

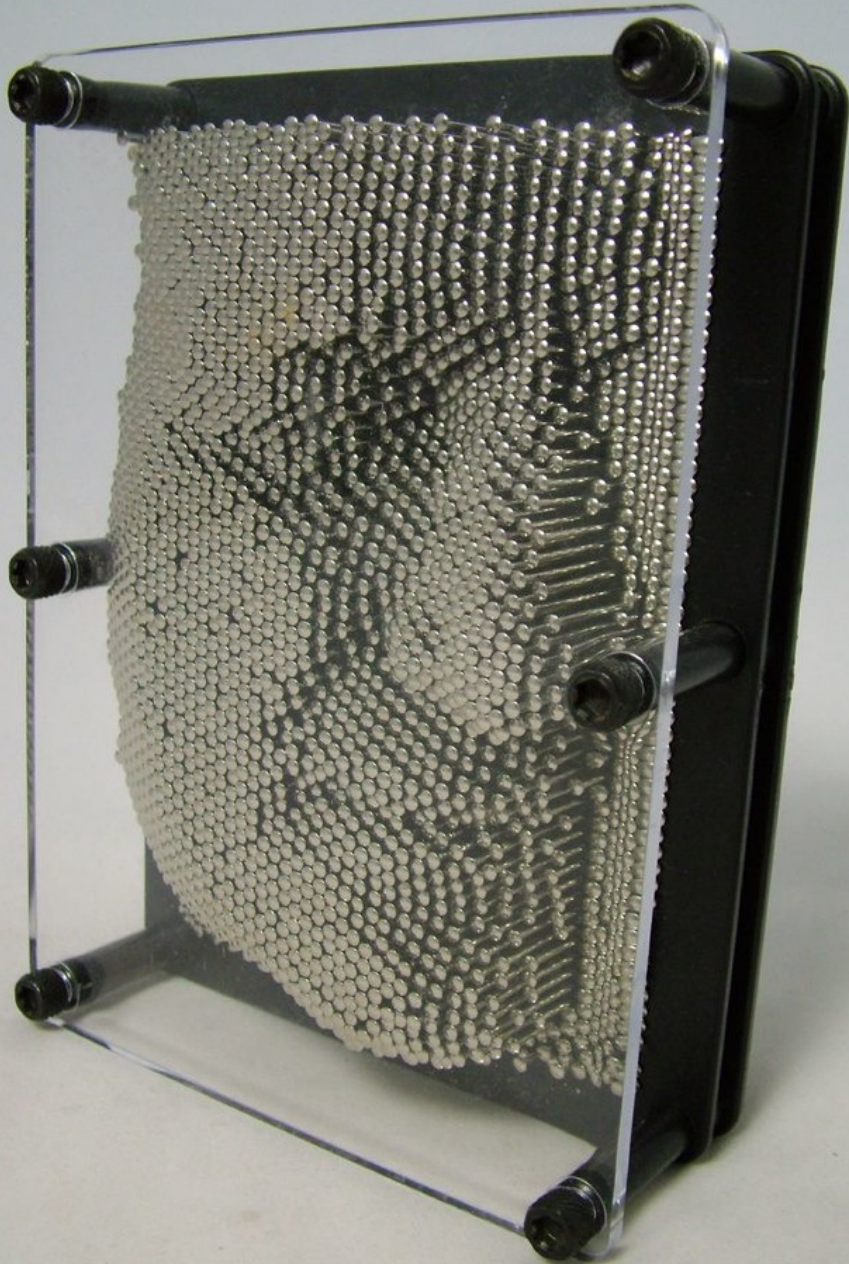
**“Representation” means  
exactly what you think it means**

Yohan J John

Research Assistant Professor

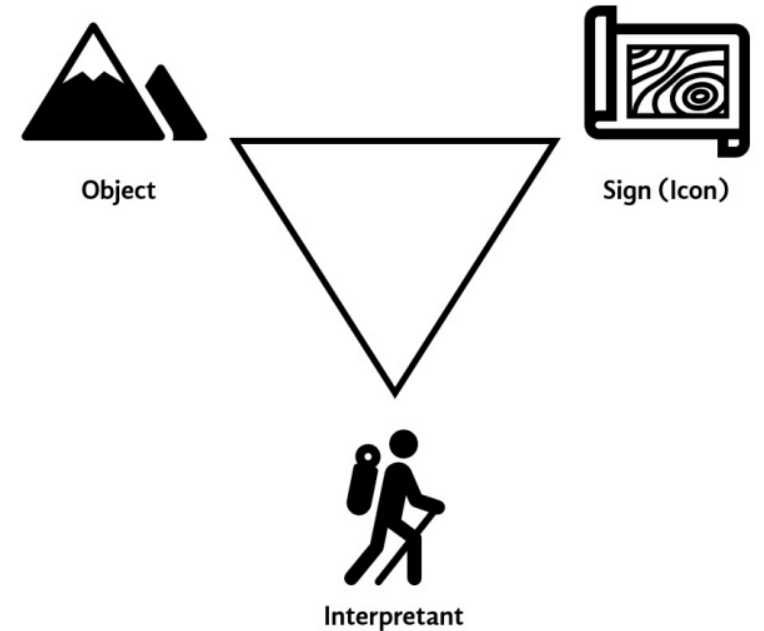
Neural Systems Laboratory

**BOSTON  
UNIVERSITY**



# How should we represent “representation”?

1. Establish the **colloquial, literal meanings** of “representation” outside of science, before applying it in a metaphorical way to mind/brain
2. Examine the **history** of the use of the word “representation”
3. Acknowledge that representation is often a **placeholder**, and that is this to be expected before consensus theories are established
4. Use computational models that close the perception-action loop to identify how **representation can be more than a placeholder**



# 1. What do *most* people mean by “representing”?



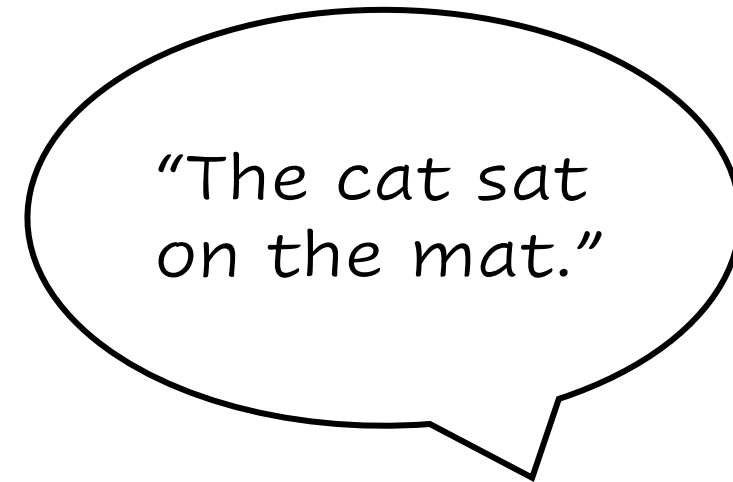
“The **painting** is a representation of **animals** to **the viewer**.”



“The **statue** is a representation of **the Buddha** to **the viewer**.”



“The **record** is a representation of **music** to **the record player**.”



“The **utterance** is a representation of **a scene** to **the listener**.”

“The **printed sentence** is a representation of **an utterance** to **the reader**.”

“**X** is a representation of **Y** to **Z**.”

“**X** is a map of **Y** for **Z**.”

# 1. Can we define “representation” before using it metaphorically?



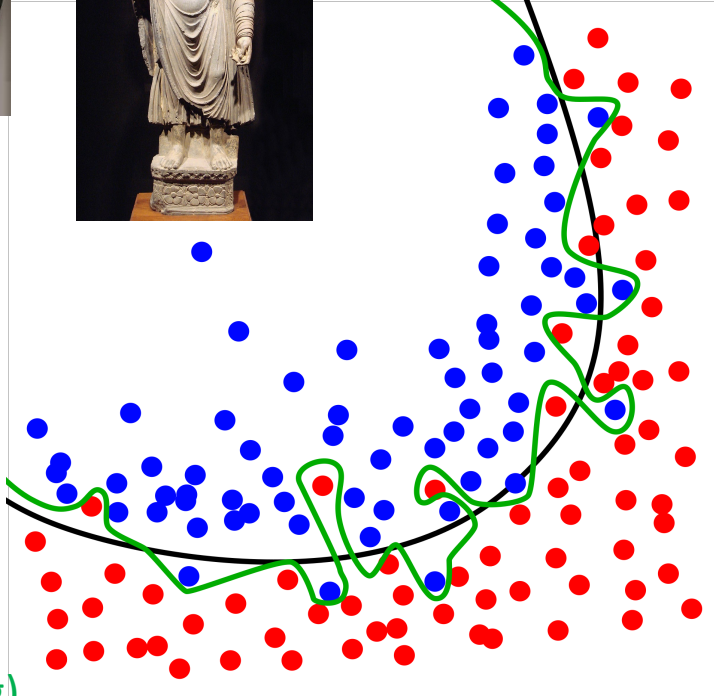
Useful intensional definition



“The cat sat on the mat.”



Nitpicky intensional definition (overfitting)



“The **painting** is *not* a representation of **anything** to **the baffled museum-goer**.”

*Extensional definition*: a list of examples of correct (●) and incorrect (●) uses of “representation”

*Intensional definition*: a set of necessary and sufficient conditions for the use of the term “representation”

# 1. Intensional definitions of “representation”

All uses of representation seem to imply:

**X**: an object/phenomenon doing the representing

**Y**: an object/phenomenon being represented

**Z**: an interpretant of X

Representations occupy a spectrum between two poles:

Representations based on **similarity**, which can be recognized cross-culturally without learning a convention or interpretation system

Representations based on **arbitrary convention**, which require learning/decoding

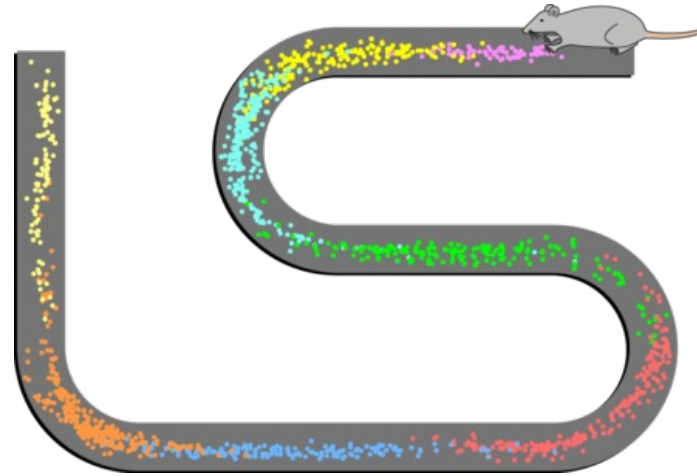
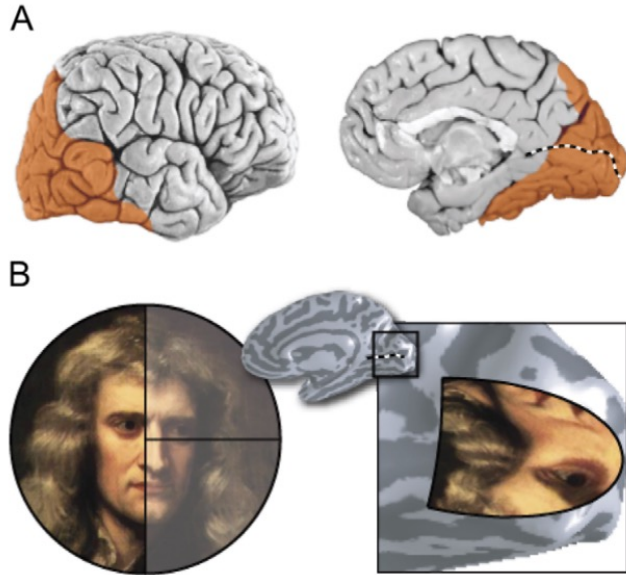


Similarity-based: The **main etching** represents **a seated person** to **many viewers**

Convention-based: The **markings on the top** represent **(unknown) conventional symbols** to **many viewers**



# 1. Which of these uses is confusing/controversial?



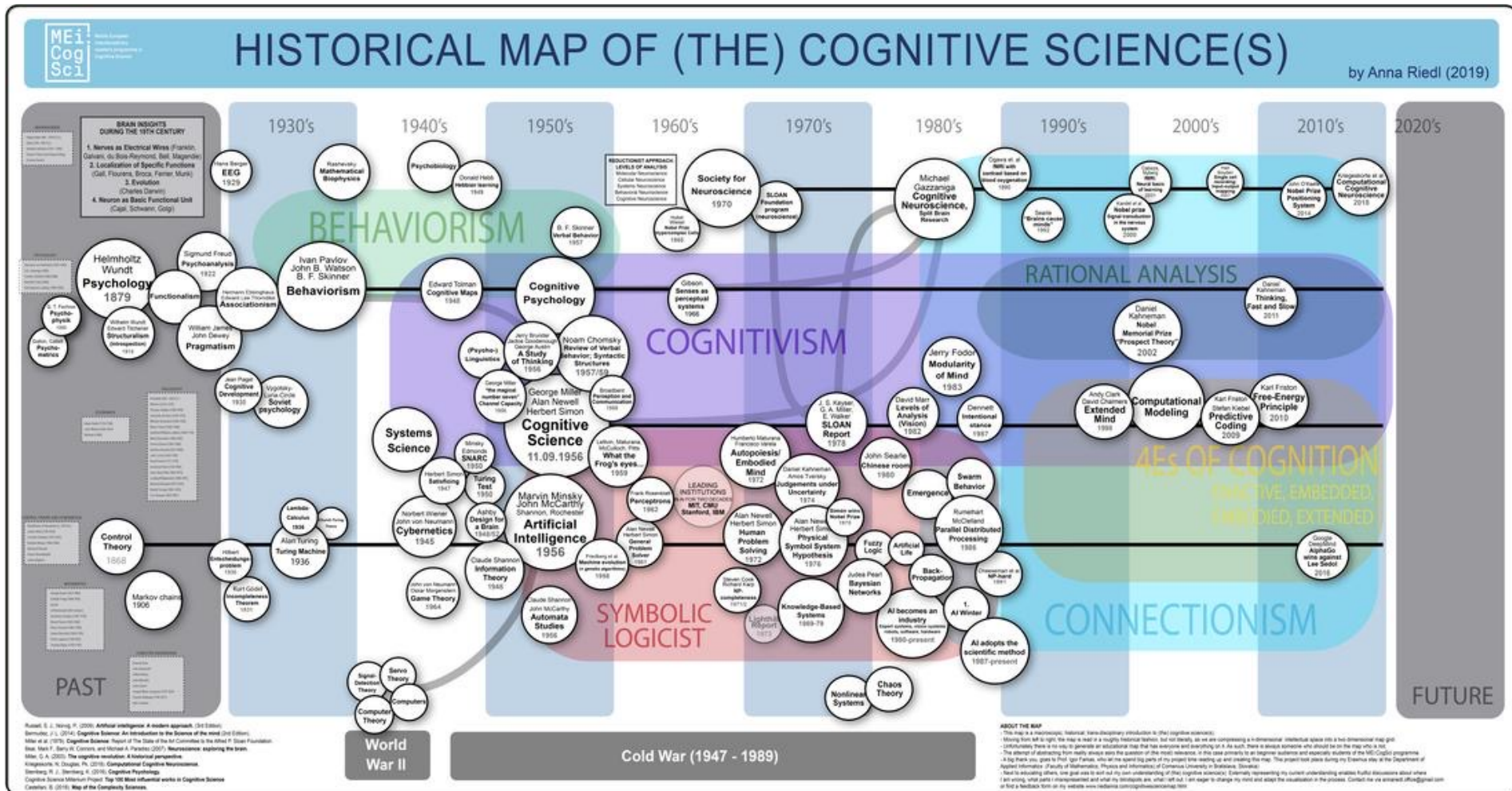
“X is a representation of Y to the rest of the brain.”

If someone claims that V1 does *not* represent anything, do they mean:

- The **pattern of firing** in V1 does not resemble the **pattern of light on the retina**? (*Demonstrably false*)
- The **rest of the brain** does not receive/use the signals coming from V1? (*Demonstrably false*)

What *else* might it mean to say that V1 does not represent?

# 2. Where did (anti)representationalism come from?



“This map is a representation of the history of cognitive science to us.”

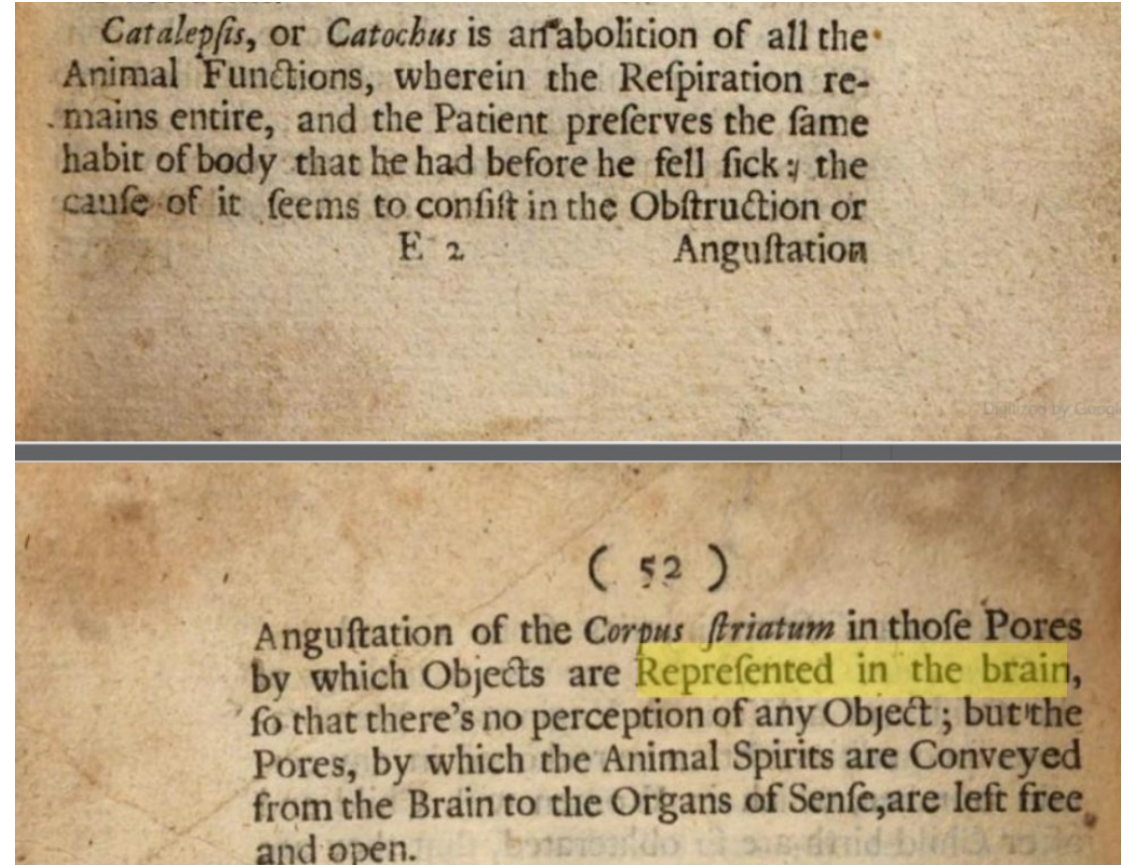
<https://www.riedlanna.com/cognitivesciencemap.html>



## 2. Cogitivism = Symbols = Representations ?

This narrative seems very common among critics of representation, but it may be biased in several ways including:

- Overreliance on recent history (post WWII)
- Overreliance on debates from psychology and philosophy
- Neglect of the older, looser uses of “representation”, especially in medicine
- Neglect of actual usage by neuroscientists



*Lexicon medicum græco-latinum. A Physical Dictionary; in which all the terms relating either to anatomy, chirurgery, pharmacy, or chymistry, are very accurately explain'd* - Stephen Blancard (1684)

# Interlude: “No information without representation”



The Indus Valley symbols represent written language to some (but not all) researchers.

Do they contain/convey information?

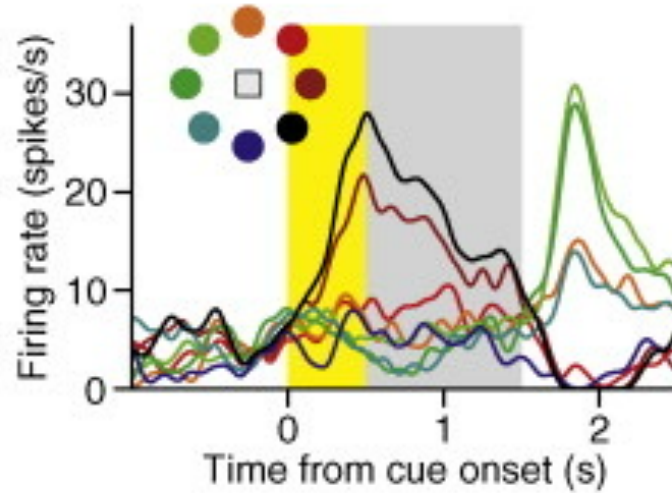
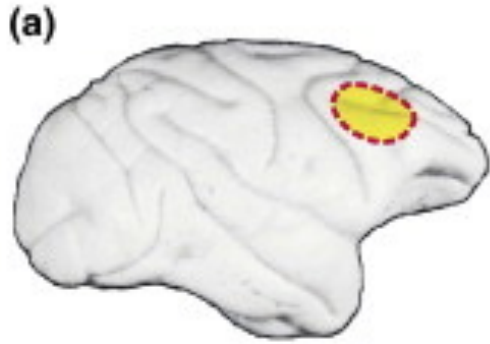
Information is a popular concept that is used by some critics of representation

What is the *difference* between “X represents Y to Z” and “X conveys information about Y to Z”?

How can information *out there* in the world influence action (e.g., via the motor cortex) without being somehow made *present* (represented) in the brain?

As in the case of the undeciphered Indus Valley symbols, the *form* of neural activity suggests information-carrying *before a complete theoretical understanding*.

# 3. Representation as placeholder



The firing rates of some PFC neurons represent items in working memory to the rest of the brain.

“...an internal representation of relevant information must be created and maintained until it can be used later to guide behavior.”

Experimentalists do not necessarily have a fleshed-out mechanistic scheme in place when they use the term “representation”.

What word should experimentalists use for the hypothetical process a neuron or group of neurons, when:

1. its activity correlates with some information (either sensory or inferred by the experimentalist to be necessary to perform the task)?
2. damage hampers the use of this information?

“The neuron/network \_\_\_\_\_s the information.”

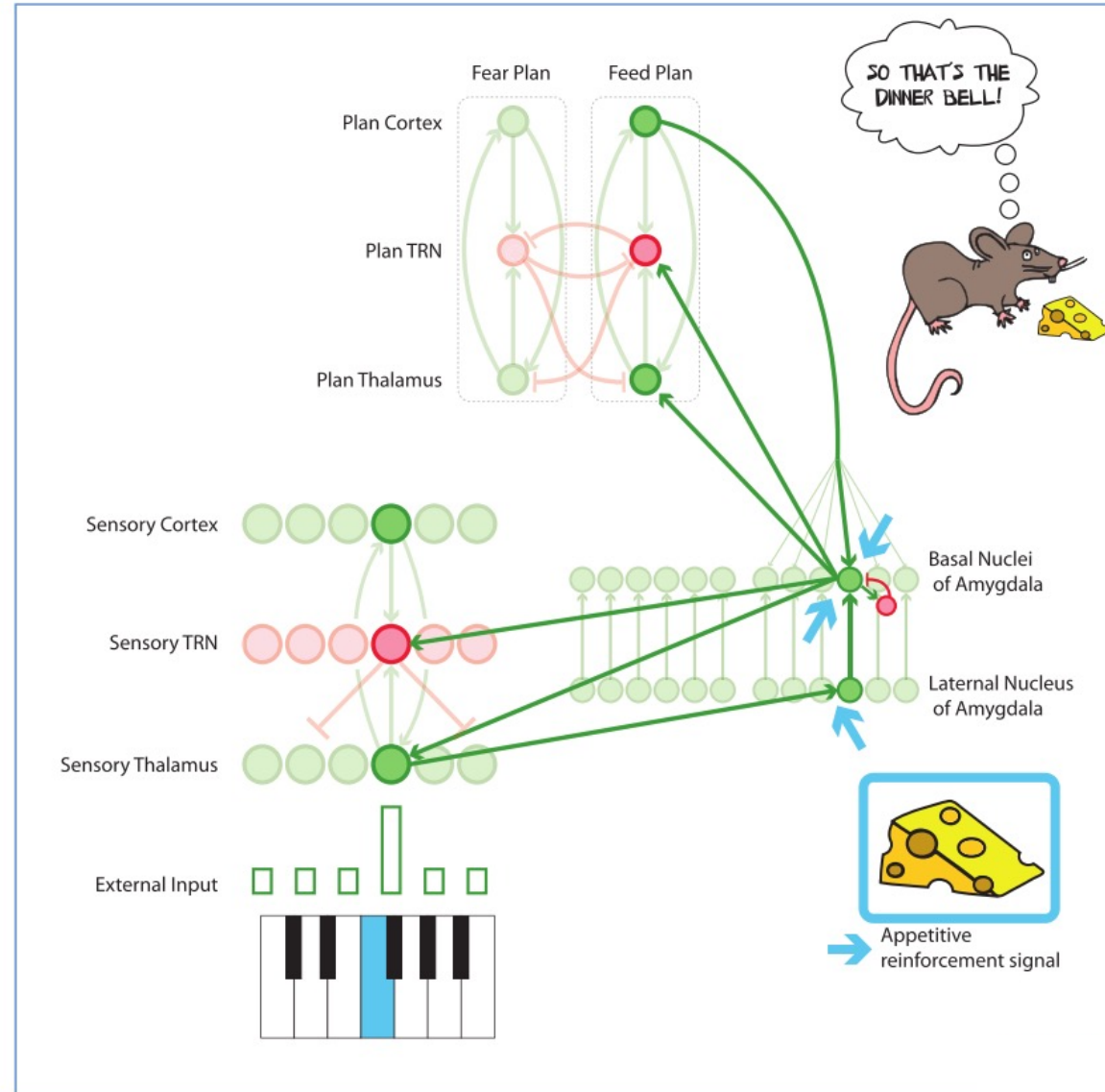
What verb should be used?

# 4. Representation in computational models

In computational models, “representation” is *not a placeholder*, but an evocative label for a set of processes that is fully specified, e.g., by differential equations.

**The amygdala** represents **salience and valence** to **the prefrontal cortex, thalamus, and TRN**.

What other term might be used for these processes?



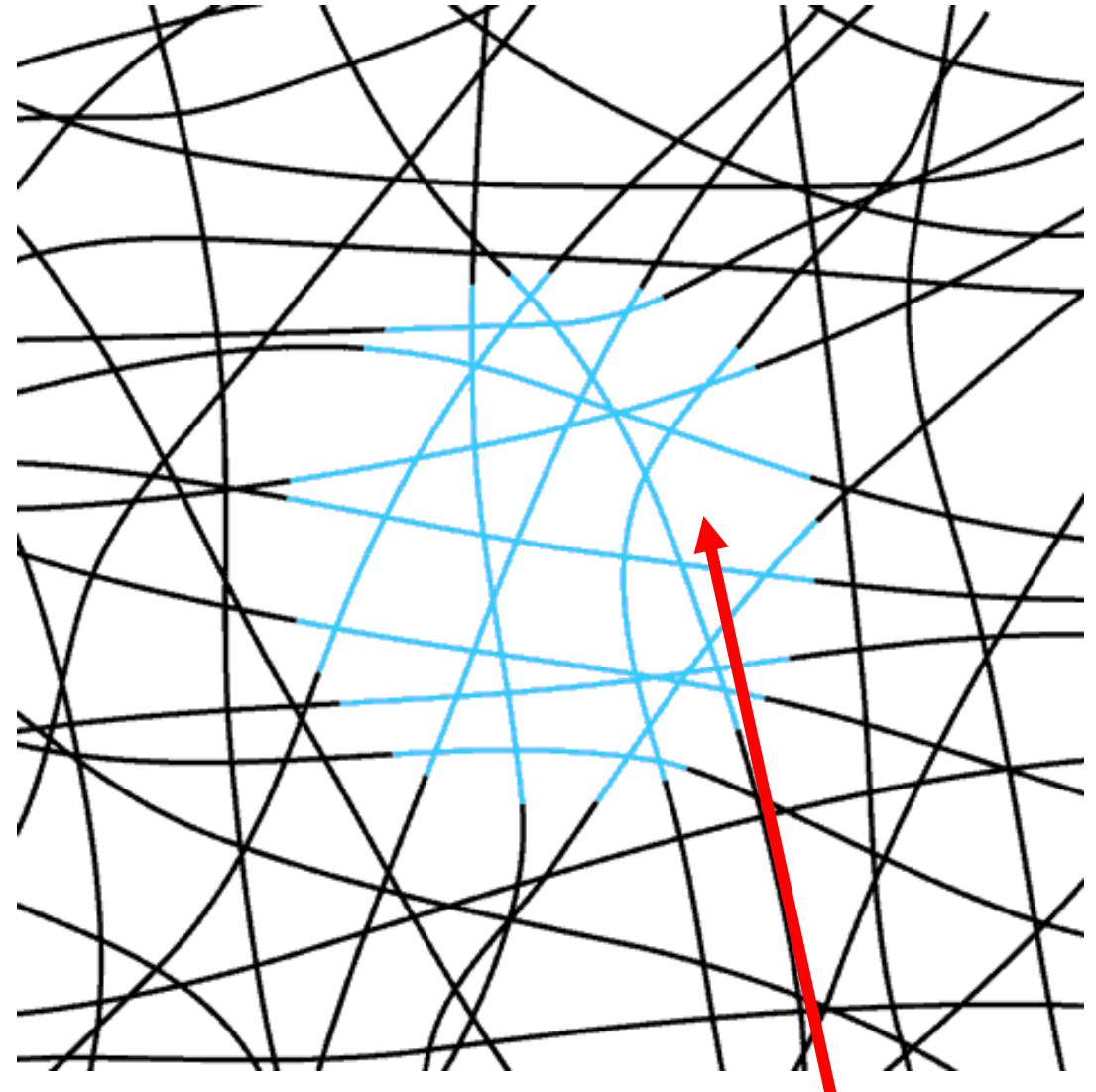
John, Y. J., Zikopoulos, B., Bullock, D., & Barbas, H. (2016). The emotional gatekeeper: a computational model of attentional selection and suppression through the pathway from the amygdala to the inhibitory thalamic reticular nucleus. *PLoS Computational Biology*, 12(2), e1004722.

# Final thoughts

For those of us who regularly describe neural processes using the term “representation”, it is strange to imagine that **memory**, **imagination** and **language** are not representations *in every possible sense* of the word.

Antirep: “Okay, memory, imagination and language might involve representation, but *perception* is ‘direct’!”

Rep: “Given the growing consensus that top-down processes (including memory and unconscious inference) are central to perception, there can be no sharp between perception and other neural/cognitive processes.”



Is **your brain** representing **the color blue** to **you**?



# Thank you!

## Acknowledgements

Neural Systems Lab, BU

Cognitive and Neural Studies  
discussion group

Twitter friends!

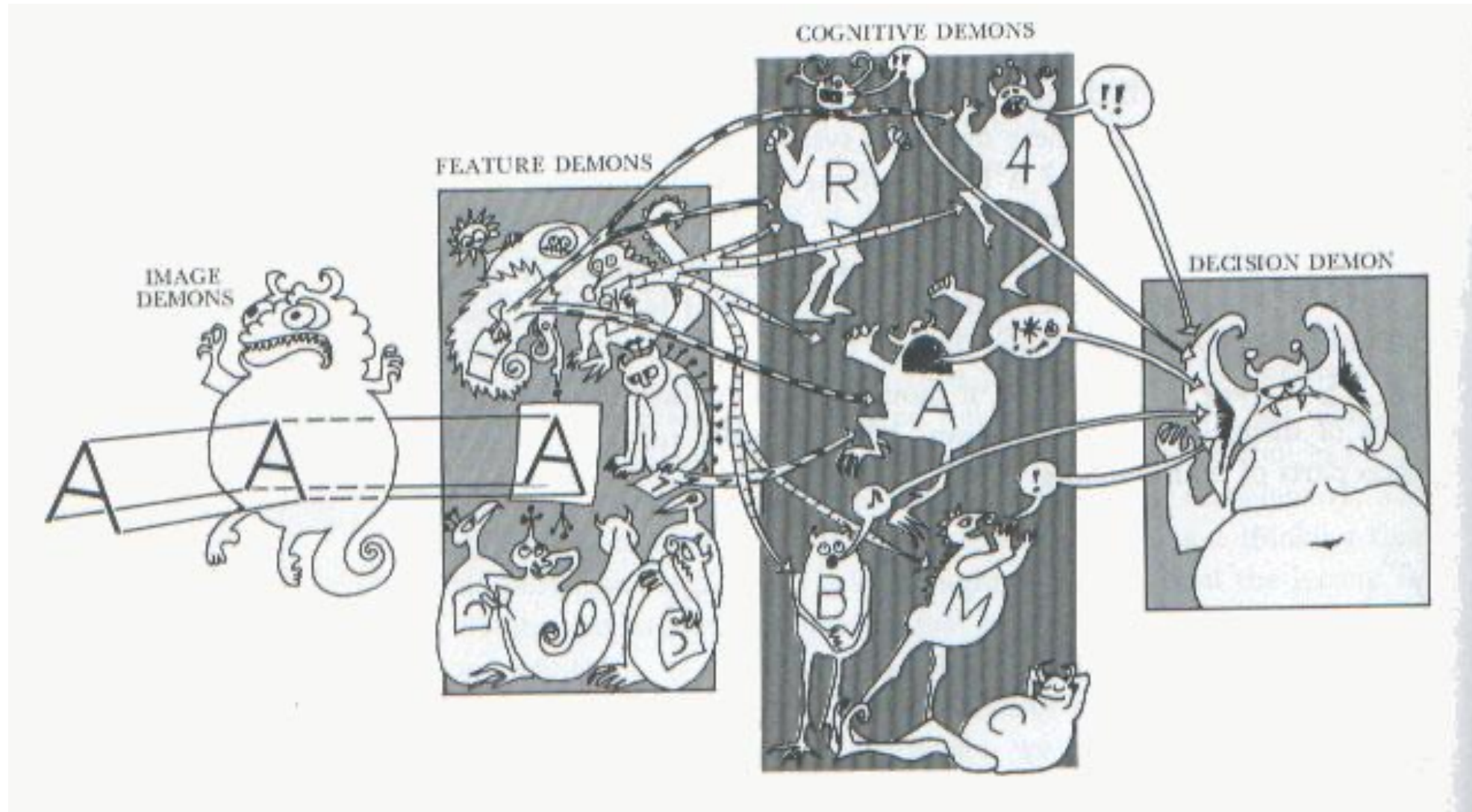


**Chimps** representing **members of parliament** to **us**.

**MPs** representing **their constituents** to ...**each other? the queen?**

**Banksy** representing **democracy** to **us**.

# Representation as (causal) mediation





# An unavoidable tension?

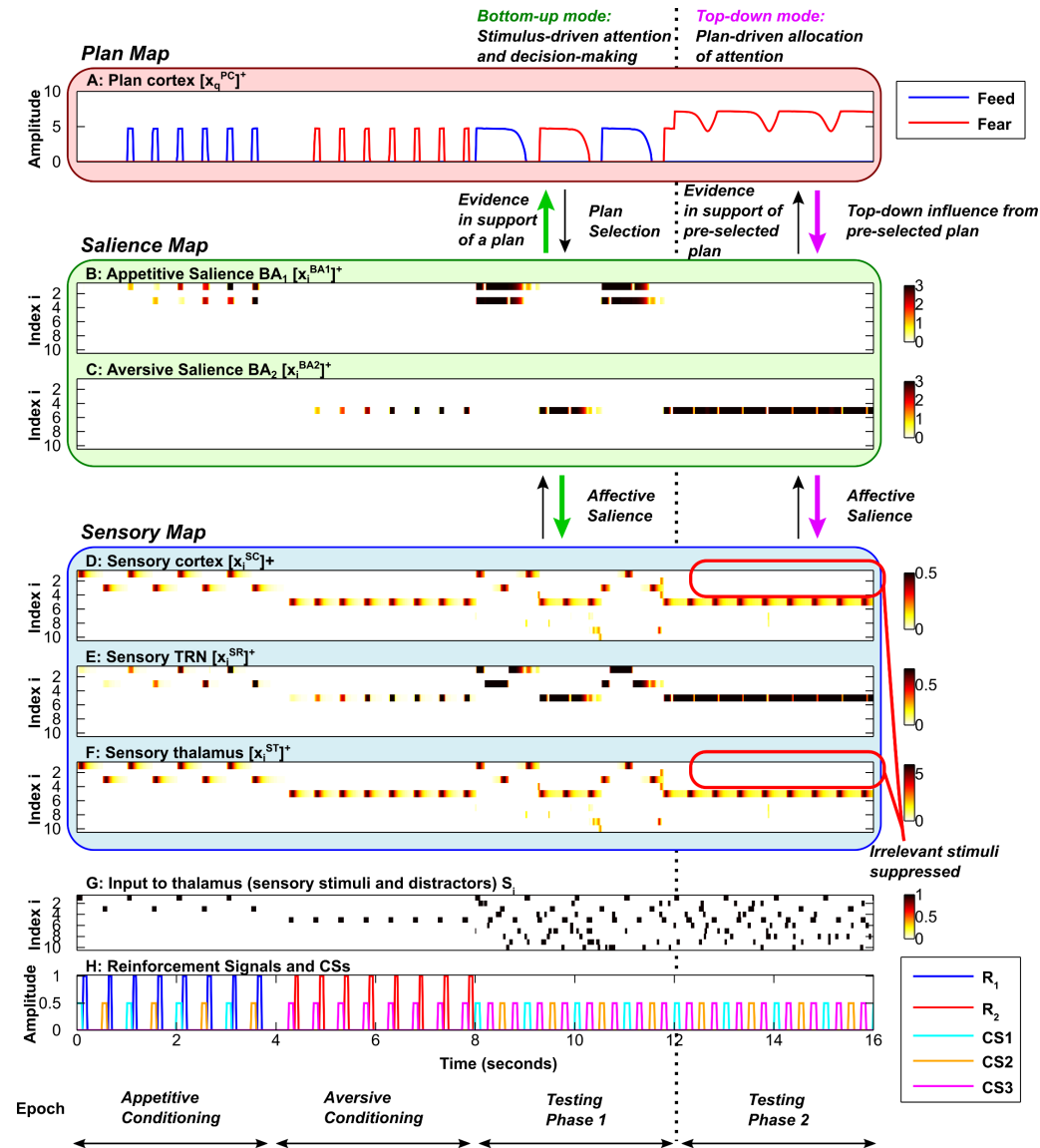
“What can an ‘idea,’ say of a ball, be, except a process representing certain impressions of surface and particular muscular adjustments? What is recollection, but a revivification of such processes which, in the past, have become part of the organism itself?”

Hughlings Jackson 1985/1931, p.26

“Thus there is a tension in Jackson between the idea of **representations as the states of the sensory-motor system** which constitute the elements of reductive explanations of behaviour and mental life, and a holistic picture in which **representations serve to co-ordinate the actions of all parts of the organism**, an “integrative action” (to borrow Sherrington's phrase) which cannot be understood unless the goals of the animal are taken into account.”

This same tension still pervades neuroscience and we cannot *wish it away* prior to a consensus theory of brain/mind function.

# Representation in computational models



John, Y. J., Zikopoulos, B., Bullock, D., & Barbas, H. (2016). The emotional gatekeeper: a computational model of attentional selection and suppression through the pathway from the amygdala to the inhibitory thalamic reticular nucleus. *PLoS computational biology*, 12(2), e1004722.