobic bacteria and the reptile's body states which varies with the heat of the sun (Burge 2010: 300, 303–4); hormone concentrations in our blood (Schulte 2015: 119–20); detection of light in amoebas or planarias for phototaxis (Fodor 1986: 10–1), etc. The objection is that teleological or causal theories imply that these states are genuinely representational, but it is a mistake to treat these states as genuine representations—they clearly do not represent anything. Thus, the conclusion is that these naturalistic reductionist theories should be ruled out or at least highly refined.[3]

Let us consider some of these cases. First, when the osmolarity of human blood (roughly, the number of particles in a given quantity of blood plasma) reaches above a certain level, the antidiuretic hormone (also called "vasopressin") is produced by the hypothalamus. The antidiuretic hormone then acts on the kidneys, where it triggers an increase in the quantity of water that is reabsorbed during urine formation, resulting in the lowering of the osmolality of the blood. So, why not claim that the antidiuretic hormone, produced by

---

[2] The assumption that primitive representational states are sensory representations is widespread in this debate (Burge 2010: 317, Millikan 2004: 80, 158–70, Schulte 2015). Here I also assume that primitive representations are *single modal* sensory states.

[3] For an assessment of the objection that causal-informational or teleosemantic theories of representation imply that the aforementioned sensory states are representational, see Artiga 2016, Burge 2010: 291–308. For an assessment of how the objection of liberality threatens another proposal for the limits of intentionality—the success pattern proposal—see Souza Filho 2022: 8-11.

the hypothalamus, represents high plasma osmolarity? Second, the anaerobic bacterium's magnetosomes are attracted to the prevailing magnetic field because it correlates with the direction of oxygen-free water – the oxygenated water is lethal for the bacterium. So, why not claim that magnetosomes genuinely represent the presence of oxygen-free water? The problem that arises here is why it is the case that sensory states like the antidiuretic hormone and anaerobic bacterium's magnetosome states are not genuinely representational. What is the reason behind it? This verdict only stands in light of the establishment of one or more minimal conditions for intentionality, in such a way that the aforementioned sensory states fail to satisfy them. This is precisely what the constancy mechanism proposal tries to do.

3 The constancy mechanism proposal[4]

The debate on the minimal conditions for intentionality is focused on the distinctive *explanatory role* of representational states in intentional explanations of behaviour. The explanatory criterion establishes that for a given state to be a mental representation, its representational status should have a distinctive explanatory role. That is, the positing of a mental representation by a given cognitive theory must earn its explanatory keep.[5]

Cognitive science and other sciences of mind have developed theories that posit representational states in order to explain the behavioural outputs of organisms. So, there are cognitive theories that posit representational states to explain the behaviour of rats (O'Keefe & Nadel 1978), honeybees (von Frisch 1967), great apes (Call & Tomasello 2007), etc. The positing of a representational state by a cognitive theory is supposed to play a distinctive explanatory role, i.e., the state should have an explanatory power that justifies

[4] I have previously assessed the constancy mechanism proposal (Souza Filho 2017), in this section I further develop such assessment. Following Marc Artiga's suggestion (2016: 422), I objected that the counterfactual robustness argument only shows that there is a distinction between representational states with *proximal content* and those with *distal content*, not a distinction between *representational* and *non-representational states* (2017: 39–40). However, now it strikes me that this objection is ultimately flawed. Now I argue that the minimal distance problem, to be developed in this section, is what really threatens both the counterfactual robustness argument and the constancy mechanism proposal.

[5] The explanatory requirement for the positing of mental representations has been proposed by several philosophers of mental representation, see Burge 2010, Ramsey 2007: 24–34, Rescorla 2013 and Shea 2007.

its positing as a representation.[6] Consider a cognitive theory that posits a representational state in a given organism to explain its behaviour. According to the intentional explanation, the organism has this behaviour because it represents the presence of a given external condition. Provided that this state plays the distinctive explanatory role of representational states in the intentional explanation, the assignment of a representational status is explanatorily justified and, so, it is a genuine representational state. By contrast, if the positing of a representational state fails to have any explanatory purchase and, hence, the intentional and non-intentional explanations have the same explanatory powers, then the non-intentional explanation should be preferred in light of parsimony considerations.

The lesson to be taken is that if a given demarcation proposal considers certain states as representational but they do not play the distinctive explanatory role of representational states, then this proposal should be rejected. The explanatory criterion is meant to verify whether the states that satisfy the minimal conditions established by the proposal are explanatorily justified as representations. If so, the satisfaction of these conditions by a given state guarantees that its being posited as a representation is explanatorily justified, and, so, it is genuinely representational.[7] But what is the distinctive explanatory power of representational states in intentional explanations that justifies the positing of these states? This is what the constancy mechanism proposal tries to specify. In fact, the specification of the explanatory power of representational states goes together with the specification of conditions for minimal intentionality.

Some sensory systems are capable of producing tokens of the same sensory state that covary with tokens of a certain distal feature, even though at different instants there is a huge variety of proximal stimuli reaching the system's sensory apparatus. That is, despite the multiplicity of intermediate stimuli between the sensory apparatus and the distal feature, the sensory state remains the same. But how is it possible? How the sensory state covaries with the distal feature and not with the proximal stimulus? How the system keeps constant the sensory state despite the variety of proximal stimuli coming from the same distal feature? In order to do that, the system should employ a *constancy mechanism*. This is

---

[6] Cognitive scientists usually do not assess the demarcation problem, they just assume the notion of representation in the development of their scientific investigations (Palmer 1978: 259). I take to be part of the philosopher's job to assess the demarcation problem that cognitive scientists put aside.

[7] Is the explanatory role the only relevant criterion for the assessment of demarcation proposals? This is an open question, but in this paper I focus only on the explanatory aspect of the debate on the problem of demarcation.

a mechanism that guarantees that the system produces tokens of the same sensory state despite great varieties (to a given extent) in proximal stimuli coming from the surrounding environment.[8] There are cases of colour, shape, size and position constancy. Let us consider the first two.

There is *colour constancy* when the visual system sees an object with the same colour even when there are great differences in the environmental light conditions. The system sees the same colour even though there is a great variety of light reflected by the object under different lighting conditions, i.e., there is a multiplicity of light rays reaching the retina. To illustrate colour constancy, consider a white paper which appears to have a uniform colour under a highly uneven illumination, in spite of the fact that the light reflected by the paper's shaded part is different from the light reflected by the unshaded part. Here there is colour constancy since we see all parts of the paper with the same colour, despite a great variation in the illumination which reaches these parts. Now let us move on to *shape constancy*: a coin looks round when viewed head-on as well as when viewed from acute angles, despite the areas projected by the coin on the retina are hugely different under these two conditions. This is a case of shape constancy because the coin looks as having the same shape, even though it is seen from very different angles and, so, there is a huge variety in the light intensities reflected by the coin.[9]

It is here that several philosophers have come up with the idea that constancy mechanisms draw the limits of intentionality. Burge proposes that the limits of perception are the limits of intentionality and that it is the employment of constancy mechanisms that distinguishes perceptual from non-perceptual states: "perception marks the lower border of representation. Perception lies not only at the root of empirical objectivity. It is, I think, where states with veridicality conditions first clearly emerge" (Burge 2010: 549). So, what distinguishes representational states from non-representational states is that only the production of the first states employ constancy mechanisms:

---

[8] To put it in another way, "[p]erceptual constancy refers to the perception of these object properties as remaining constant even when things such as illumination, distance, and viewing angle cause changes in the stimulation these objects present to the receptors" (Goldstein 2010: 309).

[9] This is a simplified and as neutral as possible account of constancy mechanisms. A proper account would have to assess different problems, like the ones on similarities and dissimilarities aspects of the appearance of the object throughout changes of proximal stimuli (e.g., colour, shape, etc.). It is not my goal to enter into this debate here, for a discussion on the proper account of constancy mechanisms, see Cohen 2015, Hilbert 2005, Schulte 2021.

> [C]ertain processes in perceptual systems systematically distinguish effects of stimulation that are special to the individual and the context from perspective-independent attributes of the wider environment. Explanation of the formation of perception keys on processes in perceptual systems that make this distinction. Such processes constitute the ground of perception, representation, and objectivity. (Burge 2010: 23)

Arguably, Sterelny also proposes that the employment of the constancy mechanisms is a minimal condition for intentionality, but he is not committed to Burge's thesis that the limits of intentionality are the limits of perception. Sterelny proposes that it is a minimal condition for a sensory state to constitute a representation that "there is a sufficient variety of proximal routes and sufficient stability of distal sources" (Sterelny 1995: 261–2). In short, the constancy mechanism proposal is the thesis that it is a minimal condition for a system to be genuinely representational that it employs a constancy mechanism. But what are the reasons for the constancy mechanism proposal? Why is it the case that the employment of a constancy mechanism is a condition for minimal intentionality?[10]

The main argument for the constancy mechanism proposal available on the market is *the counterfactual robustness argument*. It was originally proposed by Sterelny (1995: 258–62) and, followed by him, Schulte (2015: 132–4). The argument is based on a distinction between two kinds of explanations—*robust-process* and *actual-sequence explanations*.[11] Let us illustrate this distinction with two different explanations for the result of a football match. First, the actual-sequence explanation: a detailed description of everything that happened in the match, describing every header, kick, dribble, goal, shooting and so on—i.e., a complete description of every move in the football match. Second, the robust-process explanation: a description of the abilities of every footballer in both teams, concluding that in light of the fact that one team is much better than the other, it would probably win anyway. As it happens, one explanation has advantages over the other, which is the reason that one is not fully replaceable by the other. The actual-sequence explanation is *more specific*—after all,

---

[10] Sterelny does not *explicitly* claim that the employment of a constancy mechanism is required for intentionality. However, his position implies the constancy mechanism proposal, since it claims that it is a minimal condition for intentionality that the sensory system produces the same sensory state, despite the variety of proximal stimuli coming from the same distal condition (Peter Schulte (2015: 134) observes that Sterelny's position supports the constancy mechanism proposal).

[11] Here Sterelny appeals to a distinction originally drawn by Frank Jackson & Philip Pettit (1992) between comparative and contrastive explanations.

it consists in a complete description of every move in the match, including a description of why one team defeated the other. But only the robust-process explanation is capable of explaining what *would probably have happened* if things were a bit different: for instance, if the teams would have played in a different stadium, the stronger team would still probably defeat the weaker team.

The starting point of the counterfactual robustness argument is that intentional explanations are *robust-process explanations*, not actual-sequence ones. The actual-sequence explanation of behaviour fully specifies the exact sequence of physical-chemical events which resulted in this behavioural output. In opposition, the intentional explanation of behaviour is a robust-process explanation since it is capable of explaining how the organisms *would* have behaved if things were different. The fundamental *explanatory difference* between intentional and non-intentional explanations of behaviour is that only intentional explanations are capable of giving an account of what would have happened in *counterfactual scenarios*.

Contrast the actual-sequence and robust-process explanations of the approaching behaviour of the organism when it finds something in the surrounding environment. Why did the organism move in this direction? The actual-sequence explanation specifies everything that happened in the motor and visual systems of the organism: the shadow in the organism's retina, the exact position of the perceived object and of everything else on the organism's visual field, etc. But it fails to explain what would have happened if the object were in a different position or angle on the organism's visual field. It is here that the great advantage of the intentional explanation as a robust-process explanation arises: the organism had this approach behaviour because it saw some food. That is, the organism *represented* that the perceived object is a food. For instance, in case of the food were not in position x but in position y or in case of the food were not in the angle α but β, the organism would still have the behaviour of moving toward its direction.

The distinctive explanatory role of representational states in the explanation of behaviour is that they have *counterfactual robustness*. As an actual example, Sterelny mentions the piping plover's behaviour when it sees a predator approaching its nest, which consists in the feigning of a broken wing in order to seduce the predator into its direction and, hence, to keep it away from the nest (Ristau 1991, Sterelny 1995: 261). The piping plover responds in this way precisely because it represents the presence of the predator in a given location, and it would have the same behaviour in counterfactual scenarios in which the predator lies in nearby locations.

Intentional explanations of behaviour have the advantage of explaining counterfac-

tual situations. But this gain in systematicity always comes with the loss of richness of detail—after all, you cannot have your cake and eat it. Actual-sequence explanations of behaviour appeal to *proximal features* (e.g., a light ray or a shadow in the retina). Intentional explanations explain behaviour as a response to the represented *distal feature* (e.g., some food or a male mate), not to the proximal one.

In light of this characterization of intentional explanations as robust-process explanations, it becomes clear why only systems that employ constancy mechanisms have behavioural patterns that are prone to intentional explanations. In the absence of the constancy mechanism, the system's responsive behaviour is triggered only by a specific proximal stimulus. The behaviour is fully *chained* to this proximal stimulus. But as we saw, the intentional explanation of the behavioural output appeals to the system's response to the distal feature and the system can only respond to the distal feature—not to proximal stimuli—provided that it employs a constancy mechanism. Thus, the employment of a constancy mechanism is a condition for the responsive behaviour to be prone to an intentional explanation and, so, constitutes a minimal condition for intentionality.

According to the constancy mechanism proposal, the piping plover is a genuine representational system, since it employs a constancy mechanism in the tokening of the sensory state that triggers the behavioural response towards the predator. Note that the piping plover's behaviour is very flexible and is triggered by a variety of proximal stimuli routes. That is, the piping plover's sensory state which triggers this behavioural response is not chained to this or that proximal stimulus. By contrast, the anaerobic bacterium and hypothalamus are not genuine representational systems, since they fail to employ any constancy mechanism in the production of the magnetosomes states and antidiuretic hormone. Actually, both are respectively chained to a specific proximal stimulus, namely, the prevailing magnetic field and the osmolarity level of human blood.

One could object that it is not clear why a distinction between robust-process and actual-sequence explanations in fact reveals a contrast between representational and non-representational states. That is, that the distinction between robust-process and actual-sequence explanations merely demarcates a line between systems that employ and don't employ constancy mechanisms. However, this objection is not fair.[12] After all, this objection can be raised to every demarcation proposal for the limits of intentionality, no matter how

[12] Marc Artiga assesses this objection, but in the end he concludes that "this question is hard to settle" (2016: 422).

appropriate it may be. For every demarcation proposal which specifies whatever property *p* as a minimal condition for intentionality, it is always possible to object that this proposal does not draw a line between representational and non-representational states, but merely a line between *p* states and *non-p* states. Then, the whole debate on the limits of intentionality would become sterile. Here I take a different route. In what follows, I will develop an objection that applies to both the constancy mechanism proposal and the counterfactual robustness argument. It is based on the fact that the distinction between proximal and distal features comes in degree, which gives rise to a serious problem for the constancy mechanism proposal in general and the counterfactual robustness argument in particular.

4 The minimal distance problem

It is a consequence of the counterfactual robustness argument that there is no representation of proximal features. That is, that there is no behavioural pattern prone to be explained by a supposed proximal content intentional explanation. The kernel of the argument is that the distinctive explanatory power of intentional explanations is the account of counterfactual scenarios. It follows that there is no such a thing as proximal content explanation since it is not possible to explain counterfactual cases based on proximal content representations. Only distal content explanations are capable of explaining counterfactual scenarios because constancy mechanisms are required for the state to represent the same distal feature, despite huge proximal stimuli variations. However, there is no variation between the sensory apparatus and the proximal stimulus which could give rise to counterfactual scenarios left to be explained by a proximal content explanation. The huge variation of proximal stimuli between the sensory apparatus and the distal feature constitutes the counterfactual cases that the distal content explanation is capable of explaining. But there is no variation which a proximal content explanation could explain. The conclusion is that the supposed proximal content explanation has no counterfactual robustness.

The defender of the counterfactual robustness argument may justify, from an explanatory point of view, that there is no genuine proximal content explanation in the following way. The positing of a state as a proximal content representation, in order to explain behaviour, makes no explanatory difference, since a non-intentional explanation that posits no representation has the same explanatory power as a supposed proximal content explanation. So, only representations of distal features have distinctive explanatory power and, since only systems with constancy mechanisms can represent distal features, the conclusion is that their employment is a minimal condition for intentionality.

But here a serious problem arises for both constancy mechanism proposal and counterfactual robustness argument—*the minimal distance problem*. The starting point is that the distinction between proximal and distal stimuli comes *in degrees*. It is not clear-cut. There are no strict groups of proximal and distal stimuli because there is no non-arbitrary line which strictly divides proximal and distal stimuli. Evidently, certain stimuli are more proximal than others, in the sense that there is a *hierarchy* of stimuli classified by proximity. But any strict line drawn in order to divide it into two groups—proximal and distal stimuli—is always blurry.

The fact that the distinction between proximal and distal stimuli is not strict raises a problem to both the counterfactual robustness argument and constancy mechanism proposal: what is the *minimal distance* between the system's sensory apparatus and the stimulus in order for a state to genuinely *represent* it? Notice that the employment of a constancy mechanism presupposes that there is *some* distance between the representational state and what is being represented. Otherwise there would be no distance based on which a constancy mechanism would keep the state representing the same feature, despite huge variations in the stimuli reaching the system's sensory apparatus. But what is the extent of this minimal distance? The counterfactual robustness argument gives rise to one response, while the constancy mechanism proposal gives rise to another one. I will assess each response by turn, let us start with the former.

Maybe this is the direct response to this problem. The minimal distance between the stimulus and the token state, for the later to represent the former, is any stimulus but the most proximal one—the most proximal stimulus cannot be represented. The main argument for this response comes from the constancy mechanism proposal itself. No state can represent the most proximal stimulus because in this case no employment of a constancy mechanism is possible. There would be no distance for the state to represent the same feature despite the stimuli variation reaching the system's sensory apparatus. Some minimal distance between the tokening of the state and the triggering stimulus is required for the employment of a constancy mechanism to be possible in the first place. Furthermore, notice that in the case of the most proximal stimulus which triggers a certain state, it would be explanatory idle to posit that the state represents this stimulus because the non-intentional explanation of behaviour would have the same explanatory power. Therefore, the minimal distance between the stimulus and the token state cannot be the most proximal one.

What about the *second* most proximal stimulus? The counterfactual robustness argument gives rise to the following justification for the second most proximal stimulus constituting the minimal distance, in contrast with the most proximal stimulus. The second most

proximal stimulus is the closest stimulus to the sensory apparatus that a given stimulus may be that still gives rise to a robust explanation of behaviour, i.e., an explanation that gives an account of counterfactual situations. The explanation that posits the state as representing the second most stimulus in the stimuli chain gives an account of counterfactual situations. By contrast, an explanation that posits the state as representing the most proximal stimulus cannot give an account of counterfactual situations. Therefore, the latter is not a robust explanation and, thus, not an intentional explanation. This is how the counterfactual robustness argument gives rise to a justification for the minimal distance being the second most proximal stimulus, not the most proximal one.

However, this justification is problematic. At the second most proximal stimulus level the group of counterfactual situations is so small that threatens the viability of the claim that there is a genuine *robust* intentional explanation of behaviour, in contrast with non-intentional explanations that lack such a robust character. Note that according to the counterfactual robustness argument, it is the amount of counterfactual situations, that an explanation is capable of giving an account, which determines whether the explanation is sufficiently robust or not. But if the amount of counterfactual situations is so small, an explanation which gives an account of it is not explanatorily relevant in order to make this explanation genuinely robust.

Here it could be replied that there is at least one counterfactual situation that an intentional explanation gives an account because it posits that the state represents the second most proximal stimulus, while a non-intentional explanation cannot give an account of this counterfactual situation. Hence, in this case there is a small (but nonzero) explanatory purchase. However, it does not follow from the fact that an explanation gives an account of *just one* counterfactual situation that it is *sufficiently* robust. The explanation is not sufficiently robust in order to justify the positing of the state as representing the second most proximal stimulus. Note that the constancy mechanism proposal, as developed by Kim Sterelny, itself requires "sufficient variation" in proximal stimuli between the state and the distal external condition for the first to represent the latter (Sterelny 1995: 261–2). So, according to the constancy mechanism proposal, intentional explanations are sufficiently robust. But once again, how huge should this stimuli variation be to constitute *sufficient* variation for intentionality? The *very specification* of the constancy mechanism proposal makes it clear that it is not whatever variation that is required for intentionality. This is another reason for the conclusion that the indeterminacy of the required amount of stimuli variation is problematic for the viability of the constancy mechanism proposal.

This is what is problematic with the above justification for the second most proximal

stimulus as constituting the minimal distance. Suppose that a stimuli chain triggers the tokening of a certain state in a given system—s^1 is the most proximal stimulus of the chain, s^2 is the second most proximal stimulus, s^3 is the third one and so on. As previously shown, the state cannot represent s^1 because there is no distance between the sensory apparatus and the stimulus. Now consider the hypothesis that the state represents s^2 because an intentional explanation that posits this state as representing s^2 provides an account of the counterfactual situations constituted of variations of s^1. However, the amount of counterfactual situations, constituted by variations of s^1, is so small for the explanation of behaviour that an account of them is not sufficient to make the explanation robust. That is, the positing of the state as representing s^2 fails to make the explanation sufficiently robust, and, according to the counterfactual robustness argument, the minimal distance is not s^2.

What about the hypothesis that the minimal distance is s^3? The same problem arises again. The amount of counterfactual situations constituted by variation of s^1 and s^2 is still small. It is not clear that an explanation which gives an account of these counterfactual situations, by positing the state as representing s^3, is sufficiently robust. Once again, the positing of the state as representational is not explanatorily justified.

What is the lesson we should draw from this whole exercise? It is not clear what is the first stimulus in the causal chain (s^2, or s^3, or, s^4…) for which the positing of the state as representing this first stimulus gives rise to a *sufficiently robust* intentional explanation. That is, an explanation which gives an account of a sufficiently large and relevant amount of counterfactual situations constituted by variations in the intermediate stimuli between the tokening state and the first stimulus in the causal chain. In the absence of some justified criterion, it is indeterminate what is the minimal distance from the sensory apparatus that a stimulus may be for the state to genuinely represent it. This is *the indeterminacy objection* for the response that the counterfactual robustness argument gives rise to the minimal

distance problem.[13]

Let us illustrate the minimal distance problem with concrete cases. On one hand, it is clear that the paramecium's tropistic states are strictly chained to a specific proximal stimulus, namely, light. On the other hand, when the sensory state of the frog detects the fly, it is clear that there are enough intermediate stimuli to constitute a large amount of counterfactual situations required for the sufficient robustness of the intentional explanation. In fact, the frog's visual system employs a *size-distance constancy mechanism*: roughly, the frog responds to the real size of the object, not to the size of retinal stimuli (Ingle & Cook 1977). The paramecium's tropistic state and the frog's sensory state are, respectively, clear cases when there is and when there is not a sufficiently large amount of counterfactual situations required for the robustness of the intentional explanation. Thus, in light of the constancy mechanism proposal, only the frog's sensory state is representational. But now let us assess a third case.

Consider the case of vervet monkeys (Cheney & Seyfarth 1990), as described by Kim Sterelny:

> Adult vervets fear martial eagles. When they see them, their warning and avoidance behaviours are quite distinctive (Cheney and Seyfarth 1990). If there are a *variety of routes* through which eagle recognition and appropriate responses to them are constructed, then we should explain this as learning about eagles and how to avoid them. But if the vervet behavioural changes are cued by *a few highly specific stimuli*, there is nothing about vervet behaviour or change in behaviour that is robustly tracked by a feature of the world which we can take the organism to come to represent. (Sterelny 1995: 260; emphasis added)

This is how the minimal distance problem arises: how *huge* should this variety of routes be

---

[13] Notice that this indeterminacy objection should not be confused with the content indeterminacy problem for naturalistic theories of representation (e.g., teleosemantic and causal-informational theories). The latter is a problem about the determination of the *content* of representational states, in the sense that maybe naturalist theories cannot fully determine the content of representational states. In contrast, the indeterminacy objection is an objection to a specific demarcation proposal of the minimal conditions for the assignment of representational *status* to sensory states, namely, the constancy mechanism proposal. In sum, the indeterminacy objection regards the *representational status* of sensory states, while the indeterminacy problem regards the *representational content* of representational states. So, it does not follow, from the indeterminacy objection, that there is a content indeterminacy for those naturalistic theories that appeal to constancy mechanisms.

in order to the vervet monkey's sensory state represent martial eagles? Sterelny claims that a *few* highly specific stimuli are not enough for genuine representation, but, once again, how much proximal stimuli routes would be enough? In order to make this point clearer, suppose that the vervet monkey's sensory state is triggered by a *median* variety of proximal stimuli routes (i.e., neither huge nor small) coming from the martial eagles. So, there is a median amount of counterfactual situations constituted by such a variety. Is it sufficient for the sensory state to be genuinely representational? This is indeterminate. It is not clear that an explanation which gives an account of these counterfactual situations, constituted by a median variety of routes coming from the martial eagle, is sufficiently robust. The solution to the minimal distance problem requires a *principled criterion* to specify how much proximal stimuli routes is enough for the sensory state to represent the distal condition. Otherwise, the constancy mechanism proposal is still too vague.

Now let us move to the response which the constancy mechanism proposal gives rise to to this problem. The constancy mechanism proposal maintains that the employment of a constancy mechanism is a minimal condition for a certain state to be representational. Accordingly, it is a condition for a state to represent a given feature that it represents the same feature despite *significant* variations of proximal stimuli reaching the system's sensory apparatus. Based on this proposal, it could be argued that since there are variations of s^1 between the sensory apparatus and the second most proximal stimulus s^2, then the room is open for the employment of a constancy mechanism and, so, s^2 is the minimal distance. However, this move is problematic. In this case, the only possible variation of proximal stimuli reaching the sensory apparatus is constituted of variations of s^1. Is this *enough of a variation* for the state to genuinely represent s^2? Maybe not. It is not clear that there is sufficient variation, maybe there is genuine representation only when there is *more* variation. The same problem arises for the hypothesis of s^3 as the minimal distance. The only possible variation of proximal stimuli is constituted of variations of s^1 and s^2. Is this now enough variation for the state to genuinely represent s^3? Once again, it is not clear that now there is enough variation for the state to genuinely represent this stimulus.

The conclusion is that the minimal distance between the sensory apparatus and the stimulus in order to have enough intermediate stimuli variation—i.e, enough intermediate stimuli variation for the relevant sensory state to be genuinely representational—is *indeterminate*. In the end, it is an *open question* which stimulus in the causal chain is the first, such that there is enough stimuli variation for the sensory state to represent it.

The indeterminacy objection threatens both responses to the minimal distance problem, the response based on the counterfactual robustness argument in particular and the

response based on the constancy mechanism proposal in general. Its neutralization requires a *non-arbitrary criterion* to determine the extent of the minimal distance from the system's sensory apparatus that a stimulus may be in order for the state to genuinely represent it. If the minimal distance is indeterminate, then the constancy mechanism proposal and the counterfactual robustness argument are in dangerous. In the absence of a solution to the minimal distance problem, the viability of the constancy mechanism proposal is threatened since it depends on a strict distinction between proximal and distal stimuli. That is, if this distinction is indeterminate, it follows that the constancy mechanism proposal's assumption that there is only representation of distal features is also indeterminate and, so, the constancy mechanism proposal is in trouble.

Now it could be objected that the indeterminacy objection runs the risk of being a slippery slope argument. This risk is illustrated by this analogy. Baldness is a vague notion: it is not fully determinate the point in which someone that is not bald becomes bald. How many strands of hair should a non-bald person lose to become bald? There is no strict turning point, but a continuum between full baldness (no hair) and full non-baldness (full head of hair). Somewhere in the middle lies the limits of baldness. Now suppose that a supermodel agency prohibits bald male supermodels. It would be nonsense to claim that this is not a viable condition for supermodel selection since it does not provide a principled way of determining exactly when someone non-bald turns bald. After all, baldness is a useful notion and it works quite well for the selection of supermodels. Analogously, the constancy mechanism proposal cannot be rejected because it does not provide a principled way of determining the extent of the minimal distance from the system that an external condition may be for the state to represent it. Rather, there is a continuum between representationality and non-representationality in which there are clear cases of representational states and clear cases of non-representational states. There is nothing problematic about it.

In fact, there is nothing problematic with the notion of baldness in its everyday use to distinguish bald from non-bald people. Nevertheless, things are different in respect of the debate on minimal conditions for intentionality. Here we are looking for minimal conditions for intentionality to develop a *categorization* of representational and non-representational states, not a continuum between these states. In order to develop such a categorization, the categories of a representational state and non-representational state cannot suffer from this *level of indeterminacy*, precisely because it threatens the viability of the delivered distinction between representational and non-representational states. In short, it is not a solution to the minimal distance problem to just claim that there is a continuum between representational and non-representational states.

Consider the constancy mechanism proposal. The employment of the constancy mechanism presupposes some distance between the representational state and what is represented. But it is indeterminate what is the minimal distance from the state that a stimulus may be for there to be enough intermediate stimuli variation. That is, it is indeterminate what is the required *sufficient* stimuli variation for the state to be representational. The distinction between sufficient and insufficient stimuli variation comes in degree—it does make sense to talk about more or less stimuli variation. So, in the delivered demarcation, the categories of representational and non-representational states also come in degree. But this is problematic: these categories are clear-cut. A non-arbitrary criterion is, hence, required to determine what is the sufficient amount of stimuli variation for the state to be genuinely representational. This is precisely what the constancy mechanism proposal misses.

Here one may reply that one should not expect the categories of representational and non-representational states to be *strictly* determinate. Rather, there is certain vagueness on the limits of these categories. So, the indeterminacy of the minimal distance from the state that a stimulus may be, for the state to genuinely represent it, is not problematic for the constancy mechanism proposal.[14] Behind this reply lies the following problem: *how strict* should the demarcation between the categories of representational and non-representational states be in the first place?

Here one may claim that the category of a representational state comes in degree, just like most biological categories (as well as other natural categories) also come in degree. There are paradigmatic and marginal cases of representational states, just like paradigmatic and marginal cases of biological states (Godfrey-Smith 2009: 6–7, 41–3). Consider the biological function of a given biological trait. According to the aetiological theory of biological function (e.g., Millikan 1989), this trait has the function F iff it was *selected* to have the effect F (e.g., the heart has the function to pump blood because it was selected to pump blood). However, biological selection is not a strict category, it comes in degree. Note that the first trait that pumped blood, in virtue of a genetic mutation, did not have the function to pump blood, and neither its first generation descendants. After all, some time and generations are required to definitely establish the fact that a certain trait was selected to pump blood. In sum, biological section is not a clear-cut category and, hence, assuming the aetiological conception, neither is biological function. The same line of reasoning applies to life and other biological categories. So, why take the category of representational states

[14] I thank the referees for pressing this issue.

to be strict, in opposition with these biological categories? Is there any reason for assuming that the category of representational states is clear-cut and, so, does not come in degree?

Talk about representations admits no distinction of degree on the representational statuses because it makes no sense to talk about one state as being *more representational* than another. After all, one cannot conceive of a sensory state as *partly* representing something, *partly as not* (i.e., a sensory state whose status is partly representational, partly non-representational). Of course, one may wonder whether a given state is representational or not, but *once it is established* that it is representational, it makes no sense to say that this state is more representational than another state. So, the distinction between representational and non-representational states is clear-cut.

The problem with this argument is that it appeals to our *intuitions* of what is a representational state. So, there is always room for someone to object that it is *not* intuitive to claim that the category of representational states is clear-cut. Furthermore, as I have formerly established, in this paper I restrict my assessment of the constancy mechanism proposal for intentionality in light of the *explanatory criterion* and, so, not in light of an intuitive criterion. Thus, here I will not follow the intuition-based argumentative strategy presented above.[15] Nevertheless, it is not required to establish a definitive proof, that the category of a representational state is clear-cut, in order to show that the indeterminacy objection jeopardises the viability of the constancy mechanism proposal.

The constancy mechanism proposal assumes that representational and non-representational states are different categories, and, so, that there is no continuum between them. After all, if there is just a continuum between representational and non-representational states, it makes no sense to establish minimal conditions for intentionality in order to draw a demarcation to distinguish representational from non-representational states (furthermore, the aforementioned objection of liberality against naturalist theories of mental representation assumes that there is no such continuum). But how strict are these categories?

In the absence of a principled criterion to specify the extent of the minimal distance from the system's sensory apparatus that a stimulus may be for the state to genuinely represent it, it follows that the constancy mechanism proposal delivers a continuum between representational and non-representational states, not the establishment of different cate-

---

[15] For an overview of the role of pre-theoretic intuitions on the assessment of naturalistic theories of representational states see Ramsey (2007: 1–37).

gories. But here the defender of the constancy mechanism proposal may argue that actually what this proposal delivers is not a continuum, but a *vague categorization*: the delivered categories of representational and non-representational states are *vague*, just like biology delivers vague biological categories. But how to establish whether the constancy mechanism proposal delivers a *continuum distinction* between representational and non-representational states or a *vague categorization*? In what follows, I will argue that the constancy mechanism proposal delivers a continuum between representational and non-representational states, not a vague categorisation.

Note that the debate on the minimal conditions for intentionality is focused on *borderline cases* of representational states. That is, the demarcation proposal is supposed to draw a distinction between the categories of representational and non-representational states among these borderline cases. So, the problem is whether, *among borderline cases*, the constancy mechanism proposal delivers a continuum between representational and non-representational states or a vague categorisation. What happens is that the *sufficient amount* of intermediate stimuli variation for the sensory state to be genuinely representational, required by the constancy mechanism proposal, is itself *continuous*. That is, there is a *continuum* between, on one hand, the first stimulus in the causal chain for there to be sufficient stimuli variation required for the sensory state to represent this stimulus, and, on the other hand, the next stimulus in the causal chain in which the stimuli variation is not sufficient anymore. So, it follows that the constancy mechanism proposal establishes a *continuum* between *borderline* representational and non-representational states, not a vague categorisation. Once again, a principled criterion is required to specify the required amount of intermediate stimuli variation for the sensory state to genuinely represent the relevant external stimulus. In the absence of such a criterion, there is just a continuum between borderline representational and non-representational states, rather than a proper categorization. In sum, it is indeterminate the amount of sufficient stimuli variation for the state to be representational.

Finally, the defender of the constancy mechanism proposal may reply that the vagueness of the required *sufficient amount* of intermediate stimuli variation does not imply that the proposal delivers a continuum between representational and non-representational states. What is the required sufficient stimuli variation is indeed vague, but the constancy mechanism proposal does not deliver a continuum between representational and non-representational states. However, it strikes me that, up to this point in the debate, the onus of showing that the constancy mechanism proposal delivers a vague categorization—not

a continuum—is on the defender of this proposal.[16] After all, as previously shown, there is a continuum between *sufficient* and *insufficient* intermediate stimuli variation, and the constancy mechanism proposal is fundamentally grounded on this continuum.

5 Conclusion

In this paper, I developed the minimal distance problem for both the constancy mechanism proposal and the main argument in support of it—the counterfactual robustness argument. What is the lesson one should draw from this problem concerning the debate on the problem of demarcation in general and the constancy mechanism proposal in particular? I think it is that the appropriate proposal for minimal conditions for intentionality should not rely on a distinction between proximal and distal stimuli *unless* this proposal has a solid and principled response to the minimal distance problem. As a result, the indeterminacy objection would then constitute no threat to this proposal. But it is hard to see how this is a viable way out for the constancy mechanism proposal defender, since they draw the limits of intentionality appealing *only* to the constancy mechanism condition. Maybe another minimal condition is required to draw these limits. However, it is not easy to conceive what this further minimal condition would be, especially because it cannot be ad hoc or arbitrary, but must be plainly justified.[17]

---

[16] Schulte has recently developed the two-track account (TTA) of perceptual constancy. He recognises that "[o]ne might object that TTA does not draw a sharp line between sensory systems that employ constancies and those that do not, due to the vagueness of expressions like 'sufficiently uniform' and 'approximate'. In my view, however, this [is] a virtue of the account rather than a vice: it just reflects the fact that perceptual constancy is a gradual phenomenon" (2021: 14). Maybe perceptual constancy in fact comes in degree, but it does not follow from it that representations also come in degree. Even if one assumes TTA, the resulting constancy mechanism proposal still delivers a continuum between representational and non-representational states. After all, TTA delivers a continuum between systems that employ and do not employ constancy mechanisms. Here one can bite the bullet that the TTA-style constancy mechanism proposal delivers a continuum between representational and non-representational states, but then one has the onus of showing why this is not problematic.

[17] In my PhD thesis, I have developed a dual proposal of minimal conditions for intentionality and argued that this proposal is not threatened by the minimal distance problem, in contrast with the constancy mechanism proposal. See Souza Filho (2018: 188–94).

Acknowledgements

This paper is the result of my PhD thesis at King's College London and my postdoctoral research at the Federal University of Rio de Janeiro. I would like to thank David Papineau, Roberto Horácio de Sá Pereira, Jake Wojtowicz, Carolyn Price and Barry Smith for their help. My gratitude to the participants of the "Segundo Workshop PPGLM/UFRJ-UNICAMP de Filosofia Contemporânea" for their comments on a talk in Rio based on this paper, as well as to Pedro Merlussi for co-organizing the workshop. I would like to thank the CAPES Foundation (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), FAPERJ Foundation (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro) and CNPq Foundation (Conselho Nacional de Desenvolvimento Científico e Tecnológico). I would also like to thank two referees for their constructive comments.

Sérgio Farias de Souza Filho
Federal Rural University of Pernambuco
Department of Social Sciences, Area of Philosophy,
Federal Rural University of Pernambuco.
Rua Dom Manuel de Medeiros,
s/n, 52171900,
Recife, PE, Brazil.
sergio.souzafilho@ufrpe.br

References

Artiga, Marc. 2016. "Liberal representationalism : a deflationist defense". *Dialectica 70*(3): 407–30.

Bates, Elizabeth, Camaioni, Luigia and Volterra, Virginia. 1975. "The acquisition of performatives prior to speech". *Merrill-Palmer Quarterly of Behavior and Development 21*(3): 205–26.

Burge, Tyler. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.

Call, Joseph and Tomasello, Michael. 2007. *The Gestural Communication of Apes and Monkeys*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cheney, Dorothy and Seyfarth, Robert. 1990. *How Monkeys See the World: Inside the Mind of Another Species*. Chicago: Chicago University Press.

Cohen, Jonathan. 2015. "Perceptual constancy". In *The Oxford Handbook of Philosophy of Perception*. Mohan Matthen (ed.). Oxford: Oxford University Press.

Dretske, Fred. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Fodor, Jerry. 1986. "Why paramecia don't have mental representations". *Midwest Studies in Philosophy 10*(1): 3–23.

Fodor, Jerry. 1987. *Psychosemantics*. Cambridge, MA: MIT Press.

Goldstein, E. Bruce. 2010. "Constancy". In *Encyclopedia of Perception*. E. Bruce Goldstein (ed.). Los Angeles: SAGE Publications.

Godfrey-Smith, Peter. 2009. *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.

Hilbert, David. 2005. "Color constancy and the complexity of color". *Philosophical Topics 33*(1): 141–58.

Ingle, David and Cook, Jefferey. 1977. "The effect of viewing distance upon size preference of frogs for prey". *Vision Research 17*(9): 1009–13.

Jackson, Frank and Pettit, Philip. 1992. "In defense of explanatory ecumenism". *Economics & Philosophy 8*(1): 1–21.

Millikan, Ruth Garrett. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.

Millikan, Ruth Garrett. 1989. "In defense of proper functions". *Philosophy of Science 56*(2): 288–302.

Millikan, Ruth Garrett. 2004. *Varieties of Meaning*. Cambridge, MA: MIT Press.

O'Keefe, John and Nadel, Lynn. 1978. *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.

Palmer, Stephen. 1978. "Fundamental aspects of cognitive representation". In *Cognition and Categorization*. Eleanor Rosch & Barbara Lloyd (eds.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Papineau, David. 1993. *Philosophical Naturalism*. Oxford: Basil Blackwell.

Papineau, David. 2016. "Teleosemantics". In *How Biology Shapes Philosophy: New Foundations for Naturalism*. David Livingstone Smith (ed.). Cambridge: Cambridge University Press.

Ramsey, William. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press.

Rescorla, Michael. 2013. "Millikan on honeybee navigation and communication". In *Millikan and Her Critics*. Dan Ryder & Justine Kingsbury & Kenneth Williford (eds.). Malden, MA: Wiley-Blackwell.

Ristau, Carolyn. 1991. "Aspects of cognitive ethology of an injury-feigning bird, the piping plover". In *Cognitive Ethology: the Minds of Other Animals*. Carolyn Ristau (ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schulte, Peter. 2015. "Perceptual representations: a teleosemantic answer to the breadth-of-application problem". *Biology & Philosophy 30*(1): 119–36.

Schulte, Peter. 2021. "The nature of perceptual constancies". *Philosophy and Phenomenological Research 103*(1): 3–20.

Shea, Nicholas. 2007. "Consumers need information: supplementing geleosemantics with an input condition". *Philosophy and Phenomenological Research 75*(2): 404–35.

Souza Filho, Sérgio. 2017. "The constancy mechanism proposal for the limits of intentionality". In *The Philosophy of Perception and Observation: Contributions of the 40th International Wittgenstein Symposium*. Christoph Limbeck-Lilienau & Friedrich Stadler (eds.). Kirchberg am Wechsel: Austrian Ludwig Wittgenstein Society.

Souza Filho, Sérgio. 2018. *Naturalising intentionality: A Teleological Approach*. PhD thesis, King's College London. Retrieved September 10, 2022. https://kclpure.kcl.ac.uk/portal/files/103712183/2018_Farias_de_Souza_Filho_S_rgio_1465852_ethesis.pdf.

Souza Filho, Sérgio. 2022. "A dual proposal of minimal conditions for intentionality". *Synthese 200*(115): 1–22.

Sterelny, Kim. 1995. "Basic minds". *Philosophical Perspectives 9*: 251–70.

von Frisch, Karl. 1967. *The Dance Language and Orientation of the Bees*. Cambridge, MA: Harvard University Press.