## The Elusive Role of Normal-Proper Function in Cognitive Science

Frances Egan

## 1. Introduction

The main goal of Karen Neander's important book (2017) is to defend a particular version of *teleosemantics*, which holds that the norms necessary for the semantic evaluation of intentional mental states in cognitive science are grounded (at least in part) in the *normal-proper functions* of cognitive systems. The account aims to be *naturalistic*, explaining how intentionality is derived from non-intentional and non-semantic properties already recognized by the natural sciences. Its scope is restricted to low-level *non-conceptual* cognition, presuming – as is standard practice among naturalistic philosophers of cognition – that insight derived from studying simple cases will eventually help illuminate sophisticated human thought. The project is admirable in eschewing appeal to intuitions about toy cases, instead building a case on the basis of careful discussion of empirical work in cognitive science. My comments will focus on two central chapters – a general methodological argument for teleosemantics and a detailed discussion of prey-capture in the toad. I argue that while Neander's work elucidating the toad case, in particular, is exemplary, she hasn't succeeded in discovering what teleosemantics needs, namely, an explanatory role for the notion of normal-proper function in cognitive science.

## 2. The Methodological Argument for Informational Teleosemantics

In chapter 4 Neander offers a methodological argument for informational teleosemantics. It is intended to show that the view is supported by the explanations of cognition currently provided by cognitive science. Here is the argument:

P1 A notion of normal-proper function is central to the multilevel componential analyses (i.e. functional analyses) of the operation of bodies and brains that are currently provided by physiologists and neurophysiologists.

P2 The brain's normal-proper functions include cognitive functions.

P3 The same notion of function (mentioned in P1) is central to the functional analyses of cognition that cognitive scientists provide.

P4 An assumption in the mainstream branches of cognitive science is that cognition involves information processing.

P5 The (relevant) notion of information involved in such talk of information processing in cognitive science is (not a semantically evaluable notion but instead) a notion of natural-factive information.

P6 Cognitive science posits "normative aboutness," with the norms derived from the aforementioned normal-proper functions and the aboutness from the aforementioned natural-factive information.

C Informational teleosemantics is supported by the explanations of cognition that the mind and brain sciences currently provide.

I shall challenge the argument at P3 and P6.

The notion of *normal-proper function* appealed to in P1 (and P2, P3, and P6), Neander argues (in chapter 3), is central to biology, and more specifically, to the analyses of bodily and brain processes provided by physiologists and neurophysiologists. The notion of N-P function has two important features: (i) it is the notion of function that underwrites talk of *normal* and *abnormal* functioning, as in, for example, talk of a normally functioning human immune system, and (ii) it supports a distinction between *function* and *accident*, which underwrites the idea that

an item is *for* something. Hearts, for example, are for pumping blood, not for making whooshing sounds, though they typically do both. *Selected* function – what an item was selected to do – is one species of normal-proper function, and it is this more specific notion that Neander in fact thinks plays a central role in both physiology and the cognitive sciences.

I think that feature (ii) – privileging *one* effect among the various things that an item typically does – presupposes *some* process of selection (typically by evolution in the case of biological systems, or by design in the case of artifacts). In any event, the methodological argument for teleosemantics appeals only to the more general notion of normal-proper function. I will challenge the claim (P3) that this notion plays a central role in cognitive science.

The branches of cognitive science at issue in the argument (see P4) are those that construe cognition as a species of *information processing*, in other words, computational cognitive psychology and computational neuroscience. The issue, then, is whether the notion of N-P function plays an explanatory role in these sciences. If we look at actual practice I think it is clear that it does not.[1] Computational theorists aim to characterize the processes underlying cognition as a species of mathematical process. Observing that a physical system succeeds at some cognitive task – detecting significant boundaries in the scene, grasping nearby objects, understanding and producing speech, and so on – the theorist hypothesizes that the system computes a well-defined function (in the *mathematical* sense) and explains how computing this function constitutes (in the system's normal environment) the exercise of the cognitive capacity that is the explanatory target of the theory.[2] The strategy is pervasive in computational cognitive science; let me cite a few examples from various cognitive domains. Marr's (1982) theory of

---

[1] See Egan 2014, 2017 for defense of the view described here.
[2] This mathematical sense of function is not to be confused with the various notions of function at issue in the present debate.

early vision purports to explain edge detection by positing the computation of the Laplacian of a Gaussian of the retinal array. The mechanism takes as input intensity values at points in the image and calculates the rate of intensity change over the image. In other words, it computes a particular smoothing function. Ullman (1979) hypothesizes that the visual system recovers the 3D structure of moving objects by computing a function from three distinct views of four non-coplanar points to the unique rigid configuration consistent with the points. Shadmehr and Wise's (2005) computational account of motor control purports to explain how a subject is able to grasp an object in view by computing the displacement of the hand from its current location to the target location, i.e. by computing vector subtraction. In a well-known example from animal cognition, Gallistel (1990) purports to explain the Tunisian desert ant's impressive navigational abilities by appeal to the computation of the displacement vector to its nest from any point along its foraging trajectory. Seung et al. (1996, 1998, 2000) hypothesize that the brain keeps track of eye movements by deploying an internal integrator. Theories employing the strategy explain a cognitive capacity by appeal to an independently well-understood mathematical function under which the physical system is subsumed. In other words, what gets computed, according to these computational models, is the value of a mathematical function (e.g., vector subtraction, the Laplacian of a Gaussian, a fast Fourier transform) for certain arguments for which the function is defined. A fully specified theory of a cognitive capacity will go on to propose an algorithm by which the computation of the value of the function(s) is effected, and describe the neural hardware that implements the computation.[3]

---

[3] The specification of the function computed, the specification of the algorithm, and the neural implementation correspond, roughly, to David Marr's three levels of description—the computational, algorithmic, and implementation, respectively. The topmost, computational, level of theory also adverts to general environmental facts ('constraints') essential to the explanation of the cognitive capacity, as explained below. See Egan (1995) for elaboration and defense of

The mathematical characterization provides a domain-general, environment-neutral characterization of a mechanism. It prescinds not only from the cognitive capacity that is the explanatory target of the theory (vision, motor control, etc.), but also from the environment in which the capacity is normally exercised. In fact, the abstract nature of the computational characterization—in particular, the fact that as an independently characterized mathematical object the function can be decoupled from both the environmental context and the cognitive domain that it subserves—accounts for an important explanatory virtue of computational characterization. The mathematical functions deployed in computational models are typically well understood independently of their use in such models. Laplacian of Gaussian filters, fast Fourier transforms, vector subtraction, and so on are standard items in the applied mathematician's toolbox. To apply one of these tools to a biological system—to subsume the system under the mathematical description—makes sense of what might otherwise be a heterogeneous collection of input–output pairs. Since the computational characterization specifies the function intensionally, typically in terms of an algorithm for computing the function, it provides the basis for predicting the output of the device in a wide range of circumstances that go well beyond the observed data set.

Of course, the cognitive theorist must explain how computing the value of the specified function, in the subject's normal environment, contributes to the exercise of the cognitive capacity that is the explanatory target of the theory. Only in some environments would computing the Laplacian of a Gaussian help an organism to see. In our environment this computation produces a smoothed output that facilitates the detection of sharp intensity gradients across the retina, which, when these intensity gradients co-occur at different scales, correspond

---

this account of Marr's computational level.

to physically significant boundaries—changes in depth, surface orientation, illumination, or reflectance—in the scene. Ullman's structure-from-motion mechanism succeeds in recovering the 3D structure of a moving object by computing the unique rigid configuration consistent with three distinct views of four non-coplanar points on the object only because, in our world, most objects are rigid in translation (the *rigidity assumption*). Thus, to yield an explanation of the target cognitive capacity, the environment-neutral, domain-general characterization given by the computational description must be supplemented by environment-specific facts that explain how computing the value of the specified mathematical function, in the subject's normal environment, contributes to the exercise of the target cognitive capacity.

Neander wants to find an explanatory role for normal-proper function in computational theorizing, but it is not at all clear that there is such. It is, however, possible to *recover* a notion of normal-proper function from a computational account of a cognitive capacity. On the plausible assumption that perceptual mechanisms are *adaptations*, detecting edges in the scene is likely to be what the Laplacian/Gaussian mechanism was selected *for* and so is likely to be the normal-proper function of the mechanism. But adaptationist assumptions play no role in computational theorizing itself, which takes no account of historical origins. In any event, it does not depend on or presuppose the notion on normal-proper function, as P3 requires.

Of course, ultimately the results of computational neuroscience and neurophysiology will have to be integrated – both theorists are engaged in explaining how the brain works – but it does not follow that notions of function essential to physiology (and biology more generally) play a central role in computational neuroscience. Moreover, it does not suffice for P3 that a computational model presupposes some notion of *normal functioning*. Computational theorists aim to explain manifest cognitive capacities, and, as noted above, they do so by positing

mathematical capacities. In attributing a mathematical capacity or *competence* to a physical system – to compute vector displacement, or a smoothing function – a computational model supports attributions of *correctness* and *mistake*. Just as the normal functioning of the system—correctly computing the specified mathematical function—explains the subject's success at a cognitive task in its normal environment, so a malfunction explains its occasional failure.

The notion of normal functioning central to computational theorizing arises in a distinctive way. A complete computational characterization of a physical system will specify a mapping from physical states of the system to the arguments and values of the computed (mathematical) function. The mapping *interprets* physical state transitions as computations. For ease of explication, let us suppose that the system under consideration is an adder. The mapping specifies that when the system goes into the physical state interpreted under the mapping as 57 and goes into the physical state interpreted under the mapping as 43, then it will go into the physical state interpreted under the mapping as 100. In doing so it (correctly) computes the sum of 57 and 43. If instead it were to go into the physical state interpreted under the mapping as 99 it would *miscompute* or *make a mistake*. In some circumstances – if, say, the system suffers severe damage – it may go into a physical state not interpreted under the mapping at all, and we can say in this case that it doesn't *miscompute*, it simply *malfunctions*. (A miscomputation may be seen as a special kind of malfunction. It is also, as all mistakes are, a partial exercise of a competence.) So the specification of the function computed (here, addition) and the interpretation (here, the mapping of physical states to addends and sums) provides all we need to partition the behavioral space into (i) correct computations, (ii) miscalculations or mistakes, and (iii) malfunctions (states that receive no interpretation under the mapping). The specification of the function computed and the interpretation together provide the *norm* necessary to underwrite

this three-fold distinction. Since a computational model is committed to the existence of a mapping of this sort, the model supports talk of 'well-functioning' even if (as is usually the case with natural computers) the mapping is not fully specified. Importantly, the norm is not grounded in the selection history of the organism or other historical facts.

In summary, the computational characterization does not make essential appeal to the notion of *normal-proper function*, as required by P3. As noted above, the notion can be *recovered* from a well-confirmed computational model of a cognitive capacity. A computational characterization will allow us to say that a system is *for* adding, or *for* detecting edges or *for* grasping objects in view, but the notion doesn't play an essential role in the theorizing.

To return to the methodological argument, Premise 5 says that the relevant notion of "information" in information-processing accounts is akin to Grice's "natural meaning"; it is factive. If smoke means fire (in this sense) then if there is smoke there must be fire. Natural-factive information is not semantically evaluable; it is not *content*.

Premise 6 says that cognitive science (in particular, information-processing or computational accounts) posits "normative aboutness", which *is* a semantically evaluable notion. Natural-factive information gets its normative *boost*, Neander claims, from the normal-proper functions of the items that carry it in cognitive systems. But I have argued that the notion of normal-proper function plays no role in the branches of cognitive science that treat cognition as information-processing. The norms that ground the distinction between well-functioning and malfunctioning in computational cognitive science are not the norms central to biology and physiology; rather they are grounded in the interpretation of the physical system as computing a well-defined function (in the mathematical sense). These are not teleological norms. Hence the explanations provided by computational cognitive science do not support informational

teleosemantics, as the conclusion of the methodological argument claims.[4]


3. A Decisive Victory in the Teleosemantic Toad Wars, but…

In chapter 5, entitled "Simple Minds", Neander tackles the disputed case of the frog and fly, first discussed in Lettvin et al. (1959). As is well known, frogs snap at small items crossing their visual field; often enough, in their natural environment, the item is a fly or other type of edible insect. Teleosemanticists argue that the internal state responsible for engaging the frog's tongue-snapping behavior represents whatever it has the natural function of detecting, but they disagree about what that is. Among the candidate contents are *bug*, *frog food*, and *small dark moving thing*, Neander's preferred content in her 1995 and 2006. Philosophical discussion rarely bothers with the empirical facts of the case, presuming that the matter can be settled without adverting to them. This chapter, which takes us through the specifics of toad prey-capture (very similar to what is known about frogs), supports Neander's earlier proposal. In fact, the chapter is a triumph of careful, detailed argumentation which should, in my opinion, close the books on the in-house debate among teleosemanticists. But I shall argue, contrary to what all teleosemanticists assume, attention to the details of the theory makes clear that there is no explanatory role for the notion of N-P function.

In summarizing the account I gloss over much interesting detail. The behavioral facts are roughly as follows: Toads typically orient toward prey-like stimuli, in particular, toward "worm-like*"* stimuli of constant width and varying length that move parallel to their longest axes. This is

---

[4] The norms implicit in the computational characterization are, in fact, sufficient to support a semantically evaluable notion of aboutness, though not a teleosemantic notion. The input states of the mechanism represent the *arguments* and the outputs states the *values* of the computed function. In the example described in the text the system *misrepresents* the sum of 43 and 57 as 99. See Egan 2014, 2018 for further discussion of these *mathematical contents*.

the *sign signal* – the configuration of environmental features – that trigger the response. Investigation of the toad's brain reveals that three kinds of retinal ganglion cells in the optic nerve, mediating between the retina and relevant mid-brain structures, respond differentially to moving stimuli, though their response patterns do not correspond to the toad's behavioral responses; further processing is required. The optic tectum in the toad's midbrain plays an especially important role. Adjacent retinal ganglion cells respond to adjacent retinal receptors, which are sensitive to light reflected from adjacent regions of the visual field, and they project to adjacent areas of the tectum. In effect, the optic tectum contains multiple maps of the visual field. The activation of a particular class of cells in the tectum – known as "T5-2" cells – produce directly, without further processing, the appropriate orienting response and are thought by neuro-ethologists to correspond most closely to the recognition of prey-like stimuli. As Neander puts it, "… this is their candidate for the relevant representation, whose content is debated by philosophers." (108)

The activation of T5-2 cells in the optic tectum is the most likely candidate for the representational *vehicle*. For convenience, Neander calls a T5-2 activation a "W" and then goes on to consider what the content of W-tokenings might be. There are at least two aspects to the content: *what* a target is represented as and *where* it is represented to be. T5-2 cells respond to stimuli in their specific receptive fields and so represent targets to be at the corresponding location in the visual field. Neander considers and decisively rebuts a suggestion in Dretske 1988 that the content of a representation is the state of affairs with the highest probability, given a tokening of the representation-type, arguing that the suggestion would favor a location content for Ws much less specific than the discriminative ability of the toad merits. Along similar lines, she argues that *what* Ws represent is a configuration of visible features, in particular, a *moving*

*worm-like stimulus*. This content ascription best captures what T5-2 cells respond selectively to –

a particular configuration of visible features that alternative content ascriptions such as *prey* or

*toad food* capture only very loosely or not at all, since if prey or toad food do not present as

moving worm-like stimuli they will not produce W-tokenings.[5]

In explaining the toad's prey-recognition capacity, the theorist must (1) isolate the neural

structures that play the appropriate role in mediating the prey-recognition process (T5-2 cells),

and (2) specify precisely the conditions of their activation (a moving worm-like stimulus). These

two tasks might be characterized, informally, in representational terms (but see below): (a),

corresponding to (1), identifying the structures that serve as the representational *vehicle*, and (b),

corresponding to (2), ascertaining their *content*. To support teleosemantics, the notion of normal

proper function must play an essential explanatory role in theoretical tasks (1) and (2). To

support the idea that *representation* is essential to the account, the characterization of T5-2 cell

activation as carrying representational content must bear a substantial explanatory burden

beyond that borne by (1) and (2), which do not advert to representational notions. I take up these

two points in turn.

It is not at all obvious that the notion of normal-proper function plays a substantive role

in either (1) or (2), though, as before, once the structures have been isolated and their activation

conditions specified, and with plausible further assumptions about the selection history of the

mechanism, its normal-proper function can be *recovered* from the account. In arguing against

alternative content proposals Neander says:

---

[5] *Moving worm-like stimulus* is close enough to *small dark moving thing* to vindicate the earlier suggestion in Neander 1995 and 2006.

> … an information-processing approach to explaining vision is not primarily concerned
> with the rational justification of behavior or with explaining why the visual capacity
> evolved. It is primarily concerned with explaining visual processing. (119)

This is exactly right, but it should be noted that the explanation of the visual processing underlying the toad's prey-capture capacity does not require appeal to selected functions; it requires appeal to the precise causal sensitivities of the cells, and to their role in processing – purely synchronic features of the mechanism – whatever the selection history of the mechanism might be. Neander points out that T5-2 cells are

> … *tuned to respond to* some things rather than others…. These pathways were adapted
> by phylogenetic natural selection and calibrated by ontogenetic maturational and learning
> processes to respond to certain specific types of stimuli in certain parts of the visual field.
> (114-115, emphasis in original)

Again, this is almost certainly right, but it is hard to see, either from Neander's excellent reconstruction of current theory about toad prey-capture, or from consideration of the information-processing approach to cognition in general, how these assumptions play a substantive role in the theory. Indeed, the normal-proper function of the T5-2 cells cannot be determined until *after* their response profiles have been mapped out, that is, until *after* the cognitive theorist has done the work required by tasks (1) and (2). At that point, the theorist's job – specifying the mechanism underlying the toad's manifest prey-recognition capacity – is essentially done. The account can be used to support further claims about the mechanism's normal-proper function, but Neander, and teleosemanticists in general, have the order of explanation backwards. There is no explanatory role for N-P function in the information-processing account.

I turn briefly to the question of whether the notion of *representation* is doing any essential explanatory work in the account. I do not think it is. Characterizing the neural structures that play the appropriate role in mediating the prey-recognition process – that is, T5-2 cells – as *representations* with the content *moving worm-like stimuli at location x* is, I suggest, just a convenient way of describing their response profiles. Neander argues convincingly that alternative teleosemantic accounts of the content of the toad's internal state – as representing *bug* or *toad food* – find no support from the science, because they are at odds with what it tells us about the toad's actual recognitional capacity.[6] They fail to capture the response profiles of T5-2 cells. But once the cognitive theorist has fully specified the state's causal role in the process the theoretical heavy lifting is done. Talk of the cell's activation *representing* its distal stimulus conditions (its sign signal) is best construed as a *gloss* that adds nothing of theoretical significance.[7] To be sure, characterizing the state in contentful terms supports attributions of *misrepresentation* in cases where, for some reason, the cell is activated by something other than a moving worm-like stimulus. I am not denying that there is some motivation for 'going representational' here. *Conceptual* thought processes are typically characterized in terms of their contents and so ascribing content to the non-conceptual processes described by cognitive scientists provides a common way of thinking about the two. Philosophers interested in cognition more generally will look for continuity with higher-level intentional thought. But the information-processing theorist tasked with explaining non-conceptual capacities has no such motivation. Everything she might want to say about toad prey-recognition, for example, can be

---

[6] This is not surprising, since Neander is alone among teleosemanticists in taking careful account of the empirical work on anuran prey-recognition.

[7] See Egan 2014, 2018 for the view that representational talk is best construed as a gloss serving various pragmatic purposes in computational models.

14

expressed more directly in straightforwardly causal terms.

References

Dretske, F. (1988). *Explaining Behavior.* Cambridge, MA: MIT Press.

Egan, F. (1995). "Computation and Content," *The Philosophical Review*, 104: 181-203.

Egan, F. (2014). "How to Think about Mental Content," *Philosophical Studies* 170: 115-135.

Egan, F. (2017), "Function-Theoretic Explanation and the Search for Neural Mechanisms,"
*Explanation and Integration in Mind and Brain Science*, David M. Kaplan (ed.), Oxford
University Press, 145-163.

Egan, F. (2018). "The Nature and Function of Content in Computational Models," *The Routledge
Handbook of the Computational Mind*, Matteo Colombo and Mark Sprevak (eds.),
Routledge, 247-258.

Gallistel, C.R. (1990). *The Organization of Learning*, Cambridge, MA: MIT Press.

Lettvin, J.Y., H.R. Matturana, W.S. McCulloch, and W.H. Pitts (1959). "What the Frog's Eye
Tells the Frog's Brain," *Proceedings the Institute of Radio Engineers* 47:1940-1951.

Marr, D. (1982). *Vision*, New York: Freeman.

Neander, K. (1995). "Misrepresenting and Malfunctioning," *Philosophical Studies* 79: 109-141.

Neander, K. (2006). "Content for Cognitive Science," in *Teleosemantics*, G. McDonald and D.
Papineau (eds.), Oxford, UK: Clarendon Press, 167-194.

Neander, K. (2017). *A Mark of the Mental*, Cambridge, MA: MIT Press.

Seung, S., Lee, D. D., Reis, B.Y., and Tank, D. W. (2000), "Stability of the Memory of Eye
Position in a Recurrent Network of Conductance-Based Model Neurons," *Neuron* 26:
259–71.

Shadmehr, R. and Wise, S. (2005). *The Neurobiology of Reaching and Pointing: A Foundation for Motor Learning*. Cambridge, MA: MIT Press.

Ullman, S. (1979). *The Interpretation of Visual Motion*, Cambridge, MA: MIT Press.