Kolors Without Colors, Representation Without Intentionality*

Angela Mendelovici and David Bourget

July 15, 2020

1 Introduction

Over the past few decades, the dominant approach to explaining intentionality has been a naturalistic approach, one appealing only to non-mental ingredients condoned by the natural sciences. Karen Neander's *A Mark of the Mental* (2017) is the latest installment in the naturalist project, proposing a detailed and systematic theory of intentionality that combines aspects of several naturalistic approaches, invoking causal relations, teleological functions, and relations of second-order similarity. In this paper, we consider the case of perceptual representations of colors, which is a challenging case for Neander's theory. This case will brings out a general methodological concern

---

*This paper is forthcoming in *Philosophy and Phenomenological Research* as part of a symposium on Karen Neander's *A Mark of the Mental*.

with Neander's and other naturalistic theories: these theories generally rest on the assumption that the mental intentionality we are acquainted with in everyday life—the phenomenon exhibited by desires for cups of coffee, perceptual experiences of dogs playing in yards, and thoughts about the weather—is the very same kind of phenomenon that cognitive science studies under labels such as "mental representation" and, in some cases, "information processing." This assumption is dubious, as the case of Neander's theory illustrates.

2 Neander's target and theory

Neander's target

Intentionality is, roughly, the "aboutness" or "directedness" of mental states. Neander gestures towards the phenomenon by appealing to everyday cases:

> When you taste a cup of coffee, recall a white sandy beach or imagine walking on Mars, your mental state represents, is about, or refers to the taste of the coffee, the white sandy beach, or your walking on Mars. Your thoughts have contents. In a way, nothing could be more familiar to us than this representational power of mental states, and yet its fundamental nature remains mysterious. (p. 1)

Neander claims that these cases exemplify the phenomenon Brentano famously described as "intentional inexistence," "reference to a content," and

"direction to an object." She takes *Brentano's problem* to be the problem of explaining intentionality.

Although Neander introduces intentionality using everyday examples and quotations from Brentano, her use of the term "intentionality" leaves open the possibility that the very same phenomenon exemplified by everyday cases is also present in the sometimes subpersonal or unconscious states posited by cognitive science. She writes, "This is in part just stipulative, but it is motivated by a desire to connect traditional discussions of intentionality with contemporary scientific explanations of what is at base the same subject matter." (p. 245–6, fn. 5)

Let us call the representation-like feature(s) posited by cognitive science *CS-representation*. The notions of intentionality and CS-representation might point to one and the same phenomenon. But they might not.

Neander's primary aim in the book is "to solve the part of Brentano's problem that is within reach" (p. 3), where the part of Brentano's problem that is allegedly within reach is that of explaining nonconceptual sensory-perceptual intentionality. This type of intentionality arguably includes most *original*, or underived, intentionality, with most other instances of intentionality ultimately deriving from it. A secondary aim of Neander's book is to understand the notion of representation implicit in cognitive science, i.e., the notion of CS-representation (see, e.g, p. 9). If we are right, Neander's theory does not succeed at its primary aim, though it may very well succeed at its secondary aim.

3

Neander's theory

Neander distinguishes between an account of representational status, which specifies when something has intentional properties, and an account of content determination, which specifies what determines an intentional item's content. Her theory of content determination can be summarized as follows, where R is a sensory-perceptual representation:

CT: If R is produced by a system whose function it is to respond to C by producing R, then R represents C.

CDAT: If R is produced by an internal system that has the function of producing analogs of external items in response to those external items, then R represents the external item of which it is an analog.

The Distality Principle: If CT or CDAT determines multiple contents, only the most causally distal of these contents is represented.

These theses are combined as follows:

Combined: If R is produced by an internal system that has the function of producing analogs of external items in response to those items, then R represents the most distal external item of which it is an analog. Else, if R is a sensory-perceptual representation produced by a system whose function is to respond to an item by producing R, then R represents the most distal such item.

The notion of a function invoked in CT is that of a "normal-proper" function, which is, roughly, one that is "malfunction-permitting" in that something might not fulfill its function. For example, a human heart has the (normal-proper) function of pumping blood. Neander endorses an etiological theory of normal-proper function, on which something's normal-proper function is what it was selected for doing.

CDAT appeals to the notion of an *analog*, where, roughly, A_i is an analog of B_i just in case they are corresponding items in analog systems of items $A_1...A_n$ and $B_1...B_n$, respectively. $A_1...A_n$ and $B_1...B_n$ are *analog systems* when the similarities and differences among $A_1...A_n$ (in certain respects) mirror the similarities and differences among $B_1...B_n$ (in certain respects). An *analog representation* is, roughly, a representation belonging to a system of representations that are analogs of their contents (p. 181).

According to Combined, representations satisfying the antecedent of CDAT get their contents by CDAT as constrained by the Distality Principle, while representations satisfying the antecedent of CT but not of CDAT get their contents by CT as constrained by the Distality Principle.

Neander's focus is on providing an account of content determination, but she also briefly suggests an account of representational status. On her account, a representational system is a system that evolved to mirror the world, and representations are the elements of such a representational system whose inter-relations are supposed to mirror the relations between things in the external world in the way specified by the correct theory of content

determination (pp. 178-9).

If we could combine Neander's view of content determination, her view of representational status, and an account of when a representation qualifies as sensory-perceptual, we would get an account of sensory-perceptual intentionality. Since the intentionality of sensory-perceptual states is supposed to be original intentionality, we would have at least a partial account of original intentionality (a full account would require a theory that covers *all* instances of original intentionality). However, it is not entirely clear to us what such a combined account would look like, since it is not clear how CT (or Combined) specifies a way that relations between representations mirror relations between things in the world. As a result, it is not clear that the correct theory of content determination (by Neander's lights) always specifies a way in which the relations between representations are supposed to mirror the relations between things in the world. In particular, it is unclear how non-analog representations get their representational status, since they are not covered by Neander's account of representational status. Still, the view is clear enough to make predictions about certain cases, which allows us to begin to assess it.

3 Assessing the theory: The case of color perception

One way to assess a theory of intentionality is to consider whether its predictions are correct. We will focus here on a case that we find particularly challenging for Neander's theory, that of perceptual representations of colors—*color representations*, for short. We will argue that the view makes false predictions about what they represent. We will see that the disagreement over color representations is based in a deeper methodological disagreement, which we will discuss in the next section and which will bring us to our central point.

How does Neander's view handle the case of perceptual color representations? At first blush, one might think that such color representations ought to get their contents from CDAT, since they seem to be analog representations—they seem to bear relations of similarity and difference to one another that mirror relations of similarity and difference between their contents (i.e., colors). But color representations are not produced by internal systems having the function of producing analogs of external items in response to those items. They are produced by systems having the function to respond to certain external items—presumably surface reflectance properties or the like—but color representations are not analogs of these external items. This is because, as is well known and as Neander accepts, there is a structural mismatch between color representations and surface reflectance properties. For

example, the representation of a particular shade of violet and a particular shade of red are more similar to one another than either is to the representation of a particular shade of yellow, but the surface reflectance properties corresponding to the representations of violet and red are more similar to the surface reflectance property corresponding to the representation of the yellow than they are to each other (see Hardin 1988).

Since color representations are sensory-perceptual representations that do not satisfy the antecedent of CDAT, Combined says that they (at best) get their contents from CT. So, what contents does CT predict? Presumably, color representations are produced by a system whose function it is to produce them in response to particular surface reflectance properties, which Neander calls "kolors" in order to distinguish them from the potentially distinct contents of color terms and concepts (p. 165). So, Neander's theory predicts that color representations represent kolors.

This, however, is the wrong answer. One reason to think that color representations do not represent kolors is that the contents of color representations and kolors exhibit a *structural* mismatch. As noted above, color *representations* are not analogs of kolors. It is even more obvious that color *contents* are not analogs of kolors. For example, the content of a representation of a particular shade of red is more similar to the content of a representation of a particular shade of violet than it is to the content of the representation of a particular shade of yellow, but the kolor properties assigned by Neander's theory to the representations of red and violet are less similar to one another

than either is to the kolor property assigned to the representation of yellow. Since the space of kolors is organized differently than the space of the contents perceptual color representations represent, the two spaces are not identical, and the contents of perceptual color representations are not kolors. Thus, Neander's view falls prey to Pautz's (2006) *structural mismatch problem.*

Another reason to think that color representations do not represent kolors is that the contents of color representations and kolors exhibit a *qualitative* mismatch. Introspective and broadly psychological considerations having to do with the psychological role of perceptual color representations do not suggest that perceptual color states represent anything to do with particular surface reflectance properties. Instead, these considerations suggest that they represent primitive, nonrelational, nondispositional, and qualitative color properties—roughly, what Chalmers (2006) calls "edenic colors," colors just as they appear to be. Edenic color properties are arguably uninstantiated, but this does not prevent us from representing them.[1] Since there is a qualitative mismatch between kolors and the contents of color representations, the two are not identical and color representations do not represent kolors. Thus, Neander's view also succumbs to Mendelovici's (2018) *qualitative mismatch problem.*

Before considering Neander's likely responses to these mismatch worries,

---

[1]It also does not prevent color representations from being useful in helping us get by in the world. This is because color experiences do not merely misrepresent but rather *reliably* misrepresent, where *reliable misrepresentation* is a matter of gettings things wrong in the same way all the time, and reliable misrepresentation, in virtue of its reliability, can be as useful as reliable veridical representation (Mendelovici 2013, 2016, and 2018).

it is worth mentioning some further related concerns with taking colors to represent kolors rather than something like edenic colors. One is that, as noted above, color representations turn out not to be analog representations, since the relations of similarity and difference between color representations do not mirror the relations of similarity and difference between kolors. This is an odd result since color representations are naturally thought of as analog in some straightforward sense. Assuming the opponent processing model of color vision, the vehicles representing red and violet both involve activation of the "red channel," while the vehicle representing yellows do not. It is tempting to say that this mirrors the greater similarity in hue between the represented colors red and violet than between either of these colors and represented yellow. Even without considering the specifics of color processing, one might argue that the systematic organization of perceptual color space suggests an analog representational system. Since the structure of edenic color space arguably mirrors that of color representations, taking color representations to represent edenic colors allows us to preserve the claim that color representations are analog.

Another related worry is that if color representations are not analog, then Neander's theory cannot accommodate "novel" color contents, color contents that were not present in function-conferring circumstances for the systems producing color representations. Neander is aware of worries surrounding the representation of such novel contents and takes accommodating such contents to be one of the central motivations for supplementing CT with CDAT.

CDAT determines the contents of representations "en masse," thereby allowing for the representation of novel contents, contents that never occurred in function-conferring circumstances for the relevant representational systems.[2] Neander uses Hume's example of the missing shade of blue to highlight this need for a way of representing novel contents. Adapting the example to be directly relevant to our discussion, we can imagine a shade of blue that was never encountered in function-conferring circumstances for the system that produces color representations. If CDAT applied to color representations, it could allow us to represent this missing shade of blue. But color representations don't get their contents from CDAT, so this purported benefit is foregone in the case of color representations. Neander is not too concerned about this, dismissing Hume's case as "too far-fetched to worry us." (p. 201) However, it is not at all implausible that some of the colors we represent never occurred in function-conferring circumstances—perhaps some neon colors, supersaturated colors, or even unremarkable everyday colors never occurred in the relevant circumstances. Neander's view cannot accommodate the representation of any such colors. In contrast, taking color representations to represent edenic colors allows us to say that they are analog representations whose contents can be determined en masse, if not by CDAT, then perhaps in some other way.

---

[2]However, it is not clear that CDAT in fact has this benefit (Mendelovici and Bourget 2019).

4   Theory-independent access to intentionality

We have argued that Neander's theory makes false predictions about the contents of color representations. Neander, however, embraces these predictions.[3] In order to adjudicate such a disagreement, we need a theory-independent way of finding out what representations represent. Neander is aware of this need and devotes an entire chapter of her book to the details of toad prey-capture in an effort to draw out general methodological principles that can help provide a theory-independent way of settling disputes over what a representation represents. In this section, we suggest that while Neander's metholodogy might provide an appropriate theory-independent way of studying CS-representation, it does not provide an appropriate theory-independent way of studying intentionality.

Neander expresses her general methodological approach as follows:

> If we are interested in the real nature of intentional phenomena, the content ascriptions entailed by our theory of mental content should cohere with the best explanation of the relevant capacities that science can provide, unless we have strong reasons to believe that even the best science to date is on the wrong track. I continue to assume that an information-processing approach is the best available approach... (p. 98)

_____

[3]In a discussion of color realism, she grants that kolors might not be the contents of color terms and concepts, but maintains that they are the contents of perceptual color representations (pp. 164–5).

The basic idea is that our best scientific explanations, and in particular those of information-processing approaches to cognition, help provide a theory-independent way of finding out what representations represent.

Neander suggests two more specific principles that are motivated by this general approach:

P1   Perceptual representations must be normally causally sensitive to what they represent.

P2   "[V]isual content must be extracted from the retinas by subsequent processing." (p. 116)

According to P1, we cannot perceptually represent something that we are not normally causally sensitive to. According to P2, retinas carry information about configurations of visible features—features that can be detected visually, like shapes and locations—rather than invisible features, like nutritional content or natural kind membership, and subsequent processing can infer the presence of invisible features only if the inference is ultimately based on the configuration of these visible features.

It is not entirely implausible that P1 and P2 are appropriate theory-independent constraints on CS-representational content attributions. If we understand an "information-processing" story as a story of how transduced stimuli are transformed to extract information about invariant features that are biologically important, P1 and P2 make sense as constraints on content attribution for an information-processing story. Such a story will arguably

be in some way constrained by the environmental features we are causally sensitive to and the information that can be extracted from the states of the transducers.

The central question for us, however, is whether P1 and P2 are adequate theory-independent constraints on intentional content attributions. We suggest that they are not, since they are not theory-independent. P1, in particular, is effectively a vague statement of a causal/informational theory of content. Anyone who rejects such a theory—say, in favor of a conceptual role or phenomenal theory of intentionality—is unlikely to accept P1. Similarly, if the term "extracted" in P2 refers to a process whereby a representation is created that carries part of the (Dretske-style) information found in another representation, P2 also presupposes a broadly informational/causal conception of content. P1 and P2 reflect Neander's substantive assumptions regarding the correct approach to intentionality; they do not offer anything like theory-independent constraints on content attribution.

This is not to say that intentional ascriptions should not "cohere with the best explanation of the relevant capacities that science can provide" (p. 98). Unless we outright assume that the notion of representation invoked in an information-processing story points to the same thing as our pre-theoretic notion of intentionality, we cannot know from the outset that an information-processing story offers the relevant scientific explanations with which our theory of intentionality should cohere.

We suggest, instead, that a fairly innocuous form of introspection and con-

siderations of psychological role, broadly construed, can provide us theory-independent ways of assessing competing theories' predictions (see also Mendelovici 2018, ch. 2). Recall that our target is the everyday/Brentanian phenomenon of intentionality, the phenomenon exhibited by everyday cases like those of recalling a white sandy beach or imagining walking on Mars. Our pre-theoretic grip on intentionality is via introspection. This grip not only tells us *that* we have intentional states but also, in at least some cases, gives us access to *which* intentional states we have (though it need not give us insight into their deep nature). For instance, we can introspectively tell that we are thinking about a white sandy beach rather than the insides of a volcano or the state of the stock market. This is something everyone should accept. Denying such claims takes us dangerously close to losing sight of the phenomenon of interest, intentionality in the everyday/Brentanian sense.

It should also be relatively uncontroversial that intentional states at least in some cases play certain psychological roles that are appropriate to their contents, including roles in generating further mental states (including higher-order mental states) and behavior and roles in contributing to our overall phenomenology, e.g., of our phenomenal experiences or experienced "grasp" of certain contents. For example, a perceptual experience of a red tomato before you is likely to lead to a belief that there is a red tomato before you, the higher-order thought that you are perceiving a red tomato, and tomato-appropriate behaviors, and it is likely to be accompanied by an experience of "grasping" the content *there is a red tomato before me*. Consid-

erations of psychological role give us some purchase on our own and others' mental contents.

In the previous section, we suggested that perceptual color representations represent edenic colors and not surface reflectance properties, as Neander's theory predicts. Our theory-independent considerations support this claim: In the same way that we can tell that we are thinking about a white sandy beach rather than the insides of a volcano, we can tell that we are perceptually representing something like edenic colors rather than dispositions to reflect light in certain ways. Introspection reveals contents involving bluish, reddish, and other color qualities that are in no way captured by content attributions in terms of kolors—i.e., surface reflectance profiles—or related physical features of putatively colored objects. Likewise, our color representations do not behave as if they represented kolors: they do not cause beliefs about particular kolors, higher-order thoughts about representing kolors, or behavior aimed at such kolors. Instead, our color representations arguably behave as if they represented edenic colors. If all this is right, then Neander's theory makes incorrect content attributions after all.

5 Conclusion

Neander's primary aim in *A Mark of the Mental* is to provide an account of nonconceptual sensory-perceptual intentionality, which encompasses most original intentionality. Her secondary aim is to understand the notion of rep-

resentation implicit in cognitive science, i.e., the notion of CS-representation. We have argued that Neander does not succeed in her primary aim, though she might succeed in her secondary aim. This is not to suggest that her view is an abject failure but rather that its focus should be redirected. Considered as a theory of CS-representation, it provides an exemplarily systematic, empirically-informed, and well-supported picture.

Worries with capturing the everyday/Brentanian phenomenon of intentionality are not special to Neander's view. Most naturalistic theories face similar problems, though they too might offer promising accounts of CS-representation. So, what's a naturalist to do? We would suggest dropping the assumption that the everyday/Brentanian phenomenon is of a piece with the notion of representation implicit in cognitive science. Once free from this assumption, we can study each without importing assumptions about the other. For example, we need not import the assumption that content attributions ought to be determinate from the case of intentionality to that of CS-representation—perhaps CS-representational contents admit of some indeterminacy (see also Cummins 1994). Rejecting the assumption that CS-representation is of a piece with intentionality might help us better understand the former and give us a more realistic sense of the progress that has been made on the latter.

References

Chalmers, D. J. (2006). Perception and the fall from eden. In Gendler, T. S. and Hawthorne, J., editors, *Perceptual Experience*, pages 49–125. Oxford University Press, Oxford.

Cummins, R. (1994). Interpretational semantics. In Steven, P. S. and Ted, A. W., editors, *Mental Representation: A Reader*, pages 278–301. Blackwell, Oxford.

Hardin, C. L. (1988). *Color for Philosophers*. Hackett.

Mendelovici, A. (2013). Reliable misrepresentation and tracking theories of mental representation. *Philosophical Studies*, 165(2):421–443.

Mendelovici, A. (2016). Why tracking theories should allow for clean cases of reliable misrepresentation. *Disputatio*, 8(42):57–92.

Mendelovici, A. (2018). *The Phenomenal Basis of Intentionality*. Oxford University Press.

Mendelovici, A. and Bourget, D. (2019). Review of Karen Neander's *A Mark of the Mental. Philosophical Review*, 128(3):378–385.

Neander, K. (2017). *A Mark of the Mental: A Defence of Informational Teleosemantics*. Cambridge, USA: MIT Press.

Pautz, A. (2006). Sensory awareness is not a wide physical relation: an empirical argument against externalist intentionalism. *Noûs*, 40(2):205–240.