Conceptual Change in the Brain Revolution

Paul Thagard
University of Waterloo

Outline

1. The brain revolution
2. Concepts
3. Semantic pointers
4. Conceptual change
5. Emotions

Keynes: The difficulty lies, not in the new ideas, but in escaping from the old ones.

The Brain Revolution

1. Neuroscience increasingly influences cognitive, social, developmental, and clinical psychology.
2. Cognitive neuroscience is radically revolutionary for folk psychology, e.g. no soul, and for phenomenological philosophy.
3. Neuroscience is moderately revolutionary for information processing psychology, and for analytic philosophy.

What are Concepts?

Folk psychology: concepts are things in the head, abstract entities.

Information processing psychology: concepts are computational structures.

Neuroscience: concepts are patterns of neural firing that encompass exemplars, typical features, and explanations.

Concepts are Semantic Pointers: Blouw, Solodkin, Thagard, and Eliasmith, in press, *Cognitive Science.*

The New Synthesis

Thesis (1950s): Intelligence results from the processing of physical symbols. (Herbert Simon, traditional AI)

Antithesis (1980s): Intelligence results from sub-symbolic processes in neural networks, operating with distributed representations.

Synthesis: Neural networks are capable of symbolic processes, using semantic pointers.

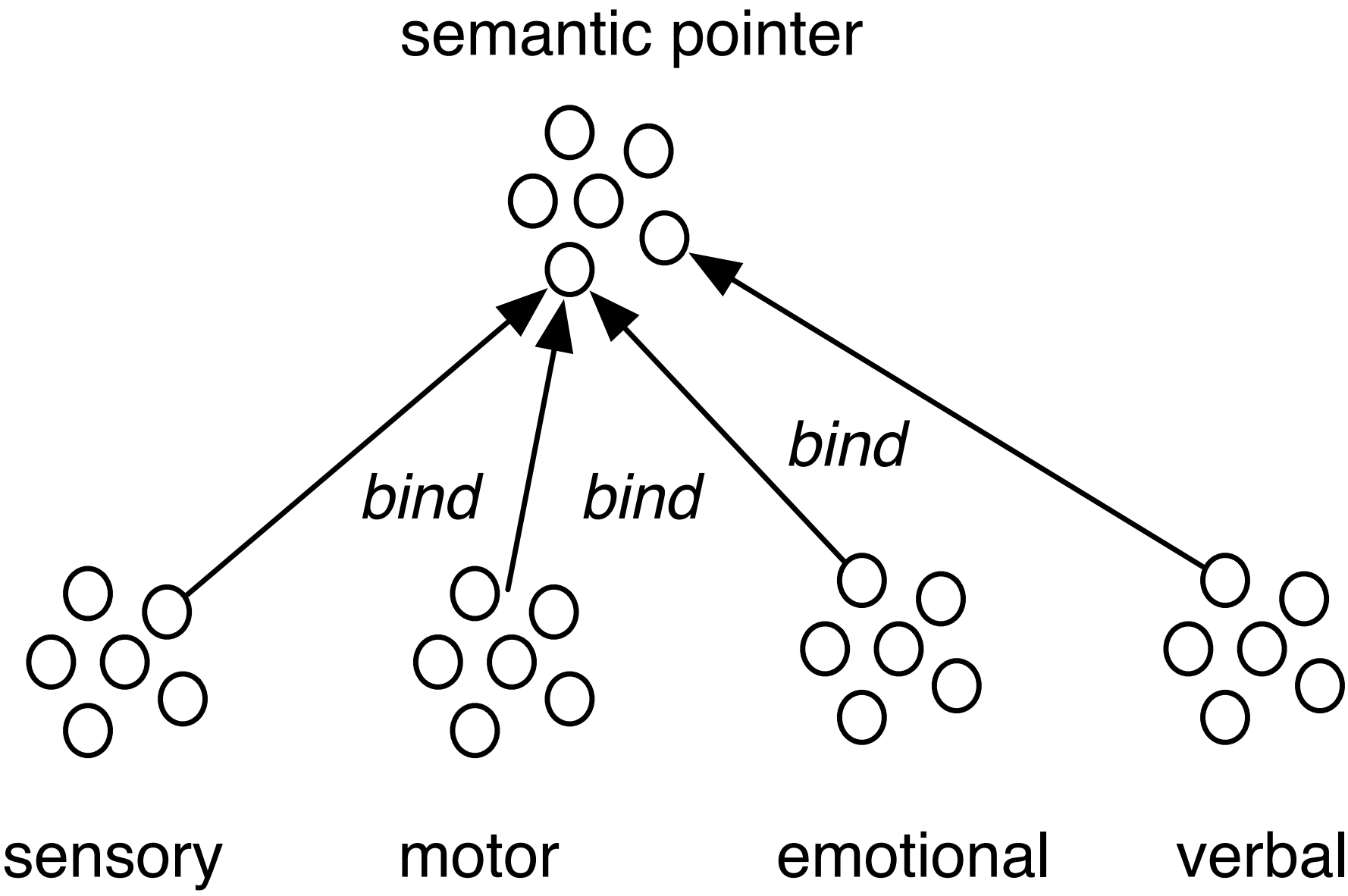
Chris Eliasmith: *How to Build a Brain*, Oxford U. Press, 2013. Eliasmith et al. (2012), *Science*.

Semantic Pointers
(Eliasmith 2013)

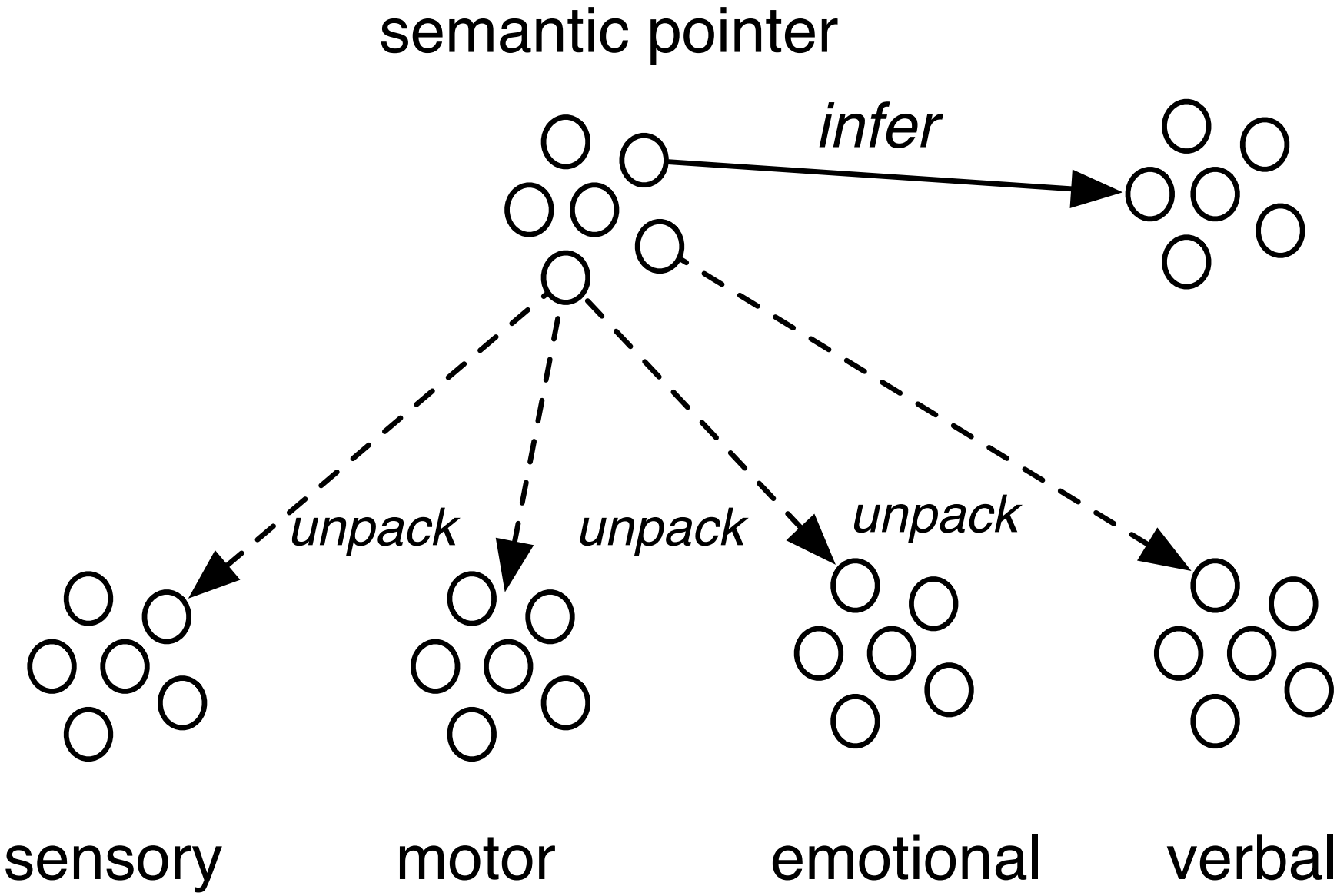
Semantic pointers are patterns of neural firing that:

1. provide *shallow meaning* through symbol-like relations to the world and other representations;
2. expand to provide *deeper meaning* with relations to perceptual, motor, and emotional information;
3. support complex syntactic operations;
4. help to control the flow of information through a cognitive system to accomplish its goals.

FORMATION


semantic pointer
bind
bind
bind
sensory
motor
emotional
verbal

FUNCTION


semantic pointer
infer
unpack
unpack
unpack
sensory
motor
emotional
verbal

What is a Taxonomy? 3-analysis

Exemplars: folk biology, Linnaeus, Wikipedia, computer ontology, DSM, ...

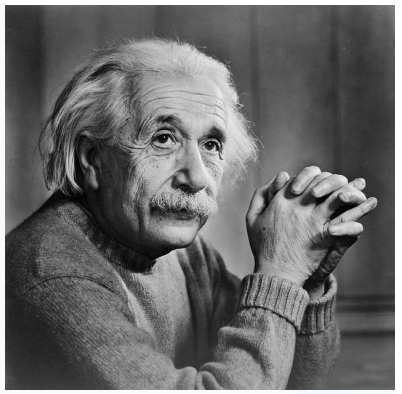
Typical features: domain, concepts, tree structure, classification, ...

Explains: how people classify, infer

Explained by: psychological processes of concept formation, classification

Concepts are Semantic Pointers: Blouw, Solodkin, Thagard, and Eliasmith, in press, *Cognitive Science.*

Kinds of Conceptual Change: A



The acceptance of explanatory identities requires substantial conceptual change.

1. **Addition** of new concepts, e.g. atom.
2. **Deletion** of concepts, e.g. caloric, aether.
3. **Differentiation**, e.g. clouds, infections.
4. **Coalescence**: electricity + magnetism + light, space-time, wave-particle

Thagard (1992) *Conceptual Revolutions.* Thagard (2012) *The Cognitive Science of Science.* Thagard (2014) "Explanatory Identities".

Kinds of Conceptual Change: B

5. Reclassification (branch jumping): transfer of kinds, e.g. earth & air -> mixture, water -> compound, human -> animal, fire & heat & weight -> **process.**

6. Metaclassification (tree switching): whole method of classification changes, e.g.

a. Darwin's classification by historical descent
b. element classification by atomic number
c. disease classification by cause, not symptom.

The Brain Revolution: A

1. **Addition** of new concepts, e.g. neuron, cell assembly, population code, firing pattern, parallel constraint satisfaction, semantic pointer, connectome, etc.
2. **Deletion** of concepts from:
   a. Folk psychology: soul, immortality, free will
   b. Analytic philosophy: propositional attitudes, supervenience, modularity
   c. Information processing psychology: bits, buffer, serial search, executive

The Brain Revolution: B

3. **Differentiation**:
   a. Digital computing vs. neural computing
   b. Representation: local vs. distributed
   c. Change: simple vs. emergent vs. multilevel emergence
   d. Information: Shannon vs. meaningful
   e. Consciousness: simple, self, self-in-society
4. **Coalescence**: cognition-emotion

The Brain Revolution: C

5. **Reclassification** (change of kinds):
   a. Mental representations are processes, not things, e.g. concepts, beliefs, rules, emotions.
   b. Meaning is a multidimensional process.
   c. The self is a system of multilevel mechanisms: Thagard (2014), *Philosophical Psychology.*
   d. Habits are multimodal rules.

The Brain Revolution: D

6. Metaclassification (change of method of classification):

a. Mental processes classified by interactions of brain areas, not folk psychology and localization.

b. Mental illness classified by causes rather than symptoms (NIMH research domain criteria vs. DSM). Should lead to differentiations, coalescences.

Premature Elimination

Examples: self (Hume, Dennett, Metzinger), concepts (Machery), mental representation (Skinner, Chemero) Griffith (emotions)

Retain a theoretical concept when:

1. It contributes to the best explanation of many phenomena.
2. It has underlying mechanisms.

Theories of Emotion

1. Emotion is a property of non-material souls.
2. Emotion is cognitive appraisal (Nussbaum, Ortony, Scherer, etc.).
3. Emotion is physiological perception & embodiment (James, Damasio, Prinz, Niedenthal, etc.).
4. Emotion is social/linguistic construction (Harre, etc.).

5. Emotions are Semantic Pointers

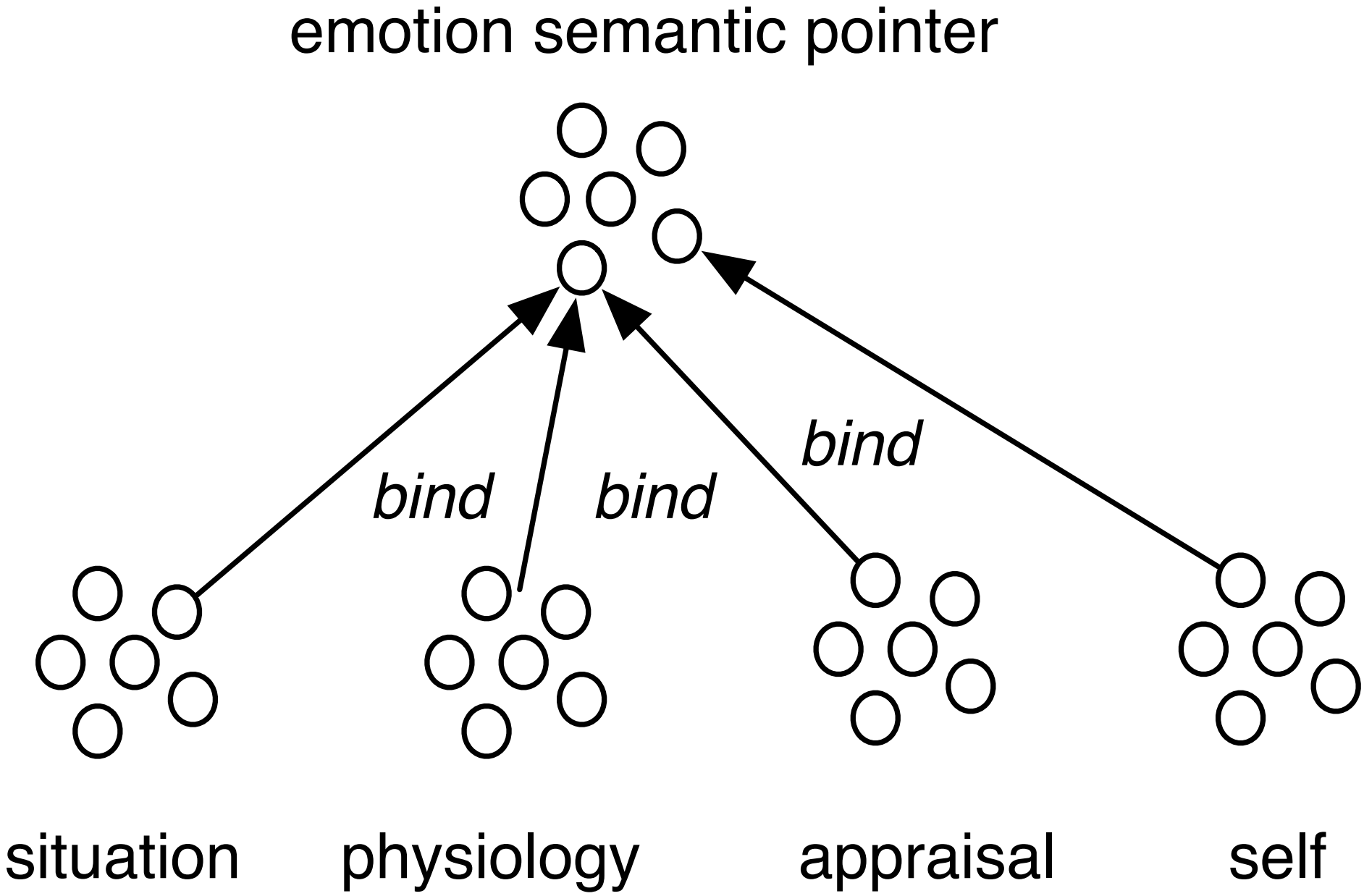
Emotion = bind (concept or belief, cognitive appraisal, physiological perception)

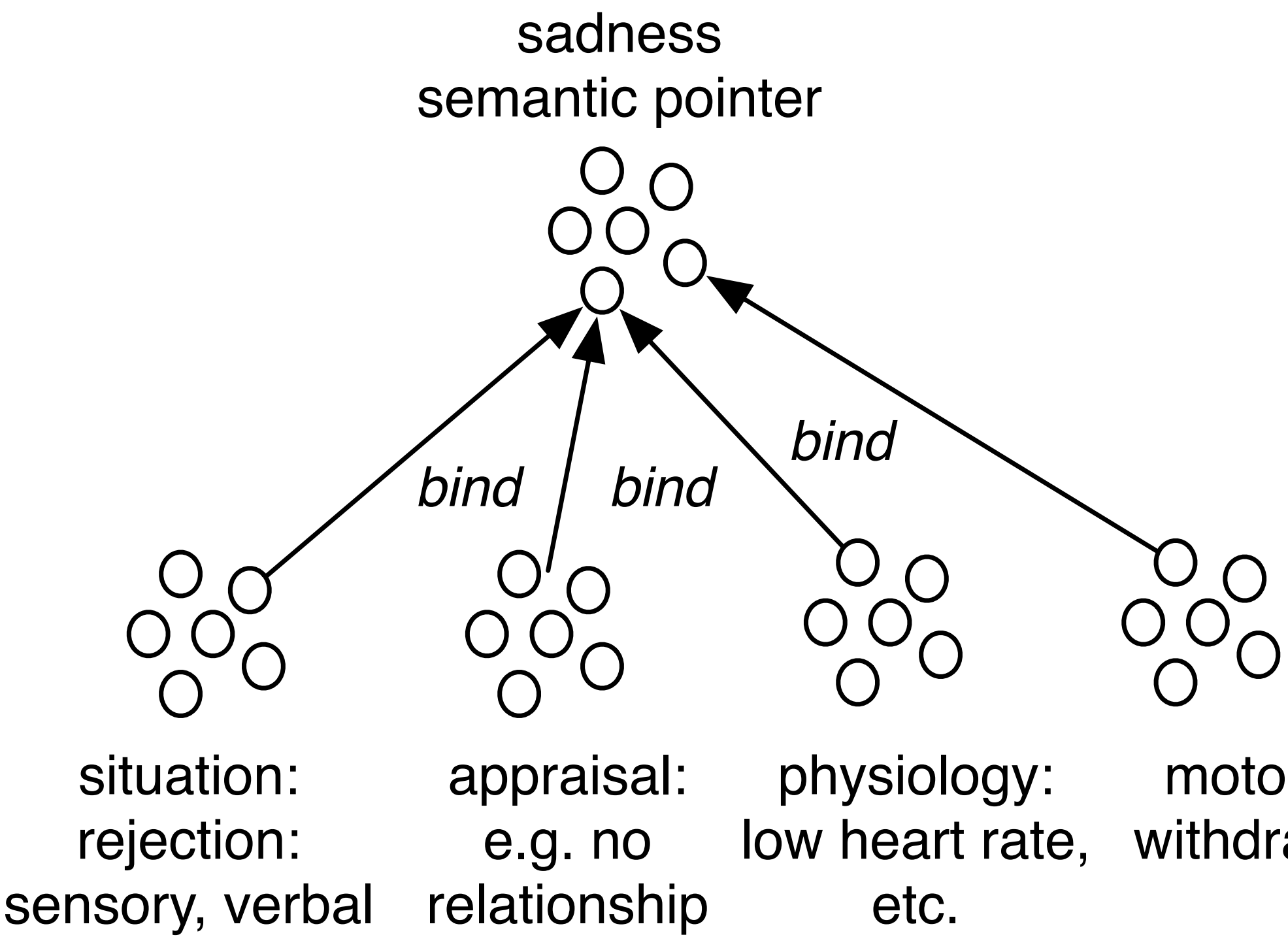
Example: being happy to be in Paris = bind (Paris, appraisal, physiology)

Concepts, beliefs, appraisal, and physiology are all patterns of neural firing.

Binding is by convolution as performed in the Semantic Pointer Architecture.

Thagard and Schröder, "Emotions as Semantic Pointers", 2014.


emotion semantic pointer
bind
bind
bind
situation
physiology
appraisal
self


sadness
semantic pointer
bind
bind
bind
situation:
rejection:
sensory, verbal
appraisal:
e.g. no
relationship
physiology:
low heart rate,
etc.
motor
withdra

Emotions and Conceptual Change: A

1. **Addition** of new concepts: e.g. semantic pointers, nested emotions like fear of humiliation
2. **Deletion** of concepts: propositional attitude, "basic" emotions
3. **Differentiation**: kinds of empathy, kinds of emotions (simple, social)

Emotions and Conceptual Change: B

4. **Coalescence**: cognition-emotion, basic-derivative emotions

5. **Reclassification:** emotions as neural processes, multilevel emergence

6. **Metaclassification**: use neural mechanisms and brain data to generate a new taxonomy of emotions?

Conclusions

1. Brain science alters folk and psychological taxonomies.
2. Fundamental concepts like *concept* and *emotion* require reclassification.
3. Metaclassification is underway with respect to mental processes and disorders.