Theories of Mental Content, Part II:
Teleological Theories of Content

> **Million Dollar Question:** What explains the intentionality of mental states? What explains why an individual's thoughts/beliefs/intentions succeed in representing what they represent?

Last week, we looked at one popular approach to explaining the intentionality of mental states: *causal theories*. This week we'll look at an alternative approach – *a teleological approach*.

The main idea behind a teleological approach is to explain the content (aboutness) of a mental representation in terms of that representation's *function*. To introduce the teleological approach, it will be helpful to first go over a close variant of the causal approach: an *indication theory*.

1. Indication Theory of Mental Representation

Recall Grice's notion of "natural meaning". Natural meaning arises when some object/event *indicates* – or carries information about - the state of the environment.

Examples:

- Smoke indicates (/means$_N$) fire.
- Those spots indicate (/means$_N$) measles.
- The rings on a tree trunk indicate (/means$_N$) that the tree has a certain age.

> **Indication:** A indicates B if and only if whenever A is present, B is also present.

Q: Can you think of any other examples of indication?

Some people have suggested that we can understand intentionality in terms of indication:

> *Simple Indication Theory:* A representation M is about some entity X if and only if tokenings of M indicate X.
>
> - For example, my concept DOG is about dogs because whenever I token the concept DOG, there is a dog present.

Two comments:

- Note that this basically amounts to the idea that mental representations can be understood in terms of natural meaning. So if you combined this view with Grice's project for reducing linguistic intentionality to mental intentionality, you would get the view that all non-natural meaning ultimately can be explained in terms of natural meaning.

- Note the Simple Indication Theory is very similar to the causal theory. Can you think of any cases where they come apart – i.e., where there is indication without causation?

Given the Indication Theory's close connection to the Simple Causal Theory, it runs into many of the same problems.

1) *Problem of misrepresentation:* If my concept DOG is sometimes caused by foxes or funny-looking cats, then my concept DOG doesn't indicate dogs (given the def. of indication). But then, by the lights of the Simple Indication Theory, my concept DOG isn't about dogs.
2) *Disjunction problem:* Doesn't explain why my concept DOG isn't about *dogs-or-foxes*, or *dogs-or-retinal-images of dog.*

Further difficulty:

3) *Overproliferation of content:* The fuel gauge indicates the level of fuel. It also indicates the presence of a large downward force on the car frame. Bu intuitively the fuel gauge only *represents* the level of fuel; it doesn't represent the presence of a large downward force.
   - Q: is this also a problem for causal theories of content?

2. From Indication Theories to Teleological Theories

Proposed solution: Appeal to functions!

Teleological-Indication Theory: A representation M is about some entity X if and only if the *function* of M is to indicate X.

- For example, my concept DOG is about dogs because it is the function of my concept DOG to indicate dogs.

How it helps:

1) Allows for misrepresentation, since a representation could fail to fulfill its function.
2) Given an adequate theory of functions, hopefully the theory could predict that the function of my concept DOG is to indicate dogs rather than to indicate *dog-or-retinal-image-of-a-dog.*
3) Given an adequate theory of functions, hopefully the theory avoids overproliferation: the function of the fuel gauge is to indicate the fuel level, not the presence or absence of a downward force on the car frame.

Big Question: Where do these functions come from? Why is the function of my concept DOG to indicate dogs, rather than foxes (or trees, or…)?

3. Functions: A Closer Look

One way of trying to understand functions is to look at natural selection. This route is explored by Millikan, who is perhaps the most prominent defender of a teleological approach.

An example: the heart does a lot of things. It:

- pumps blood
- it makes a rhythmic, percussive sound
- serves as a food to certain types of carnivores.

But plausibly it was only selected for one of these things – pumping blood. This is its *biological function*.

> *Biological functions:* The biological function of a trait T is Z if and only if T was selected for Z-ing.

How would this apply to the case of mental representations? An illustration:

> Suppose our ancestors had a state in their head that tended to be caused by predators. Suppose this state caused them to flee, which led them to survive. Then the biological function of that state is to indicate predators. If we inherited the same state from our ancestors, then the function of this brain state in us is to represent predators.

Q: Is Teleological Indication + Biological function > Simple Indication Theory?
That is, suppose one combines the Teleological Indication Theory with a biological theory of functions. Does this solve the three problems for the Simple Indication Theory? Does it face any additional problems?

4. Millikan's Theory

Millikan's theory is similar to the Teleological Indication Theory, but it is a bit more complicated. Millikan stresses the distinction between the *producers* of representations and the *consumers* of representations.

- In some cases, these will be distinct individuals – as one when one beaver (*producer*) splashes its tail in the water (*representation*), which serves to warn the other beavers (*consumers*).
- In other cases, these will different systems within a single individual, as when say a perceptual system produces a representation, which is then used by other systems of the organism (for example, the digestive system).

On Millikan's view, in order to determine the content of a representation, we need to look at the functions of its consumers.

Roughly, the idea is:

> A representation M is about some entity X for a consumer system S iff S has whatever function it has only because in the past, there was a reliable enough correlation between M and X

where "reliable enough" means "reliable enough to give S its function" – compatible with being highly unreliable (cf. predator detection systems)

Q: Are there any advantages of Millikan's proposal over the Teleological Indication Theory?

5. Challenges to Teleological Theories

Braddon-Mitchell & Jackson raise the following objection:

- Imagine we develop a technology for prosthetic brain states. These prosthetic brain states were never selected for (at least in the biological sense: they did not confer reproductive fitness on our ancestors). But wouldn't they still have content?

A somewhat related – and much more widely discussed – challenge to teleological theories comes from *Swampman*:

> Suppose a swamp is struck by lightning. In an extremely improbable event, the lightning triggers a chemical reaction which causes the molecules in the swamp to reassemble in a molecule-for-molecule copy of some human being. This "swampman" will act in ways that are indistinguishable from his real-life doppelganger: he'll eat when he's hungry; he'll tell stories about his past, etc. The teleological account seems to entail that he doesn't have any mental states with content: he might go out to a restaurant when he's hungry, but this isn't because he believes that the restaurant serves food, or because he desires to have food, etc. Many find this counterintuitive.

Question for discussion: Is the swampman objection convincing? Can you think of any way of defending a teleological approach in the face of this objection?