No Evidential Test for Naturalistic Theories of Content

Naturalistic Theories of Content (NTCs) aim to account for representational content in non-intentional, non-semantic terms. *Varitel semantics* (Shea, 2018) and *informational teleosemantics* (Neander, 2017) are two recent proposals offering contrasting views about content. On what grounds should we adjudicate among NTCs? One influential response in the literature is to put the NTCs to an *evidential test*, i.e., to a comparison between the content ascriptions predicted by the NTC and those found in our best scientific explanations. I assess two common ways to put NTCs to an evidential test: by comparing them with (1) explicit scientific content ascriptions, and (2) reconstructed content ascriptions. I highlight issues related to both approaches. I argue that (1) does not provide unambiguous content ascriptions and that (2) relies on the very theories it is meant to test. In the absence of an evidential test, philosophers should think about new ways to adjudicate among NTCs.

The notion of an evidential test captures a common strategy when discussing NTCs. Philosophers interpret convergences between NTCs' claims and scientific content-ascriptions as confirming the NTC. "If the scientists' ascriptions concur with the ones our theory entails, this can help to confirm the theory" (Neander, 2017, p. 115). Conversely, divergences are taken to signal an issue for the theory. This approach approximates what I call an *evidential test* for NTCs. An evidential test is a comparison between the content ascription predicted by the NTC in a given situation and that assumed by scientific theories in the same situation. I define what it means for *ST (scientific theory) to constitute an evidential test for an NTC*:

i. The NTC predicts that a mental state S has content X in circumstances C

ii. ST claims that S has content Y in circumstances C, and would ascribe the same content if C were to occur again (robustness requirement)
iii. ST is a reliable source of evidence about content (reliability requirement)
iv. ST is independent from NTC (independence requirement)

v. NTC is confirmed if, given C, X = Y,
vi. NTC is disconfirmed if, given C, X ≠ Y.

"Independence" means that ST's content ascriptions should not depend on NTC. For example, a change in NTC should not modify any of ST's content ascriptions.

Could *explicit scientific content ascriptions* form an evidential test for NTCs? In explaining behavior, scientists mention explicit representational contents, which they attribute to internal vehicles or to steps in computational explanations. Why not test NTCs against them? I raise two problems.

First, whether scientific explanations involve representations or not is not transparent. Psychologists and neuroscientists use highly context-sensitive expressions, like "*response", "representation", "being elicited by", "being about", "recognizing", "signal", "being sensitive to"*. Those terms may or may not denote genuine representational talk. Moreover, neuroscientists, psychologists and philosophers tend to be uncertain about how to apply the concept of representation, suggesting that the very concept of representation is imprecise (Favela & Machery, 2023).

The second problem is that explicit content ascriptions vary across and within studies. For example, studies on A2 cells in the auditory system of the moth tend to provide inconsistent content ascriptions. While one study mentions "bat's echolocation calls" (Ratcliffe et al. 2009), another refers to a "neural representation of bat predation risk" (Goerlitz et al., 2020) and yet another mentions ultrasound (ter Hofstede et al., 2013). These variations show that explicit content ascriptions are inconsistent across studies, and therefore fail to meet the robustness requirement (ii).

If *explicit scientific content ascriptions* do not constitute a good evidential test for NTCs, then *reconstructed content ascriptions* might. Moving away from explicit ascriptions to reconstructed ascriptions is common in the literature (Neander 2017, Shea 2018). These reconstructions typically insist on the explanatory role of content. I suggest that some reconstructions are subject to a circularity risk: they are partly determined by the theory they are meant to test. These reconstructions then fail to meet the independence requirement (iv). I focus on Neander (2017).

Neander reconstructs the explanatory role of content by analyzing neuroethology explanations of how the toad detects "wormlike" stimuli (Camhi, 1984). Since these explanations describe toads' visual T5-2 neurons as causally sensitive to worm-like stimuli rather than to prey, Neander concludes that the content playing an actual explanatory role is worm-like stimulus, not prey. Yet this conclusion cannot be reached from these considerations alone. It needs the additional premise that content implies a *causal sensitivity* to the represented condition (Neander, 2017, p. 156). This additional premise does not derive from the scientific explanation, but from Neander's theory of content. The reconstruction therefore fails to meet the independence requirement.

What should we do then? One option is to give up on the idea that content properties are objective. Deflationism (Egan 2020, 2025) suggests that content properties

are ultimately determined by research interests. From the deflationist's perspective, the divergences among NTCs stem naturally from the plurality of research interests. Deflationism does fit some of the observations made above about inconsistent content ascriptions and imprecision in representational talk within the scientific literature. However, there remain strong reasons to stick to objective representations, not least because mainstream psychology routinely appeals to accuracy conditions in its explanations (Burge 2010).

Another option would be to give up on the idea of evidential tests. Theories can be assessed through a great variety of means, including internal coherence, external coherence, or simplicity. The idea of testing theories of content against scientific explanations remains appealing, if only in the minimal sense that NTCs should not contradict our best scientific explanations.

Finally, it must be noted that while Neander (2017) is subject to the circularity risk, Shea (2018) is more cautious and makes no explicit appeal to evidential tests. Further work should expand on the different ways that NTCs relate to scientific explanations.

References:

Burge, T. (2010). *Origins of Objectivity*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199581405.001.0001

Camhi, J. M. (1984). *Neuroethology: Nerve Cells and the Natural Behavior of Animals*. Sinauer Associates Inc.

Egan, F. (2025). *Deflating Mental Representation*. The MIT Press. https://doi.org/10.7551/mitpress/15666.001.0001

Favela, L. H., & Machery, E. (2023). Investigating the concept of representation in the neural and psychological sciences. *Frontiers in Psychology*, *14*.

Goerlitz, H. R., Hofstede, H. M. ter, & Holderied, M. W. (2020). Neural representation of bat predation risk and evasive flight in moths: A modelling approach. *Journal of Theoretical Biology*. https://doi.org/10.1016/j.jtbi.2019.110082

ter Hofstede, H. M., Goerlitz, H. R., Ratcliffe, J. M., Holderied, M. W., & Surlykke, A. (2013). The simple ears of noctuoid moths are tuned to the calls of their sympatric bat community. Journal of Experimental Biology, 216(21), 3954–3962. https://doi.org/10.1242/jeb.093294

Neander, K. (2017). *A Mark of the Mental: In Defense of Informational Teleosemantics*. MIT Press.

Ratcliffe, J. M., Fullard, J. H., Arthur, B. J., & Hoy, R. R. (2009). Tiger moths and the threat of bats: Decision-making based on the activity of a single sensory neuron. *Biology Letters*, *5*(3), 368–371. https://doi.org/10.1098/rsbl.2009.0079

Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press.