

# Teleosemantics and Indeterminacy

Manolo MARTÍNEZ (mail@manolomartinez.net)

## Abstract

In the first part of the paper, I present a framework for the description and evaluation of teleosemantic theories of intentionality, and use it to argue that several different objections to these theories (the various *indeterminacy* and *adequacy* problems) are, in a certain precise sense, manifestations of the same underlying issue. I then use the framework to show that Millikan’s biosemantics, her own recent declarations to the contrary notwithstanding, presents indeterminacy.

In the second part, I develop a novel teleosemantic proposal which makes progress in the treatment of this family of problems. I describe a procedure to derive a (unique) homeostatic property cluster [HPC] from facts having to do with the properties that a certain indicator relied on, in the events leading to its fixation in a certain population. This HPC is the one that should figure in the content attribution to the indicator in question.

## 1 Introduction

*Teleosemantic* accounts of mental content (e.g. Millikan 1984; Papineau 1998; Neander 2013) propose that the content of mental states depends on causal-historical properties of these, or appropriately related, states, which warrant attributions of function, in the teleological sense. The appeal to functions provides a promising way of unpacking the normativity implicit in the idea of *misrepresentation*—it would be related, in ways that differ from one teleosemantic account to another, to the normative notion of *malfunctioning*—, and this is considered a major milestone in the project of naturalising intentionality. To complete this project, an appropriately naturalistic account of (mal)function has to be provided. There are several promising candidates (see, e.g., Wright 1976; Price 1998; Millikan 2002, and section 2 of this paper).

While teleosemantics is widely regarded as one of the most attractive approaches to the naturalisation of intentionality, many, perhaps all, teleosemantic theories on offer are subject to an embarrassing *indeterminacy problem*: they are not able to make univocal predictions as regards the content of mental states. This problem has been around for at least twenty-five years, and, while it was ardently discussed some fifteen years ago, it has largely fallen out of fashion, with next to no new contributions as of late. I suspect that, as is sometimes the case in philosophy, this has not happened because researchers feel that the issue

has been solved but, rather, because there is an impression that everybody has said what they had to say and a stalemate has been reached. This paper is an attempt at unblocking the situation.

I will follow the usual strategy: first, in a negative part, I give reasons to think that, some declarations to the contrary notwithstanding, the indeterminacy problem and some related difficulties are still very much open. Then, in the positive part, I propose a solution to these difficulties.

It is not possible for a piece of the dimensions of a journal article to establish that *no* teleosemantic theory has ever solved the indeterminacy problem. I will, thus, take the following approach to making my case. I first briefly introduce a model and some terminology that will allow us to see that (and in which sense) many problems with teleosemantic accounts are, in fact, different manifestations of the same underlying issue—sections 2 and 3. Then, in section 4, I use this framework to discuss the best known (and, arguably, the best *simpliciter*) teleosemantic theory of intentionality: Millikan’s *biosemantics* and her latest discussion—in Millikan (2009), drawing from the notion of natural local information as developed in Millikan (2004)—of the so-called *output problem*. The hope is that the more general discussion of sections 2 and 3, together with its application to a detailed and sophisticated concrete proposal, will help the reader see in which way other teleosemantic strategies fall prey to indeterminacy-related difficulties.

In the subsequent positive discussion I offer a general diagnosis of why teleosemantics fails in securing univocity, and provide a way around the difficulty identified. In a nutshell, theorists have tried to solve indeterminacy by providing recipes that *choose* one among many content candidates. In section 5 I propose that, instead, the solution is to *construct* the content of the target mental state out of the many candidates provided by classic teleosemantic accounts. In subsection 5.3 I formulate a detailed content-attribution recipe, for the model presented in section 2, that encapsulates this solution. One of the main lessons of section 5 is that much of the work necessary for solving the indeterminacy problem needs to be done in the theory not of representation but of the entities *represented*—for very simple contentful states, the theory of real kinds.

Section 6 shows how to apply the recipe. Finally, section 7 offers some concluding remarks, and discusses whether the resulting account is an account of *content* as opposed to some *ersatz* substitute thereof.

## 2 A Model, and Some Terminology

In the first sections, and in order to keep the discussion as focused as possible, I will be concerned with the following simple model:

1. A token mechanism of type M is inside the brain of a creature, Democritus. Mechanisms of type M can be in one out of two possible states: *on* or *off*, and they go *on* only if (but whenever) the property *Being orange and*

*round* is instantiated near its possessor. M's going *on* causes its possessor to grab and eat the orange round thing near him<sup>1</sup>.

2. Most (but not all) orange round things near Democritus are peaches, and they have been for generations. Eating peaches increases the fitness<sup>2</sup> of creatures such as Democritus.
3. M was mutated into existence many generations ago, but it is heritable, and the whole population of which Democritus is a member has a token of M these days, partly as a result of the fitness increase provided by M.

I will discuss content attributions that teleosemantic theories provide for the mental state consisting of Democritus's token of M's being *on*—I will also use “*M<sub>on</sub>*” as a name for this mental state.

Our intuitions about the content of such simple mental states are not extremely fine grained, and, at any rate, do not seem to zero in on any particular content attribution<sup>3</sup>. On the other hand, we do seem to have coarse-grained intuitions that *rule out* candidates such as, say, *There is a fitness-enhancing orange thing here* and do not rule out as *prima facie* plausible others such as *There is a peach here*, or *There is food here*. These (coarse-grained) intuitions create the possibility of a certain kind of misbehaviour on the part of a teleosemantic theory: warranting one of the content attributions that the former had ruled out. In general, that is, our teleosemantic theory of content should not be subject to the *adequacy problem*:

**Adequacy:** A teleosemantic theory is subject to the *adequacy problem* if it warrants content attributions that our intuitions had ruled out.

A more catastrophic misbehaviour is showing indeterminacy:

**Indeterminacy:** A teleosemantic theory is subject to the *indeterminacy problem* if it warrants more than one content attribution in cases in which there is just one correct one.

---

<sup>1</sup>We can think of the type M as being causally and historically individuated: something is an M iff it derives in the appropriate way from other Ms and has the causal profile outlined in the main text.

<sup>2</sup>I will simply sidestep the very interesting philosophical issues surrounding the concept of fitness—for an introduction see Sober (2000); also Godfrey-Smith (2009). For the purposes of this discussion, we may suppose that the fitness of an individual is its actual number of offspring, and that eating peaches directly causes an increase in such number. We may also assume that creatures such as Democritus are asexual.

<sup>3</sup>The wide variety of candidates for the content of frog mental states in the teleosemantic literature attests to this. An informative, and entertaining, list of such candidates can be found in Neander (2006).

That is, while probably there are genuine cases of content indeterminacy<sup>4</sup>, it is surely the case—at any rate I will assume that it is—that many mental states have only one content and, thus, there is only one correct content attribution for them. The theory to be defended in the second part of this paper is aimed at making progress in both the indeterminacy and the adequacy problems, so characterised.

In the following subsection, I introduce the main teleosemantic insight, and some important terminology, by presenting a simple version of teleosemantics, loosely based on the theory Dretske defended in his (1988).

## 2.1 A Simple Teleosemantic Proposal

A *representational system*, according to Dretske is

any system whose *function* it is to *indicate* how things stand with respect to some other object, condition or magnitude. (Dretske 1988, p. 52, my emphasis)

The idea is that one of the main building blocks of content is the relation of *indication*, or not-by-chance covariation, between the contentful state and the condition that gives its content. The other main, and the paradigmatically teleosemantic, building block is provided by the idea that content does not depend on a mechanism *actually* indicating thus and so, but on its having *the function to so indicate*. It is this appeal to functions that makes room for misrepresentation: something may indicate the presence of Fs, say, without having the function of doing so; or fail to indicate it when it is supposed to.

I will use *indication* in the following way:

**Indication:** A mechanism M indicates the presence of Fs iff

**I1:**  $P(on|F) > P(on)$ <sup>5</sup> and

**I2:** The difference in probabilities in I1 is not a matter of chance.

That is, iff the probability of M being *on* conditional on F being instantiated near M is higher than its unconditional probability, and this inequality is not a matter of chance. Many different relations between M's being *on* and F satisfy I2: M's going *on* causing the presence of F, vice versa, a third event causing both, etc.

---

<sup>4</sup>For example, if, as a popular account has it, vagueness is semantic indecision (Lewis 1986), there might be cases in which there is legitimate indeterminacy as to what content a certain mental state has. The teleosemantic accounts I will be presently discussing yield indeterminacy that goes far beyond the one induced by vagueness—indeterminacy, for example, that remains once every vague term has been substituted by an admissible precisification. It is *this* kind of indeterminacy that is objectionable.

<sup>5</sup>Or, equivalently,  $P(F|on) > P(F)$ . The formulation in the main text lends itself more aptly to the formulation of a general indication profile—see below.

It will be useful to introduce the notion of an *indication profile*, which gives a fuller picture of M’s behaviour in relation to F—see also Godfrey-Smith (1991); Godfrey-Smith (1996). M’s *indication profile for Fs* is:

$$IP_F = \begin{bmatrix} P(on|F) & P(off|F) \\ P(on|\neg F) & P(off|\neg F) \end{bmatrix}$$

where  $P(on|F)$  records the probability of  $M_{on}$  when there is an F around—what signal-detection theorists call a *hit*.  $P(off|F)$  records the probability of M’s being *off* when there is an F around—a *miss*.  $P(on|\neg F)$  is the probability of a *false alarm* and  $P(off|\neg F)$  that of a *correct rejection*.

Hits, misses, false alarms and correct rejections are, each, coupled with a different fitness value: usually, the fitness value of a hit is high, that of a miss is low, etc. Hence M’s *fitness matrix from the indication of Fs*:

$$FM_F = \begin{bmatrix} w(on|F) & w(off|F) \\ w(on|\neg F) & w(off|\neg F) \end{bmatrix}$$

where I use  $w(on|F)$  for the fitness value of a hit,  $w(off|F)$  for the fitness value of a miss, etc. Finally we have M’s *fitness contribution from the indication of Fs*—the average of fitness values, weighed by the indication values<sup>6</sup>:

$$FC_F = P(F) \cdot [wP(on|F) + wP(off|F)] + \\ + P(\neg F) \cdot [wP(on|\neg F) + wP(off|\neg F)]$$

where  $wP(A) = w(A) \cdot P(A)$ .

On the other hand, I will use *function* in the following way:

**Function:** A token of the mechanism M in a subject S has the function of indicating the presence of Fs iff

**EF1:** According to *Indication*, mechanisms of type M have indicated the presence of Fs around their possessor in (a sufficient number of) S’s recent ancestors.

**EF2:** The fitness contribution from the indication of Fs of the token of M in S’s recent ancestors has been positive (that is,  $FC_F > 0$ ), and this is part of an explanation of the fact that S has a token of M.

**EF3:** (A sufficient number of) the conditional probabilities in  $IP_F$  are not a matter of mere chance.

**EF4:** (A sufficient number of) the fitness values in  $FM_F$  are not a matter of mere chance.

---

<sup>6</sup>This definition of fitness contribution relies on a number of idealisations. For example, that fitness values are additive and, e.g., every miss is as deleterious as any other. The purpose of the semi-formalism I’ve just introduced is to stress the common features in many problems with teleosemantic accounts—in a nutshell, that they boil down to the relabelling of hits as correct rejections, etc. and the reallocation of the fitness value of the very same episodes to one or the other labels—, and the crudeness of the model helps avoid distracting factors. Of course, for other purposes —e.g., the estimation of population sizes—more realistic models will be needed.

This is a straightforward application of the *etioloical theory of functions* (Wright 1994; Godfrey-Smith 1994; Price 1998; Millikan 2002), to the indication case that interests us. The main insight associated with the etioloical approach is that attributing a device with a function is to provide a certain kind of explanation of its existence:

[S]aying that the function of X is Z is saying at least that X is there *because* it does Z. (Wright 1994, p. 39)

EF1 and EF2 say that M's having indicated the presence of Fs helps explain the existence of S's token of M. EF3 and EF4 ensure that this is not just a happy coincidence.

Finally, *Indication* and *Function* provide a rigorous interpretation for the following, broadly Dretskean teleosemantic proposal:

**Basic Teleosemantics:** A token of  $M_{on}$  has the content *There is an F here* if M has the function of indicating the presence of Fs.

*Basic Teleosemantics* provides the bare-bones version of two main teleosemantic themes: the presence of informational relations between representeds and representations, and the reliance on teleofunctions. It is also subject to particularly stark versions of the indeterminacy and adequacy problems, which will make it a useful prop in the discussion of these difficulties.<sup>7</sup>

I will discuss now adequacy and indeterminacy in connection with *Basic Teleosemantics*. Immediately after that, I show how a host of other much more sophisticated proposals are subject to the same problems, for the same reasons.

### 3 The Indeterminacy Problem

What's the content of Democritus's  $M_{on}$  according to *Basic Teleosemantics*? Consider first whether M has the function of indicating the presence of peaches:

**EF1:** *Ex-hypothesi*, mechanisms of type M have indicated the presence of peaches around their possessor in (a sufficient number of) Democritus's recent ancestors. This is so because Ms indicate the presence of orange, round things, and for generations these things have mostly been peaches.

**EF2:** The fitness contribution from the indication of peaches of the token of M in Democritus's recent ancestors has been positive (that is,  $FC_{peach} > 0$ ),

---

<sup>7</sup>I should stress again that *Basic Teleosemantics* is a expository device and that, while distinctly Dretskean in important respects, it is certainly not Dretske's theory. At least not his theory for the content of *beliefs* as developed in Dretske (1988), where he suggests that learning processes help screen off the indeterminacy present in *Basic Teleosemantics*. It is, on the other hand, very close to Dretske's account of the content of what he, in the same book, calls *Type III Representations*.

I will not discuss Dretske's views on content determination through learning, which I find less appealing than, and have had less influence than, other proposals I do discuss.

and this is part of an explanation of the fact that Democritus has a token of M: M has helped Democritus’s ancestors grab and eat peaches and this has been an important advantage against other conspecifics which lacked a token of M.

**EF3:** M complies with *Indication*; thus  $P(on|F)$  (and, we may assume,  $P(on|\neg F)$ ) are not a matter of chance. We may further assume that M’s being *off* indicates the absence of peaches because there are very few non-orange and non-round peaches—and this is not by chance, either: there are causal processes that make most peaches orange and round. This takes care of  $P(off|\neg F)$ —and, we may assume,  $P(off|F)$ .

**EF4:** Peaches are nutritious for the likes of Democritus; and, again, it is not by chance that the properties *Being a peach* and *Being nutritious for the likes of Democritus* covary positively. This takes care of  $w(on|F)$ . We may assume that this is the value in the fitness matrix that has been instrumental in increasing the fitness of Democritus’s family line above that of its conspecifics without M.

So, M has the function of indicating the presence of peaches and, so, *Basic Teleosemantics* warrants a content attribution of *There is a peach here*, as we wished.

But Ms have also indicated instantiations of the property *Being orange and round* around their possessors, and, unfortunately, this explains equally well the fact that there are still Ms around: most orange round things near possessors of a token of M have been peaches, and eating peaches is fitness-conducive. An analogous exercise will show that M has the function of indicating the presence of orange round things, and thus that *Basic Teleosemantics* warrants a content attribution of *There is an orange, round thing here*: content indeterminacy. In an analogous way it may be shown that *Basic Teleosemantics* warrants indefinitely many content attributions, involving orange round things, peaches, food, good stuff for digestive systems, etc.

Slightly more formally, there are very many properties  $F_1, F_2, \dots, F_n$  such that M indicates their presence, and such that, for any two of them,  $FC_{F_i} = FC_{F_j}$ . For example, if Democritus’s fitness matrix from the indication of peaches is:

$$FM_{peach} = \begin{bmatrix} w(on|peach) & w(off|peach) \\ w(on|\neg peach) & w(off|\neg peach) \end{bmatrix}$$

it can be shown (by trivial, if tedious, probability calculus) that the fitness values from the indication of orange round things are:

$$w(on|ort) = \frac{1}{P(on|ort)} \left[ P(peach|ort) wP(on|peach) + P(\neg peach|ort) wP(on|ort \wedge \neg peach) \right]$$

$$w(off|ort) = \frac{1}{P(off|ort)} \left[ P(peach|ort) wP(off|peach) + P(\neg peach|ort) wP(off|ort \wedge \neg peach) \right]$$

$$w(on|\neg ort) = w(on|\neg peach)$$

$$w(off|\neg ort) = w(off|\neg peach)$$

Where *ort* stands for *There is an orange round thing here* and *peach* stands for *There is a peach here*. This, together with the fact that, *ex-hypotesi*,

$$IP_{ort} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

can be shown to entail that  $FC_{peach} = FC_{ort}$ . This is the result that underlies *Basic Teleosemantics*'s failure in securing content univocity: M is equally fitness contributing regardless of whether we describe it as an indicator of peaches or an indicator of orange round things.

### 3.1 Related Difficulties

Philosophers interested in the indeterminacy problem (e.g., Agar 1993; Neander 1995; Rowlands 1997; Price 1998) have put forward two closely related worries:

**The Input Problem.** First, what Price (op. cit.) calls *the input problem*: we can formulate a *lowest* content attribution, finding the property L such that the indication profile for Ls,  $IP_L$ , is as close to the 2x2 identity matrix  $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  as possible—we may call L the *lowest property*. Between the lowest content attribution *There is an L around* and some (let us call it) “natural” attribution such as *There is a peach around* there lie indefinitely many attributions.

A teleosemanticist might try to screen off the indeterminacy in *Basic Teleosemantics* by defending that the right attribution for a state is the lowest, with, maybe, some additional constraints. This approach gives rise to what Rowlands (op. cit.) calls *stimulus-based* teleosemantic accounts, Papineau (op. cit.) calls *producer* teleosemantics, and Neander (op. cit.) calls *low church* teleosemantics.<sup>8</sup> The input problem is the problem of not being able to stop your theory

<sup>8</sup>(Neander 2013) is her most recent defence of this approach. In this piece, Neander circumscribes the scope of her approach to “sensory representations”, as described by neuroethologists. It is unclear whether the very low content attributions she endorses are objectionable in this special context. On the other hand, she also expresses the belief that her preferred brand of *informational teleosemantics* is the right kind of foundation for building a theory of “much more sophisticated representational contents” (op. cit., p. 36). The very low attributions that Neanderian teleosemantics recommends would probably be counterintuitive in the case of, say, human beliefs, but it is difficult to adjudicate this issue, as Neander has so far provided no indication of how her theory extends to the sophisticated cases.



from endorsing ever lower, to the point of totally implausible, content attributions. For example, endorsing contents that make reference to retinal shadows, activation of groups of neighbouring neurons, etc. which would be implausibly low (because implausibly proximal).

If your theory is able to single out just one among the very low content attributions, then your input problem is a variety of the adequacy problem: the content attribution you endorse is, simply, implausible. If, on the other hand, there is no clear candidate for *the* lowest content attribution that your theory endorses, you are worse off: your input problem is a variety of the indeterminacy problem, because your theory cannot decide among the different candidates to be the lowest content attribution.

**The Output Problem.** Second, the complementary *output problem*. We can formulate a *highest* content attribution, by looking for the highest property, H: the property such that the fitness matrix from the indication of Hs,  $FM_H$ , has highest diagonal values and lowest anti-diagonal values. Exactly what function should be maximised ( $|w(on|F) + w(off|-F)| - |w(off|F) + w(on|-F)|$ , or what) is unclear, but the suggestion is that the highest content attribution involves the property that *in fact* accounts for the fitness of the possessor of M: it is not enough that M helps Democritus eat peaches; they must also be *nutritious* peaches, *far-from-predators* peaches, *not-covered-with-poison* peaches, etc. for the possessor of M to improve its fitness thanks to it. In the limit, they must be, simply, good for Democritus. All of these conditions for success correspond to a different proposal regarding the content of  $M_{on}$ .

Trying to secure univocity by defending the highest content attribution as the right one (again, with some possible additional constraints) gives rise to what Rowlands (op. cit.) calls *benefit-based* teleosemantic accounts, Papineau (op. cit.) calls *consumer* teleosemantics, and Neander calls *high church* teleosemantics. The most popular brands of teleosemantics are at least extensionally equivalent with some version of consumer teleosemantics as just characterised. In particular, in the following section, I will show that this is the case with Millikan's biosemantics and that it is, indeed, subject to the output problem, at least in cases such as Democritus's. Again, depending on the actual details of your consumer teleosemantics, the particular output problem you face may be a variety of the adequacy problem or of the indeterminacy problem.

What emerges from this discussion is a continuum of content attributions (see figure 1), from lowest to highest, depending on the choices we make for Indication Profile or Fitness Matrix.<sup>9</sup> This was to be expected: the critical parameter behind selection is the Fitness Contribution of a state, which underdetermines the Indication Profile - Fitness Matrix pair that is needed to fix a content attribution. This extra degree of freedom gives rise to the multiple

---

<sup>9</sup>More carefully: although, mathematically, there is a real continuum of Indication Profile - Fitness Matrix pairs, nothing ensures that, for each  $\langle IP_i, FM_i \rangle$  pair, there will be a real-world property  $i$  such that M indicates  $i$  and M's Indication Profile from the perspective of indicating  $i$  is  $IP_i$ . The indeterminacy need not be as huge as the formalism suggests, although it will definitely be big enough for the basic-teleosemantic project to be in serious problems.

<b>Indication Profile</b>	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	...		
<b>Fitness Matrix</b>		...		$\begin{bmatrix} \textit{max} & \textit{min} \\ \textit{min} & \textit{max} \end{bmatrix}$
<b>Candidates for F in “There is an F around”</b>	Orange, round things	Peaches	Food	Good stuff
	<b>Lowest</b>	<b>“Natural”</b>	<b>Highest</b>	

Figure 1: A Continuum of Content Attributions

indeterminacy in content pointed out by critics of teleosemantics. In summary, input and output, indeterminacy and adequacy: all of these problems are simply different perspectives on the same underlying issue.

## 4 Millikan’s Biosemantics

In Millikan’s theory, to a good approximation, contents are to be identified with *(a subset of) the conditions that held in the occasions in which past responses to the output of the representations helped explain selection for the representation’s producer*. It will be easier to see the theory at work on our little model:

- A certain mechanism (M, in our example) produces a couple of representations— $M_{on}$  and  $M_{of}$ .
- $M_{on}$  is consumed in the following way: it causes its possessor to grab and eat the orange round thing nearby. This has been useful in a number of occasions because...

Now, at the right-hand side of the ‘because’ in the last sentence, biosemantics asks us to write the *most proximate Normal explanation* of the success of the grabbing-and-eating mechanism:

[The most proximate Normal explanation of a type of device R performing a function F] is the least detailed explanation

possible that starts by noting some features of the structure of members of R, adds some conditions in which R has historically been when it actually performed F—these conditions being uniform over as large a number of historical cases as possible—adds natural laws, and deduces, i.e., show in detail without gaps, how this setup leads to the performance of F. (Millikan 1984, p. 33)

A plausible candidate for the role of most proximate Normal explanation is that the grabbed-and-eaten thing is usually nutritious—that it is food.

- M's being *on* and its being *off*, in the relevant situations correspond to (or, in Millikanian terminology, *map onto*) the conditions of *There being food around M's possessor* and *There not being food around M's possessor*, respectively. These are to be considered the contents of the two states.

As Millikan writes discussing the well-known example of frogs snapping at flies,

[T]he systems that use, that respond to, the frog's fly detector signals, don't care at all whether these correspond to anything black or ambient or specklike, but only whether they correspond to frog food. . . . So the firing means frog food. (Millikan 1991, p. 163)

I have stated that Millikan's biosemantics is subject to the output problem. We can now see why: it is unclear that the most proximate Normal explanation in this case is, univocally, *food*. In fact, examination of the "conditions in which R has historically been when it actually performed F" would reveal that the food in question was *not sprinkled with poison*—it couldn't have, on pain of killing the ancestor of Democritus in that situation—so this is what a biosemanticist should settle in for as content. Or, maybe, it should be *non-poisonous food such that there is no predator near it*. Or. . . etc.

Millikan's account, in cases such as Democritus's, seems to find a resting place only in a very high content attribution; maybe something along the lines of *There is something good for digestive systems here*. Again, if there is just one such highest content attribution, Millikan's output problem is an adequacy problem; if there is not, it is an indeterminacy problem.

In a relatively recent summary of her views, there is a brief (but, as far as I am aware, the most explicit) discussion of the output problem:

Taking for her example the female-hoverfly detector in a male hoverfly's visual system, Karen Neander (1995) has objected that among the external conditions needed for the detector's consumers to perform all their functions are that the female is fertile and that she won't be eaten before she reproduces, hence that on the biosemantic theory these facts about the female must be part of what is represented by the detector in the male's visual system. What this overlooks, however, is that an intentional icon [for our purposes, this

is a contentful state—MM] must also have a producer and that it must be a function of the producer to make an icon that corresponds to the condition it represents. If the producer has a function there must be a normal mechanism by which it performs that function. This, however, would require the male hoverfly’s visual systems to be sensitive to natural signs of fertility in female hoverflies and of liability not to be eaten. But on no theory of information, certainly not on the theory of local natural information, does the male hoverfly use or even encounter any such natural information. (Millikan 2009)

This discussion presupposes the following setting: male hoverflies have a detector (we will call it HOV) that fires when a shadow of a certain shape and at a certain speed crosses the hoverfly’s retina. HOV’s going *on* causes the hoverfly to dart in a certain direction, calculated from the speed and angle of the shadow, which in a sufficient number of times helps the hoverfly reach a fertile female to mate with. The case is described in (Millikan 1990, 1993).

The passage just quoted is characteristic of Millikan’s more recent (post 2004) writings in that it portrays her theory, more clearly than before, as a variety of consumer semantics according to the characterisation I made in section 3.1: the theory warrants attribution of contents involving the highest property with the additional constraint that the attributee mental state carries (local natural) information about the instantiation of such a property.

But it is unclear that the appeal to information can lower a biosemantic content attribution away from the implausibly-high end of the content-attribution continuum. It is, to begin with, not true that “on no theory of information” does the male hoverfly encounter information about fertility in female hoverflies. Take *Indication*: in many cases we may perfectly well have that

**I1:**  $P(on|Fertile) > P(on)$ —where *Fertile* stands for the fact that the property of *Being a fertile female hoverfly* is instantiated near HOV’s possessor. And

**I2:** The difference in probabilities in I1 is not a matter of chance.

This may be so even if HOV is unable to distinguish at all between fertile and infertile female hoverflies, i.e., even if  $P(on|Fertile) = P(on|Infertile) = P(on|Female)$ . This is because

$$P(on) = P(Female) \cdot P(on|Female) + P(\neg Female) \cdot P(on|\neg Female)$$

and, so, in the relevant circumstances (i.e, when the rate of false alarms is sufficiently low and/or female hoverflies are sufficiently scarce), I1 will be met. In such a case, HOV indicates instantiations of the property of *Being a fertile female hoverfly* without needing to exploit a natural sign that is specific to fertile (as opposed to infertile) females.

What about Millikan’s own theory of local natural information? Although, in the passage quoted, she explicitly denies that HOV carries local natural information about fertile female hoverflies, it is unclear that this follows from the characterisation she has made (in her 2004, 2007) of this notion:

[A sign carrying natural local information] is one that corresponds to its represented in the same way, and for the same reason, that other signs of the same recurrent type correspond to theirs, and where there is a reason why examples of this kind of correspondence (with the same kind of cause) tend to spread from one location into nearby space-time locations. (Millikan 2007, p. 453)

If this is a definition, the matter is easy to settle: HOV *does* carry natural local information about fertile female hoverflies. There are causal explanations of the fact that a sign that loosely corresponds to female hoverflies also loosely corresponds to fertile female hoverflies, and a reason why this kind of correspondence tends to spread: the percentage of female hoverflies that are fertile stays approximately constant, and higher than the percentage of infertile female hoverflies, for reasons having to do (I hypothesise) with the rate of genetic mutation and environmental conditions leading to infertility, and which provide the causal grounding for the relevant probabilities. Also, the rate of fertile hoverflies gets copied with hoverfly reproduction, relevant environmental conditions stay put, etc.

Millikan has claimed that the reason why HOV carries local natural information about female-hoverflyhood but not fertile-female-hoverflyhood is that,

The domain in which the hoverfly operates is one in which the chance that the shadow crossing its retina, assuming that it is of a female hoverfly, is also of a fertile female not about to be eaten is no higher than the chance of any arbitrary female hoverfly being fertile and not about to be eaten. By contrast, assuming that it is the shadow of a hoverfly, the chance of the shadow being that of a female is considerably higher than the chance of an arbitrary hoverfly being female. (Millikan 2004, p. 85f)

This may be paraphrased as saying that

- While, on the one hand, the probability of there being a fertile female hoverfly around, conditional on [HOV’s being *on* and there being a female hoverfly around] is no higher than the probability of there being a fertile female hoverfly around conditional on there being a female hoverfly around—that is,

$$P(\textit{fertile female hoverfly} | \textit{HOV is on} \wedge \textit{female hoverfly}) \leq P(\textit{fertile female hoverfly} | \textit{female hoverfly})$$

- On the other hand, the probability of there being a female hoverfly conditional on [HOV's being *on* and there being a hoverfly around] is higher than the probability of there being a female hoverfly conditional on there being a hoverfly around. That is,

$$\frac{P(\text{female hoverfly} | \text{HOV is on} \wedge \text{hoverfly})}{P(\text{female hoverfly} | \text{hoverfly})} >$$

How this helps settle the issue about the information that HOV carries is unclear. For example, it is also true that

$$\frac{P(\text{fertile female hoverfly} | \text{HOV is on} \wedge \text{hoverfly})}{P(\text{fertile female hoverfly} | \text{hoverfly})} >$$

which, paraphrasing Millikan, means that

... By contrast, assuming that it is the shadow of a hoverfly, the chance of the shadow being that of a *fertile* female is considerably higher than the chance of an arbitrary hoverfly being a *fertile* female.

Again, this is because the rate of fertile females hoverflies in the female hoverfly population is sufficiently large, and a number of causal processes ensure that this remains so. But according to Millikan this is, if we are to judge by the passage just quoted and for reasons that are quite unclear, irrelevant to HOV's carrying local information about fertile-female-hoverflyhood.

Moreover, in the same passage, Millikan defends that HOV carries local natural information about hoverflyhood because

the chance of the shadow crossing [the hoverfly's] retina being that of a hoverfly rather than that of some other small particle of matter is also very much raised. (*ibid.*)

Which, I take it, can be rendered as

$$\frac{P(\text{hoverfly} | \text{HOV is on} \wedge \text{some particle})}{P(\text{hoverfly} | \text{some particle})} >$$

But, again, it is also true that

$$\frac{P(\text{fertile female hoverfly} | \text{HOV is on} \wedge \text{some particle})}{P(\text{fertile female hoverfly} | \text{some particle})} >$$

That is, paraphrasing Millikan again,

the chance of the shadow crossing [the hoverfly's] retina being that of a *female fertile* hoverfly rather than that of some other small particle of matter is also very much raised.

All in all, there is no principled reason to deny that HOV carries local natural information about instantiations of fertile-female-hoverflyhood. And, without such a reason, the “additional constraint” that natural local information puts to biosemantics attributions is unable to keep this property out of the content attribution. This is the output problem.

Millikan (in personal communication) has suggested that, while this discussion might show that she is forced to accept *There is a fertile female hoverfly around* as the right content attribution to HOV’s being *on*, this is as far as we need to get: further attempts to push biosemantics towards higher attributions will involve properties which HOV does not carry natural information about. Consider, for example, *Being a fertile female hoverfly such that it won’t be eaten before it reproduces*. Surely that’s not the kind of thing HOV can carry information about?

In fact, I think it is a virtue both of Millikan’s theory of local natural information and of *Indication* that, according to them, HOV does carry information about (indicates) this property. Again, the flying stuff that causes the right kind of shadows on the hoverfly’s retina is, more often than not and for a reason, a female hoverfly; female hoverflies are, more often than not and for a reason, fertile; fertile female hoverflies in turn are, more often than not and for a reason, lucky enough not to be eaten before reproducing—the reason: the density of predators is what it is, and the homeostatic mechanisms of the ecosystem will ensure that this remains so. Of course, if you nest enough [*more often than not, for a reason*] operators, you will reach a property that HOV does not indicate (carry information about). But the highest property that HOV indicates will already be too high to be an intuitive content-candidate. This is enough to substantiate the claim that Millikan’s biosemantics is subject to the output problem as a variety of the adequacy problem. If, moreover, as seems to be the case, there is no clear candidate for the highest property such that HOV carries information about it, the theory is also subject to the output problem as a variety of the indeterminacy problem.

I turn now to presenting my own positive proposal on how to make progress in the project of providing for content univocity.

## 5 An Alternative Strategy

The foregoing discussion shows that the appeal to information is no guarantee of solving adequacy and indeterminacy problems in consumer teleosemantics. The problem is not with hybrid (consumer + information) theories per se. I too believe that the right teleosemantic theory must be hybrid: the simplest contentful states do exploit (and exist because they exploit) a correlation between detectable (low) properties (*Being shiny and black*, say) and useful (high) properties (e.g., *Being nutritious*, or *Being dangerous*). The problem is in the way in which the theories on offer use these properties in the formulation of sufficient conditions for the attribution of content. Consumer-semantic accounts such as Millikan’s start from the high-properties end of a continuum of properties, such

as the one depicted in figure 1, and try to recover the low-properties end: contents involve, among the properties which consumers of the representation find useful, those that the producer of the representation has build a sensitivity to (carries information about). We have just seen that this strategy is subject to an adequacy (and maybe an indeterminacy) problem: the appeal to the needs of consumers push content attributions to the high, useful-properties end; the constraint that representations must be produced in normal conditions and, hence, carry local natural information helps lower content attributions somewhat, but not all the way down to the point that they solve the output problem: the highest property such that a representation carries information about it is still too high.

Trying to pin down the right property in the continuum— trying, that is, to screen off all but *one* property from the raw indeterminacy present in *Basic Teleosemantics*—turns out to be the wrong kind of strategy. The alternative I will describe here starts from *both* the high and low ends at the same time: the actual existence of mechanisms such as Democritus’s M or the male hoverfly’s HOV is explained by the presence, throughout the time needed for selection for them, of a mix of frequently coinstantiated high and low properties. The content attribution to  $M_{on}$  should be (not recovered after filtering out the rest of the mix, but) *constructed* out of this mix.

Simplifying some (but not much), selection for M in our model has been enabled by:

**Enabler1:** The frequent coinstantiation of low (detectable) properties—in our model, *Being an orange round thing*— and high (useful) ones— *Being nutritious*.

**Enabler2:** The causal mechanisms in place in the environment that have ensured that the properties in *Enabler1* remained frequently coinstantiated for the time needed for selection for M.

Notice that this much is not only true, but also common ground between friends and foes of teleosemantics: in order to show that a certain teleosemantic theory falls prey to the indeterminacy problem you need to show that it warrants multiple content attributions, and this involves an appeal to *Enabler1* and *2*. For example, to show that *Basic Teleosemantics* warrants both *There is a peach here* and *There is an orange round thing here* as content attributions one needs to show that both indicating peaches and indicating orange round things can explain the actual existence of M; this involves, in the former case, pointing out that peaches are both nutritious (thus useful) and orange and round (thus detectable) and, in the latter, showing that (the detectable) orange round things are, in Democritus’s environment, nutritious and thus useful.

A solution to the indeterminacy problem, I submit, would then be provided by a procedure that outputs the kind F that should be present in the content attribution *There is an F here* using as input the relatively uncontested material present in *Enabler1* and *Enabler2*.



Now, these two conditions are temptingly close to what an increasingly popular theory of natural kinds, the *homeostatic property cluster* (also HPC henceforth) view, proposes as constituting some of them:

The view says (roughly) that at least some natural kinds are individuated by property clusters that are afforded imperfect yet homeostatic integrity by underlying causal mechanisms. (Wilson, Barker & Brigandt 2007)<sup>10</sup>

The causal mechanisms that sustain the homeostatic integrity of properties such as *Being orange and round* and *Being nutritious* in Democritus’s environment is the one that explains that nutritious orange things spawn other nutritious orange things. The mechanism involves, or is identical with, peach reproduction. It is at least plausible that, from these materials, it should be possible to reconstruct the HPC that is the natural kind *peach*.

The contention, then, is that something similar to Boyd’s homeostatic property clusters solves the indeterminacy problem for naturalistic accounts of content. We were wrong to think that the task of the content naturaliser was to find out the right way to make a principled choice of one property among those in the detectable-to-useful continuum—see figure 1. There is no need to choose: in fact, the content of simple representations involves structures made out of indefinitely many such properties plus the causal explanation of their coinstantiation.

To summarise: there is an obvious relation between the “property clusters that are afforded imperfect yet homeostatic integrity” in the quote above and the frequent coinstantiation of properties appealed to in *Enabler1*; another between the “underlying causal mechanisms” and *Enabler2*. My proposal is to turn this obvious relation into an account of content for simple mental states, by constructing the HPC, *F*, that is to figure in the content attribution *There is an F here* out of the material in *Enabler1* and *Enabler2*—that is, to provide a principled way to carry out the reconstruction of the HPC *peach* from the enablers that I have quickly sketched above, and other similar reconstructions.

I turn to this task now: In section 5.1 I describe a principled way to provide analogues of *Enabler1* and *Enabler2* in general. In section 5.2, a procedure that outputs a content attribution out of *Enabler*-type facts. Section 5.3 recapitulates the discussion in a formally explicit fashion.

## 5.1 Obtaining *Enabler1* and *Enabler2*

Boyd (1988, 1999, p. 143f) proposes an eleven-clause elucidation of what HPCs are. Clauses 1 and 2 are the fundamental ones:

---

<sup>10</sup>The HPC view was originally developed by Richard Boyd in papers such as (Boyd 1988, 1991, 1999). Two recent defences are Brigandt (2009) and Wilson, Barker & Brigandt (2007). More relevant papers are referenced in Wilson, Barker & Brigandt (2007). Other related developments are Millikan (2000)’s *substances* (see also Millikan (2010)) and Ryder (2004)’s *unified property clusters*.

**Boyd1:** There is a family PC of properties which are 'contingently clustered' in nature in the sense that they co-occur in an important number of cases.

**Boyd2:** Their co-occurrence is not, at least typically, a statistical artifact, but rather the result of what may be metaphorically (sometimes literally) described as a sort of *homeostasis*. Either the presence of some of the properties in F tends (under appropriate conditions) to favor the presence of the others, or there are underlying mechanisms or processes which tend to maintain the presence of the properties in F, or both. (Boyd 1999, p. 143)

All in all, something very similar to our *Enablers*. Some of the other clauses, though, introduce a normative ingredient that is ill-suited for our current purposes.<sup>11</sup> We are in the business of providing a *naturalistic* recipe for the attribution of content to *M<sub>on</sub>*, and in this context we are not allowed to make undischarged appeals to intentional or normative states of affairs. We need, then, to provide non-normative counterparts to these features of Boyd's elucidation.

Let me discuss one such normative clause:

**Boyd3:** The homeostatic clustering of the properties in PC is causally important: (theoretically or practically) important effects are produced by a conjoint occurrence of (many of) the properties in PC together with (some or all of) the underlying mechanisms in question. (op. cit.)

For Boyd, this theoretical or practical importance is to be cashed out in terms of what he calls the "accommodation demands" of a certain "disciplinary matrix" (Boyd 1999, p. 148). The idea, roughly, is that the causal structures that secure the integrity of the HPC must be well suited to the kind of explanatory and inductive tasks in which they are put to use in some intellectual project or other. Thus, the HPC *acid* is well suited to the kinds of uses chemistry (which would be an example of a disciplinary matrix) puts it to. *Acid* is also well suited, for partially overlapping and partially independent reasons, to the uses mineralogy and nutrition sciences put it to.

This project-dependent nature of HPC is unacceptable for the content naturaliser's purposes. Even if it is unescapable when we try to reconstruct the use of natural-kind terms in human science, it only works against a fully intentional backdrop—it depends on the fully intentional goals of the practitioners of the discipline in question. Fortunately, as we have seen, teleosemantics offers a plausible alternative to *Boyd3*: the clustering must be causally important *for the existence of the indicator M*. That is, one way to read the kinds of conditions *Function* puts on a contentful mental state is as providing an alternative to

---

<sup>11</sup>It is unlikely that all of them are intended by Richard Boyd as necessary conditions for the existence of HPCs. In any event, I will assume that the metaphysics of HPCs is fixed by the first two clauses, and that the rest are more fruitfully taken as elucidations that spell out consequences of the first two, or identify particularly important examples of HPCs. My discussion of *Boyd3* below should be understood in this spirit.

Boyd’s accommodation demand. The epistemic project of a disciplinary matrix is here substituted by the, let us say, *biological project* which consists on the survival of an indicator. *Importance*, in this context, boils down to the purely causal notion of *selection for*. Thus,

**Boyd3\*:** The homeostatic clustering of the properties in PC is part of a causal explanation of the selection for the indicator M.

In the way I have been putting things earlier in this paper, PC includes the properties that explain that M has the collection of pairs of Indication Profile and Fitness Matrix it has, which in turn explain its Fitness Contribution as an indicator, according to *Function*—see sections 2 and 3 above. This offers a principled way to provide a candidate for the role of *Enabler1*:

The frequent coinstantiation of properties in PC.

*Enabler2* would, then, be simply identified with the set of causal mechanisms that keep the properties in PC frequently coinstantiated.

## 5.2 From *Enabler1* and *Enabler2* to a Content Attribution – The Underlying Theory of Kinds

One of the lessons of the first part of this paper is that there are indefinitely many initial candidates for the role of *Enabler1*: let us say that a *minimal property cluster* is a pair of a high and a low property—well, most every minimal property cluster will make a decent *Enabler1*. Will each of them give rise to a different HPC, and thus the theory relapse into the indeterminacy we were trying to avoid?

No, and we don’t need to complicate our theory of intentionality proper in order to avoid this, but simply pay attention to certain facts that the theory of its *targets*, the HPCs themselves, needs to accommodate:

It is often claimed by HPC theorists that natural kinds are natural insofar as they support induction (see, for example, Boyd 1999; Millikan 2000, p. 17f). The accommodation of inductive practices is, of course, central to the idea of natural kinds; and minimal property clusters meet this demand: in fact, we could describe the very simple contentful states I have been discussing here as enabling a proto-intentional precursor of induction—inferring the presence of high properties from the presence of low ones.

But being induction-grounders is not the only important property of natural kinds, even if it’s the one that makes us develop an interest on them. Natural kinds, for example, are also what Hacking calls *Mill-Kinds*:

A Mill-Kind is a class of objects with a large or even apparently inexhaustible number of properties in common, and such that these properties are not implied by any known systematized body of law about things of this Kind. (Hacking 1991, p. 120)

Minimal property clusters are able to ground induction, but they are not Mill-Kinds in Hacking’s sense: they have *two* properties, and two is far from “inexhaustible”. What we need (read: what the theory of HPCs needs) is a way to turn minimal HPCs into Mill-Kind HPCs. The following, then, is a natural suggestion:

1. We start with what we could call the *seed* of the cluster—a set of properties such as one of the minimal property clusters, which we may have found by using *Boyd3\**.
2. We then identify the homeostatic mechanism *HM* that keeps properties in the seed frequently coinstantiated.
3. The final property cluster is the one that includes every property *P* such that *HM* keeps *P* frequently coinstantiated with properties in the seed.

See Magnus (2011) for a similar idea. There are many candidates for the role of seed: one for each minimal property cluster, and many more beyond those. But this three step process will typically make all of the seeds converge to the same final property cluster<sup>12</sup>. Let me stress again that this convergence from several minimal HPCs onto a final property cluster is something that the theory of kinds, not the theory of intentionality, needs to deliver, regardless of the content naturalisers’ agenda. Otherwise put, the following is a intentionality-independent fact: the minimal entities that are able to ground induction need not be Mill-Kinds; an account of kinds that focuses solely on meeting the desideratum that kinds ground induction will be, thus, incomplete.

The theory of HPCs also needs a way to identify homeostatic mechanisms of the right scope: the theory plays (as well it should) fast and loose with the idea of *the* homeostatic mechanism that makes properties in the cluster co-occur. This is generally fine, because the theory’s intended target is the working-scientist’s practices, and she will be usually able to single out one homeostatic mechanism as the relevant one. Content naturalisers on the other hand, need to tread carefully here: if the theory of kinds depends ineliminably on the scientists’s knack for choosing homeostatic mechanisms, this intentional ingredient, down the road, will taint the theory of content that relies, even if derivatively, on such a knack.

And, without a principled way of choosing mechanisms, we need the knack, because not every homeostatic mechanism will do. Suppose, for example, that the Gaia hypothesis is right—it is not, but let that pass—and the Earth is one immense homeostatic system. If so, there is a homeostatic mechanism that keep sundry collections of properties in the global ecosystem frequently coinstantiated—one such collection of properties is the property cluster of the HPC *peach*. If *HM* is the Gaia mechanism, the three-step process outlined above would take the properties in *peach* and adjoin to them every property that the

---

<sup>12</sup>The indeterminacy that this process removes is that between seed-candidates *that meet* *Boyd3\**. There might be other candidates that the three step process does not rule out—nor should it, as *Boyd3\** takes care of them.

Gaia-homeostatic mechanism keeps frequently coinstantiated with them; say, the presence of a core of iron and nickel well below the peaches, etc. This would disfigure the HPC beyond recognition.

I have said that devising a procedure to pick out a homeostatic mechanism of the right scope is not a priority for the HPC theory, the way it is a priority for the teleosemantics that relies on it, but that doesn't mean that the former can simply make do without such a procedure. At the very least, without it the theory fails to deliver concrete predictions about the extension of kinds without relying on the practising-scientist's knack for choosing homeostatic mechanisms. I offer *Specialised Homeostatic Mechanism* as a promising candidate procedure.

**Specialised Homeostatic Mechanism:** Consider a certain set of properties, *Seed*. For every homeostatic mechanism, HM, such that HM keeps properties in *Seed* frequently coinstantiated, apply the three step process above, to reach a final property cluster  $FPC_{HM}$ . The specialised homeostatic mechanism is the one,  $sHM$ , such that

$$\forall HM [FPC_{sHM} \subseteq FPC_{HM}]$$

That is, the specialised homeostatic mechanism is the one that keeps frequently coinstantiated with *Seed* the least number of properties, in the sense that it keeps together a set of properties that is contained in the set kept together by any other homeostatic mechanism.<sup>1314</sup>

In any event, what teleosemantics needs is a sane foundation in the theory of kinds. Much of the ingenuity that has been deployed, by Millikan and others, in making the theory of contentful states watertight—the notion of *most proximal normal explanation*, discussed above. comes to mind—should instead be used in improving the theory of its *targets*, the real kinds, which anyway needs to solve what look like suspiciously similar issues—because they are, in fact, the same issues.

---

<sup>13</sup>Given that these sets will have infinitely many members, this is probably the only available sense of “least number of properties”. Typically, there will be such a proper subset of properties. But we should keep an eye open to the possibility that, in some cases, that subset fails to exist. Such cases would call for refinement in *Specialised Homeostatic Mechanism*.

<sup>14</sup>It would be very short-sighted of me to suggest that the foregoing discussion leaves no gaps in the resulting HPC theory, even insofar as the teleosemanticist is concerned. At least the following issue requires further work: most HPCs are partially individuated by a homeostatic mechanism that, so to say, sits on top of other homeostatic mechanisms. For example, the hoverfly homeostatic mechanism (which, we would think, is more or less identical to *hoverfly sexual reproduction*) is only operative in the presence of other homeostatic mechanisms that provide for a fertile soil, oxygen-rich atmosphere, Earth-like gravity and the like. The theory of HPCs needs to provide a way of teasing apart the contributions of the hoverfly mechanism from those other, merely enabling, mechanisms. I am confident that the teasing apart can be done, for example, by appealing to informational asymmetries in the property cluster, such as: the presence of hoverfly properties makes the presence of an oxygen-rich atmosphere more likely than the presence of the latter makes the presence of the former. Turning this idea into a full-fledged proposal, though, requires more elaboration than can be undertaken here.

### 5.3 A Content Attribution Recipe

For the sake of explicitness, I now recapitulate the process outlined in the previous subsections in a more formal manner.

First, a procedure that takes a seed of properties (maybe one of the pairs of low and high properties we saw above) and a class of their instantiations—these will be the concrete instantiations that have played a causal role in selection for the indicator  $M$ —and outputs an HPC, by using the three-step process from seed to final cluster and *Specialised Homeostatic Mechanism*:

**Target HPC:** Consider a class of properties, *Seed*, and a class of their instantiations,  $Seed_{inst}$ . If there is a set of homeostatic mechanisms  $S_{HM}$  that explain the existence of a sufficient number of the instantiations in  $Seed_{inst}$ , then we identify the *final property cluster associated with*  $HM \in S_{HM}$  with the class of properties,  $FPC_{HM}$ , such that HM explains the fact that instantiations of the properties in  $FPC_{HM}$  occur sufficiently close<sup>15</sup> to the instantiations in  $Seed_{inst}$ .

The *target HPC* associated to *Seed* and  $Seed_{inst}$  is the one individuated by:

- The specialised homeostatic mechanism sHM, according to *Specialised Homeostatic Mechanism*, and
- Its final property cluster  $FPC_{sHM}$ .

Second, the content-attributing recipe: the way to tie *Target HPC* to a particular indicator  $M$  is by letting the properties that explain  $M$ 's coming to have the function to indicate fix *Seed* and  $Seed_{inst}$ :

**There is an F here:**  $M_{on}$  has the content *There is an F here* if<sup>16</sup>

**TAFH1:** *Function* warrants, for some properties  $G_i$ , the attribution to  $M$  of the function of indicating the instantiation of  $G_i$  around  $M$ 's possessor,  $S$ .

**TAFH2:** These function attributions are grounded on the frequent co-instantiation of a mix of high and low properties in  $S$ 's environment.

**TAFH3:** *Seed* is the class of the properties in the mix appealed to in TAFH2.  $Seed_{inst}$  is the class of actual instantiations of properties in *Seed* the frequent coincidence of which has been causally relevant in the grounding of the function attributions.

<sup>15</sup>*Informational* (and not necessarily spatio-temporal) closeness: there is a function (mathematical sense) from positions of instantiations in  $Seed_{inst}$  to positions of instantiations of properties in  $FPC_{HM}$  such that the probability of having an instantiation of the latter at the image of the position of instantiations of properties in  $Seed_{inst}$ , according to this function, is suitably high.

<sup>16</sup>Note that this set of conditions are not presented as necessary: for example, a human belief that there is an F here will be fixed in a much more convoluted manner.

**TAFH4:** F is the HPC that results from applying *Target HPC* to *Seed* and *Seed<sub>inst</sub>*.

*There is an F here*, I submit, is an account of the content of simple mental states that recovers all the advantages of traditional teleosemantic proposals, without falling prey to the indeterminacy problem. To put the main idea under a different light, the proposal is that the content of these simple, innate mental states should not involve (as, e.g., Millikanian teleosemantics would have it) useful properties such that there is an explanation why they can be detected but, rather, it should involve *the very structure* individuated by those properties and this explanation. This structure (the homeostatic property cluster) is not any random mereological sum of causal facts and properties, but may be plausibly regarded—and is customarily regarded—as a real kind. The final content-attributing recipe is, no doubt, somewhat cumbersome, but no less is sufficient to avoid indeterminacy. And, as Godfrey-Smith wrote (more than twenty years ago),

we should have stopped expecting to hear four-line theories of meaning by now. (Godfrey-Smith 1991)

## 6 Applying the Recipe

A quick recap: one may identify indefinitely many properties such that, on the one hand, M indicates them and, on the other, this figures in an explanation of M's Fitness Contribution. For example, for Democritus: *things superficially similar to peaches; orange round things*, etc. Some of them, M indicates very accurately (i.e., they have an associated Indication Profile that is nearly the identity matrix). Some of them less so, but, according to *Basic Teleosemantics*, content attributions involving each of them are, all, warranted. Another bunch of candidates for warranted content-attributions crop up if we focus on the properties that ground the fitness gains—that is, properties that maximise values in the diagonal of the Fitness Matrix: *nutritious stuff, food, unpoisoned food*, etc. Again, each of these properties ground a different pair of Indication Profile and Fitness Matrix, but all of them amount to the same Fitness Contribution.

Now, the way to solve the Indeterminacy Problem: we look for the properties that explain that *Function* is met for these alternative content attributions. This class of properties will be the seed of the target HPC. If Democritus's M is to be contentful, there will be a specialised homeostatic mechanism, sHM, that has made properties in *Seed* recur around Democritus and its ancestors. This mechanism is, likely, peach reproduction. According to *There is an F here*, the HPC that must figure in the content attribution is the one individuated by sHM, together with the set  $FPC_{sHM}$  of all properties such that sHM explains their recurrence with the relevant instantiations of properties in *Seed*:

$$FPC_{sHM} = \{ \textit{things superficially similar to peaches, orange round things, food, unpoisoned food, thing with such-and-such a DNA. . .} \}$$

We happen to have a name in English for such an HPC: *peach*. The content of  $M_{on}$  is *There is a peach here*.

As we have seen, the fact that  $M$  has indefinitely many functions to indicate—the root of the indeterminacy problem—is not a difficulty for *There is an  $F$  here*: the HPC that must figure in the content attribution is the structure that enables the emergence of all of those functions. On the other hand, the account accommodates misrepresentation in a manner entirely analogous to that of mainstream teleosemantics<sup>17</sup>: *There is an  $F$  here* warrants an attribution of the content *There is a peach here* to Democritus'  $M_{on}$ . But there is no guarantee that  $M_{on}$  will be tokened only in the presence of a peach: for example, in our model,  $M_{on}$  will be tokened iff the property *Being orange and round* is instantiated around Democritus. So, if an ethologist presents Democritus with a cardboard orange, a token of  $M_{on}$  will be produced; this token misrepresents the situation as one in which there is a peach there, when there is none.

## 7 How is this Content?

Most naturalistic accounts of content propose an *analysans* which have an intuitive claim to being a candidate for the content of representations. Thus, simple causal accounts: the things that cause representations have a *prima facie* appeal as possible candidates for the role of contents, even if, on *secunda facie* reflection, the proposal is full with daunting problems<sup>18</sup>. Similarly, *Basic Teleosemantics* has intuitive appeal too: contents do seem the kinds of things that representations may have the function to indicate<sup>19</sup>. On the other hand, learning to live with the acute indeterminacy problem that this account suffers is not an attractive option.

*There is an  $F$  here* faces none of these imposing difficulties. In particular I have, I hope, provided sufficient reasons to think that it is not subject to the indeterminacy problem. The worry I wish to consider now is whether this resilience to problems against naturalistic accounts of content is achieved at the cost of not being an account of *content*, at all. Indeed, the involved sufficient condition for being the content of a representation which *There is an  $F$  here*

---

<sup>17</sup>In fact, the account can co-opt most teleosemantic insights. For example, the content of non-selected-for mental states can be fixed in, roughly, the manner suggested in Millikan (1984): we postulate the existence of a mechanism that has been selected for producing further mechanisms, whose states are the ones to which content will be attributed. These latter mechanisms and their states have not been selected; their existence is way too short for that. The states are, nevertheless, contentful: their existence enjoys a content-conferring biological explanation, through the workings of their producer's producer—which has been selected for in the usual way.

<sup>18</sup>A particularly vexing one: they don't allow for the possibility of misrepresentation. See (Fodor 1990).

<sup>19</sup>Question: Why not claim that the HPCs identified by *There is an  $F$  here* are the only entities that indicators have the function to indicate, and thus enjoy the intuitive appeal of *Basic Teleosemantics*? Answer: Because I am sceptical that a theory of biological functions can be formulated which, first, yields this as a univocal result for indicators while, second, still providing the correct predictions for, say, hearts, beaks, and wings.



recommends does not seem to strike an intuitive chord the way, e.g., the appeal to functions to indicate does. I will now argue that, on *secunda facie* reflection, it does.

The first thing to note is that *There is an F here*, just like *Basic Teleosemantics* and simple causal accounts, abides by the following general principle:

**Compressed Explanation:** To provide a content attribution for a representation type R is to provide a compressed explanation of the existence of R in a sufficient number of cases.

It may be defended that the principle *Compressed Explanation* provides a partial elucidation of the essence of content<sup>20</sup>; perhaps, after sufficient reflection it is *a priori* true that content attributions are, in fact, in the business of explaining the existence of representations. Be that as it may, at any rate all of the theories we have considered here comply with it. Take simple causal accounts. Indeed, whatever causes a certain representation to token figures in an explanation—the most proximal—of the existence of the representation. Or *Basic Teleosemantics*: at least under the etiological understanding of what functions are, to say that something is whatever a representation has the function to indicate (i.e., its content) is to say that it figures in an important way in an explanation of the existence of the representation: had the ancestors of the representation not indicated it, the actual representation would not exist<sup>21</sup>.

*There is an F here* also abides by the principle. In fact, according to the proposal presented in this paper, one can recover a substantial part of the story that explains the existence of a (simple, innate) representation just from its content attribution: the HPC that figures in a content attribution explains how the emergence of functions-to-indicate, that in turn explain the actual existence of the representation, was possible in the case at hand.

A second thing to note is that what explains the appeal of simple causal accounts is also present here: there is a real kind (an HPC) lying about in the environment, and it is this real kind that brings the representation into existence. At this level of description both theories are on a par. It is only that, we have found out, not any old way of causing the representation is sufficient to fix a content; rather, we have been led to conclude that only a very concrete kind of intervention of the HPC in the events leading to the existence of the representation will secure the former a place in the content of the latter. The theory I am proposing is the wisened-up version of a simple causal account.

A final consideration in favour of *There is an F here* comes from the kinds of attributions the theory makes. The account uses as content attributions

---

<sup>20</sup>Just *partial*, in any case: e.g., helping explain behaviour is likely also part of what contents essentially do.

<sup>21</sup>It is maybe evidence in favour of *Compressed Explanation* that most contemporary accounts that rely on broadly teleosemantic insights also uphold etiological functions. A prominent counterexample to this is (Cummins 1991), but Cummins himself points out that his account is not after the content of thoughts but rather after “the content of the representations of a computational system” (id., p. 88).

Even if Cummins’s theory were correct, *Compressed Explanation* might still be true for the content of thoughts.

the entities that keep together a large number of properties that are of use to the possessor of the representation— immediately useful such as *Being food*, immediately detectable such as *Being an orange round thing*, etc. It is likely that the entity distinguished with this role will not lie at any of the two extremes that give rise to the Input and Output Problems. In Democritus’s case, for instance, the entity is *peach*; one of the attributions we would not have pretheoretically ruled out.

All in all, *There is an F here* provides a candidate for the role of content that is, upon reflection, as good from the intuitive perspective as other prominent causal theories of content. On the other hand, *There is an F here* does not suffer from the problems of adequacy, or indeterminacy. I tentatively conclude that it is, thus, to be preferred to the other accounts I have discussed.

## Acknowledgements

Financial support for this work was provided by the DGI, Spanish Government, research project FFI2010-15717, and Consolider-Ingenio project CSD2009-00056.

I would like to thank Marc Artiga, Ruth Millikan, David Pineda, Dan Ryder, Miguel Ángel Sebastián, audiences in Barcelona, Lausanne and Amsterdam, and five anonymous reviewers for their comments and suggestions on different versions of this paper.

## References

- Agar, N 1993, ‘What Do Frogs Really Believe?’, *Australasian Journal of Philosophy*, vol. 71, no. 1, pp. 1–12.
- Boyd, R 1999, ‘Homeostasis, Species, and Higher Taxa’, in RA Wilson (ed.), *Species: new interdisciplinary essays*, Mit Press, pp. 141–185.
- Boyd, R 1988, ‘How to be a moral realist’, in G Sayre-McCord (ed.), *Moral realism*, Cornell University Press, pp. 181–228.
- Boyd, R 1991, ‘Realism, anti-foundationalism and the enthusiasm for natural kinds’, *Philosophical Studies*, vol. 61, no. 1-2, pp. 127–148.
- Brigandt, I 2009, ‘Natural kinds in evolution and systematics: metaphysical and epistemological considerations’, *Acta Biotheoretica*, vol. 57, pp. 77–97.
- Cummins, R 1991, *Meaning and Mental Representation*, The MIT Press.
- Dretske, F 1988, *Explaining behavior. reasons in a world of causes*, The MIT Press.
- Fodor, J 1990, *A theory of content and other essays*, The MIT Press.
- Godfrey-Smith, P 1994, ‘A modern history theory of functions’, *Noûs*, vol. 28, no. 3, pp. 344–362.
- Godfrey-Smith, P 1996, *Complexity and the function of mind in nature*, Cambridge University Press.
- Godfrey-Smith, P 2009, *Darwinian Populations and Natural Selection*, Oxford University Press.

- Godfrey-Smith, P 1991, 'Signal, Detection, Action', *Journal of Philosophy*, vol. 88, no. 12, pp. 709–722.
- Hacking, I 1991, 'A Tradition of Natural Kinds', *Philosophical Studies*, vol. 61, no. 1/2, Springer, pp. 109–126.
- Lewis, D 1986, *On the Plurality of Worlds*, Blackwell, Oxford.
- Magnus, PD 2011, 'Drakes, Seadevils, and Similarity Fetishism', *Biology and Philosophy*, vol. 26, no. 6, pp. 857–870.
- Millikan, R 2007, 'An Input Condition for Teleosemantics? A reply to Shea (and Godfrey-Smith)', *Philosophy and Phenomenological Research*, vol. 75, no. 2.
- Millikan, R 2002, 'Biofunctions: Two Paradigms', in A Ariew, R Cummins & M Perlman (eds), *Functions: new essays in the philosophy of psychology and biology*, Oxford University Press, pp. 113–143.
- Millikan, R 2009, 'Biosemantics', in B McLaughlin & A Beckermann (eds), *The Oxford Handbook of Philosophy of Mind*, Oxford University Press.
- Millikan, R 1984, *Language, Thought and Other Biological Categories*, The MIT Press.
- Millikan, R 2000, *On Clear and Confused Ideas*, Cambridge University Press.
- Millikan, R 1991, 'Speaking Up for Darwin', in B Loewer & G Rey (eds), *Meaning in Mind. Fodor and his Critics*, Blackwell, pp. 151–164.
- Millikan, R 2004, *Varieties of Meaning*, The MIT Press.
- Millikan, R 1993, *White Queen Psychology and Other Essays for Alice*, The MIT Press. Bradford Books.
- Millikan, RG 2010, 'On Knowing the Meaning; With a Coda on Swampman', *Mind*, vol. 119, pp. 43–81.
- Millikan, RG 1990, 'Truth, Rules, Hoverflies, and the Kripke-Wittgenstein Paradox', *Philosophical Review*, vol. 99, no. 3, pp. 323–353.
- Neander, K 2006, 'Content for Cognitive Science', in G Macdonald & D Papineau (eds), *Teleosemantics*, Oxford University Press, pp. 167–194.
- Neander, K 1995, 'Misrepresenting & Malfunctioning', *Philosophical Studies*, vol. 79, pp. 109–141.
- Neander, K 2013, 'Toward an Informational Teleosemantics', in D Ryder, J Kingsbury & K Williford (eds), *Millikan and her critics*, Wiley-Blackwell, pp. 21–40.
- Papineau, D 1998, 'Teleosemantics and Indeterminacy', *Australasian Journal of Philosophy*, vol. 76, no. 1, pp. 1–14.
- Price, C 1998, 'Determinate Functions', *Noûs*, vol. 32, no. 1, pp. 54–75.
- Rowlands, M 1997, 'Teleological Semantics', *Mind*, vol. 106, no. 422, pp. 279–303.
- Ryder, D 2004, 'SINBAD Neurosemantics: A Theory of Mental Representation', *Mind & Language*, vol. 19, no. 2, pp. 211–240.
- Sober, E 2000, *Philosophy of Biology (2nd Edition)*, Westview Press.
- Wilson, R, Barker, M & Brigandt, I 2007, 'When Traditional Essentialism Fails: Biological Natural Kinds', *Philosophical Topics*, vol. 35, no. 1-2, pp. 189–215.

Wright, L 1994, 'Functions', in E Sober (ed.), *Conceptual Issues in Evolutionary Biology*, The MIT Press. Bradford Books, pp. 27–48.

Wright, L 1976, *Teleological Explanations: An Etiological Analysis of Goals and Functions*, University of California Press.