Theories of Consciousness

NCC, BCC, CCC

Neural correlates of consciousness

What neural activity is present only with awareness?

* Activity in a dedicated set of neurons?
* Special property of neural activity, e.g.,globally coherent states or synchrony?

e.g., binocular rivalry, fusion

Behavioral correlates of consciousness

Allport: memory, potential to act, and subjective confidence

Computational correlates of consciousness

principles of information processing that characterize the differences between conscious and unconscious processing

information processing: the transformation of representations

Classes of Computational Models

Cognitive architectures

general blueprint for understanding information processing

e.g., global workspace model

Specific models

explain data from specific experiments

e.g., attentional blink model

Most models are aimed at “the easy problem”

Can any computational model tackle “the hard problem”?

Major Classes of Theories That Distinguish Conscious and Unconscious States

Privileged role

- Availability to consciousness depends on activity of certain brain systems whose function is to produce subjective experience

Major Classes of Theories That Distinguish Conscious and Unconscious States

Priviliged Role

Meta-representation

When conscious of X (e.g., yellow flower), you can potentially be conscious of being conscious of X, of knowing X, of knowing this is a fact and not a belief, etc.

Availability to consciousness depends on the extent to which a representation is itself an object of representation for further systems of representation.

Major Classes of Theories That Distinguish Conscious and Unconscious States

Priviliged Role

Meta-representation

Quality of representation

Availability to consciousness depends on *quality of representation*

How to define quality?

- stability: temporal duration of representation maintenance
- strength of representation: # neurons or activity of neurons
  e.g., threshold theories and binocular fusion study
- distinctiveness of representation: overlap between other similar reps
- familiarity/experience

Note: involves active representations, as discussed by Maia & Cleeremans

Major Classes of Theories That Distinguish Conscious and Unconscious States

Priviliged Role

Meta-representation

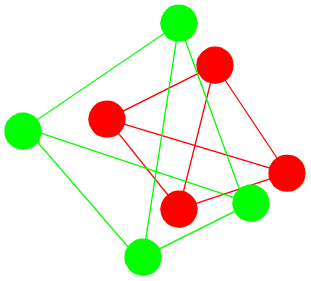
Quality of representation

Integration

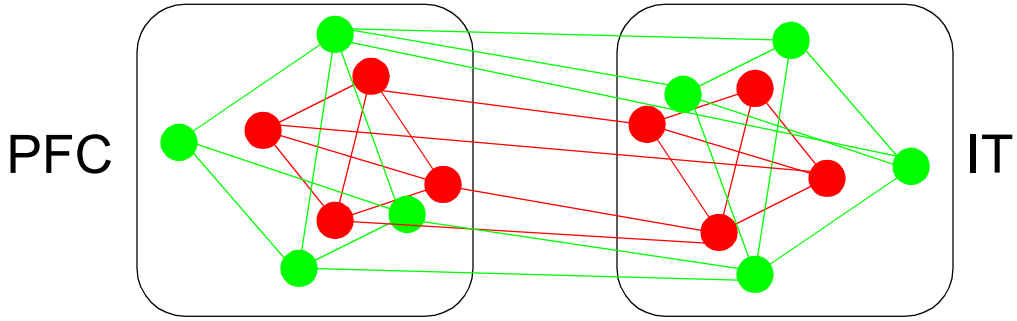
Availability to consciousness depends on processes that integrate (make coherent, synchronize) activity of multiple brain regions.

This view requires some type of competition to choose one interpretation.

"Winning coalition in a large-scale competition determines the contents of consciousness" -> constraint-satisfaction search



within region search



between region search

Major Classes of Theories That Distinguish Conscious and Unconscious States

Priviliged Role

Meta-representation

Quality of representation

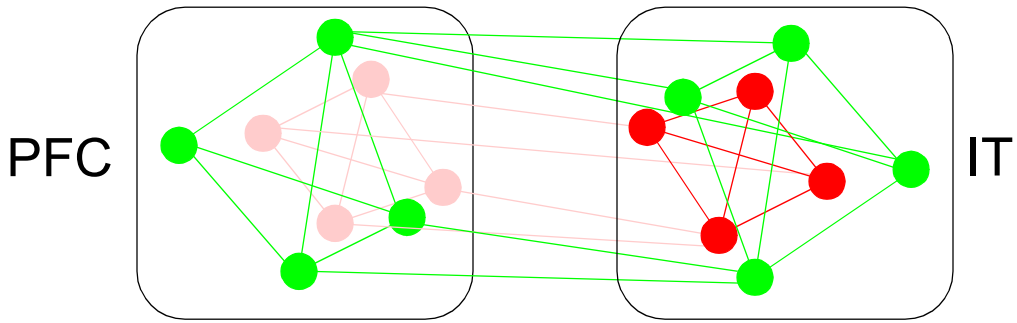
Integration

Availability to consciousness depends on processes that integrate (make coherent, synchronize) activity of multiple brain regions.

This view requires some type of competition to choose one interpretation.

"Winning coalition in a large-scale competition determines the contents of consciousness" -> constraint-satisfaction search

Can also be cast as
biased competition



Major Classes of Theories That Distinguish Conscious and Unconscious States

Priviliged Role

Meta-representation

Quality of representation

Integration

Integration and quality-of-representation views are consistent

achieving integration produces high quality representation

Integration and meta-representation views are consistent

representations created by constraint satisfaction search are fed back as the input to the next search (? according to Maia & Cleeremans)

Executive Control

A.k.a. cognitive control, executive function

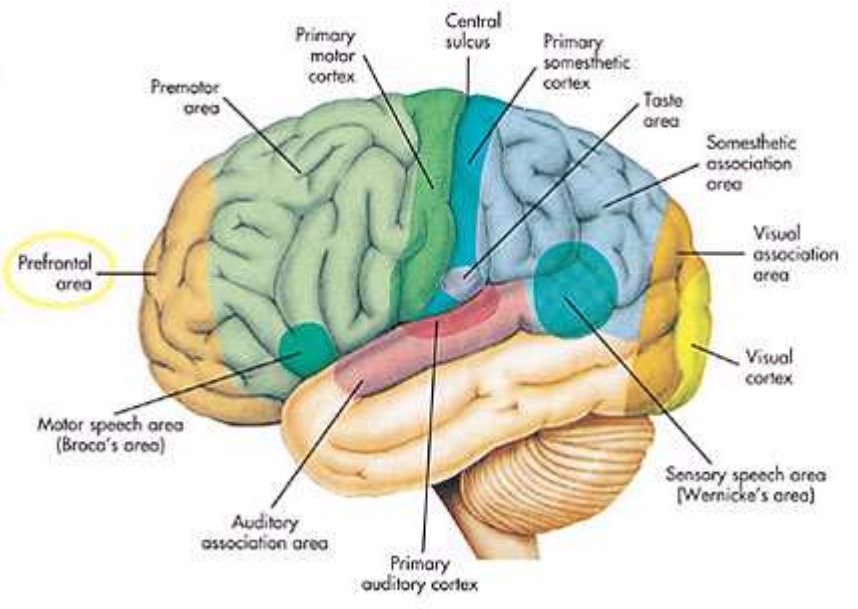
Operations that monitor and regulate ongoing processing in a goal-directed manner

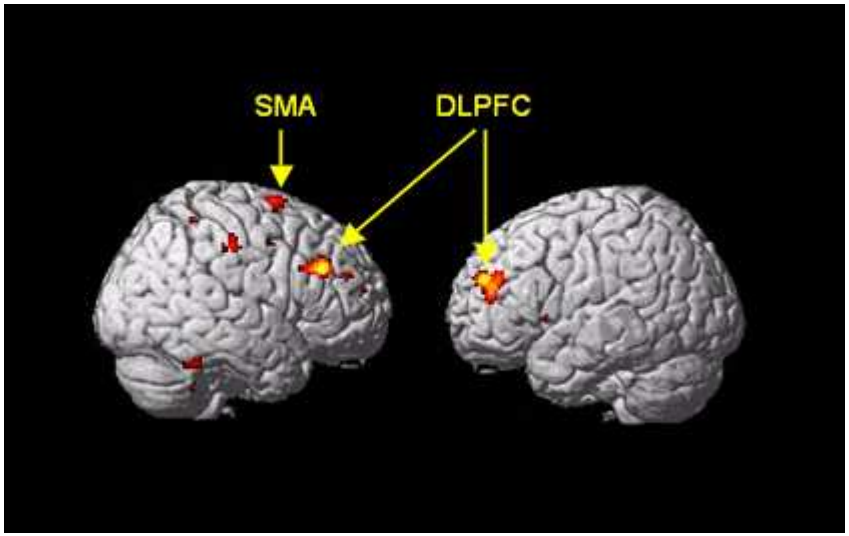
- following arbitrary directions
- processing aspects of the environment that interrupt routine action
  - e.g. ball rolling into street as you drive
- overriding default actions
  - e.g., driving to post office instead of home
- decision points
  - e.g., what to have for lunch
- maintaining information in working memory
  - e.g., phone number

When control is required, we usually become aware of the triggering events, both internal and external.

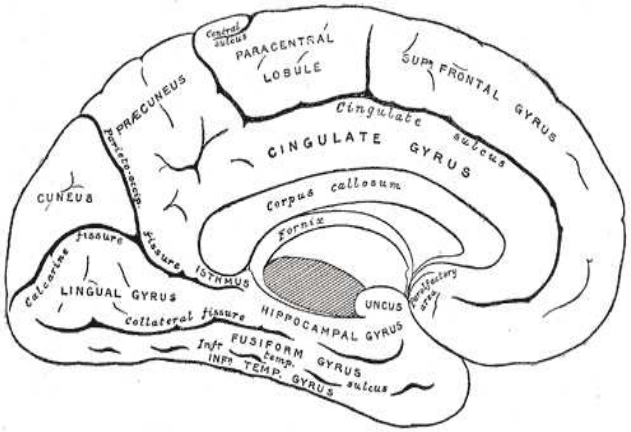
Brain Areas Involved in Executive Control

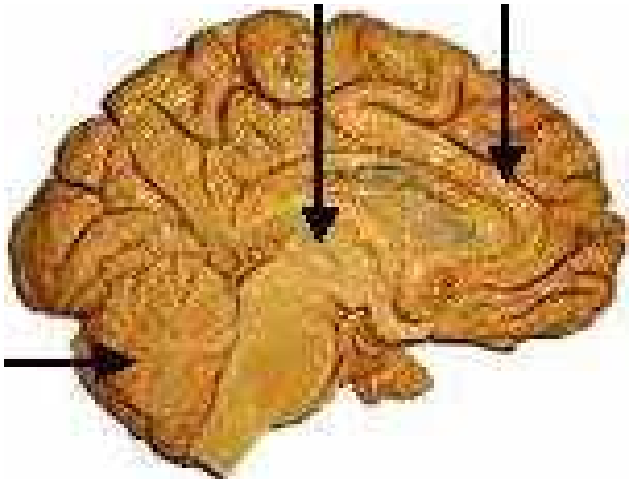
prefrontal cortex (DLPFC)





anterior cingulate cortex (ACC)





Experimental Tasks to Study Executive Control

Stroop task

Name the ink color

ORANGE BLUE GREEN

Overriding default response

Task switching

Add then subtract

9 2

4 1

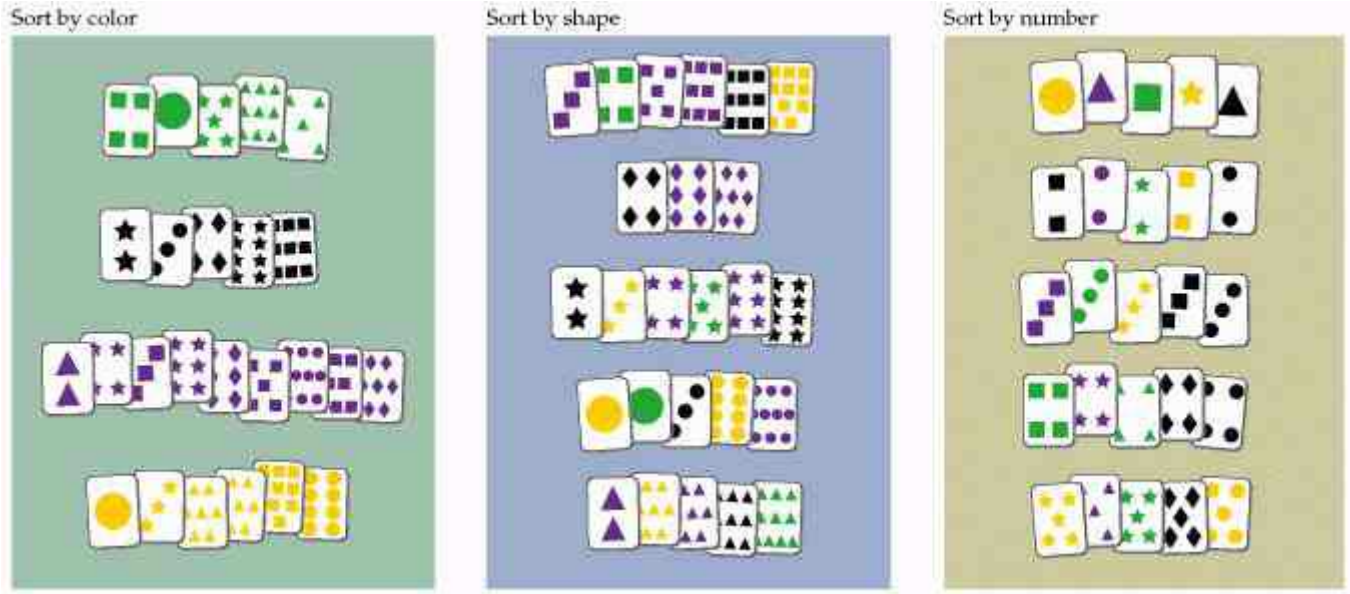
8 7

6 2

2 3

Experimental Tasks to Study Executive Control

Wisconsin Card Sorting Task



N-Back Task

Subjects view a long sequence of letters

For each letter, indicate whether it is a target or nontarget

Targets defined by condition

1-back: letter is a target if it matches the previous letter
2-back: letter is a target if it matches the letter before the previous one

What is the Relationship Between Executive Control and Awareness?

Routine, domain-specific operations do not require awareness

e.g., object recognition, motor control, reading, navigating environment

Executive control operations require awareness

Dehaene et al. (2003)

Task

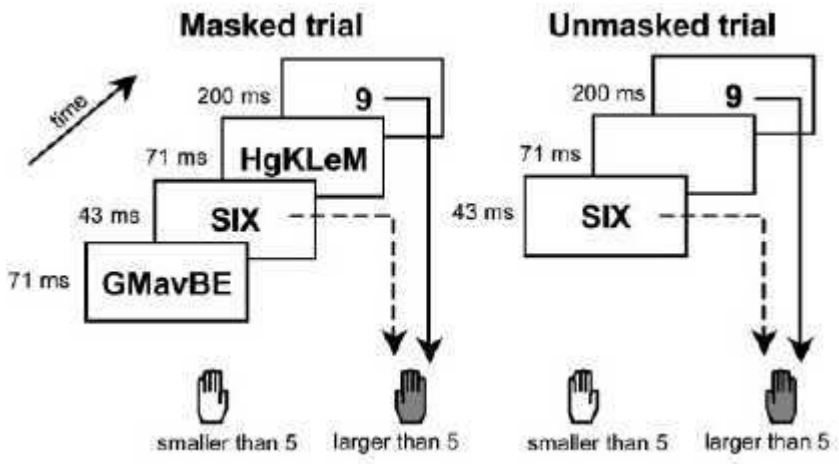
Indicate whether target digit is “less than 5” or “greater than 5”

Prime (spelled digit) precedes target

Prime can be congruent (e.g., “six - 9”) or incongruent (“four - 9”)

Prime can be masked (unconscious) or unmasked (conscious)

Unmasked primes demand cognitive control.



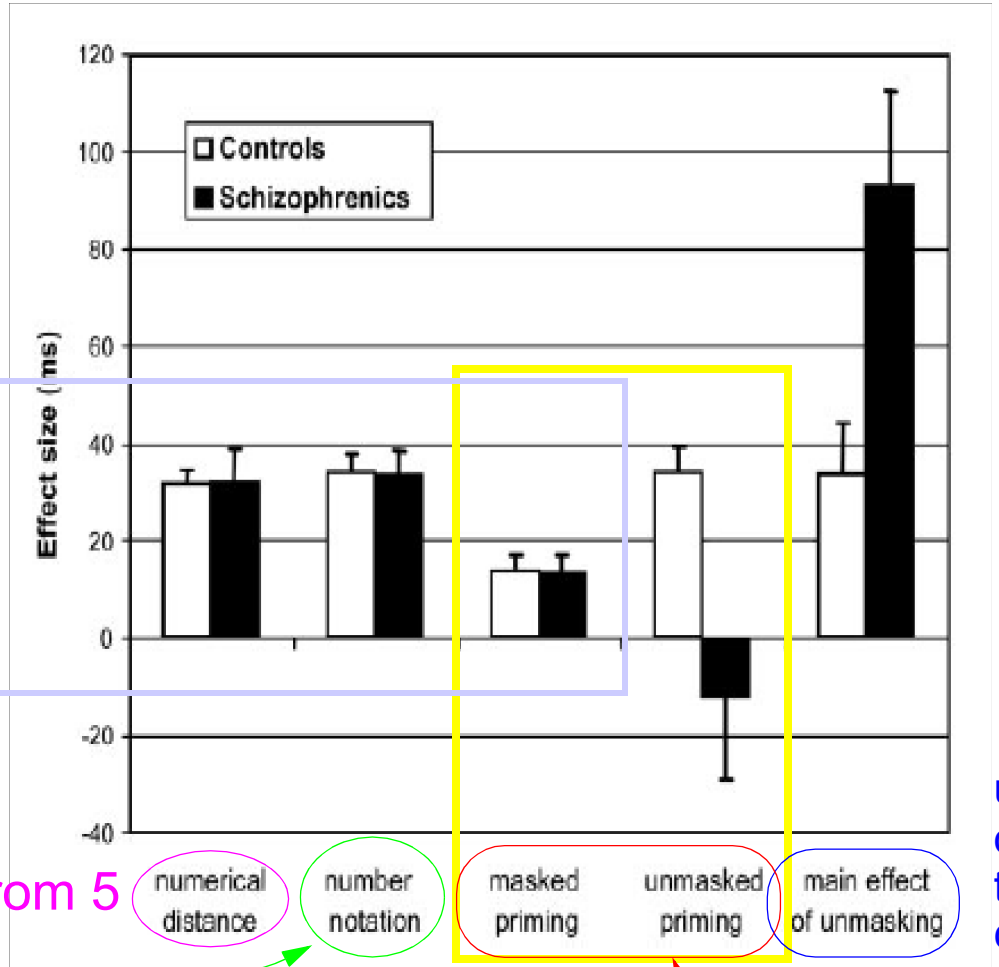
Subjects

Normals

Schizophrenics (“structural and functional abnormalities in ACC and related prefrontal areas”)

Dehaene et al. (2003)

processing of numbers not affected



how far from 5

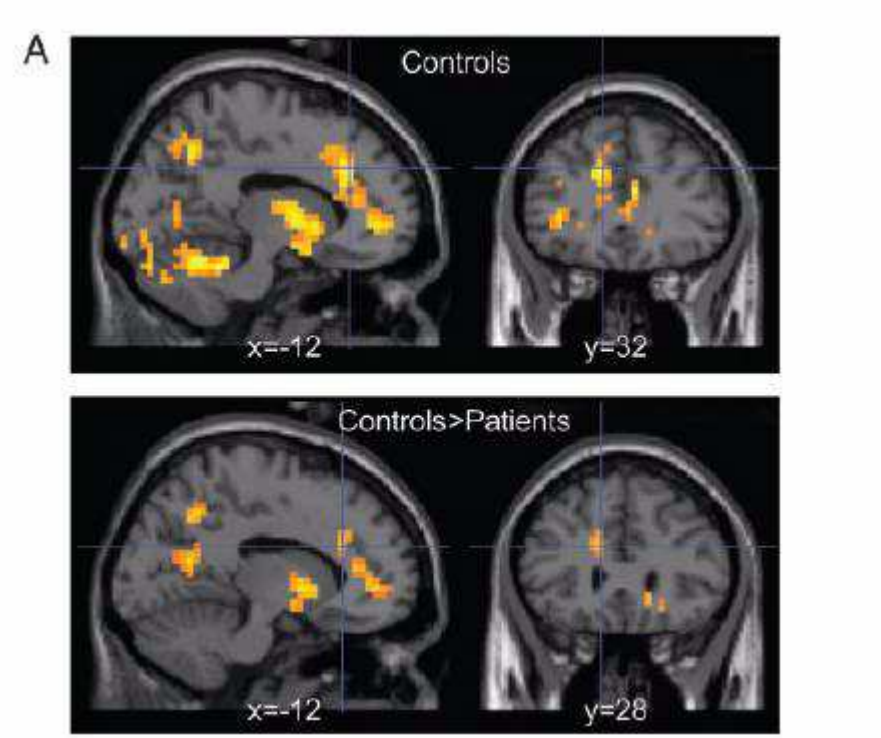
Arabic or spelled

unmasked - masked RT: operation of control to suppress potentially conflicting information

incongruent - congruent RT

Fig. 2. Behavioral performance in number comparison in control subjects and in patients with schizophrenia (mean effect size in milliseconds ± SE). Both groups show identical effects of numerical distance, number notation, and subliminal priming. However, they differed in the unmasked priming effect, which required conscious control of interference. Patients were also severely slowed in the unmasked condition compared with the masked condition.

Dehaene et al. (2003)



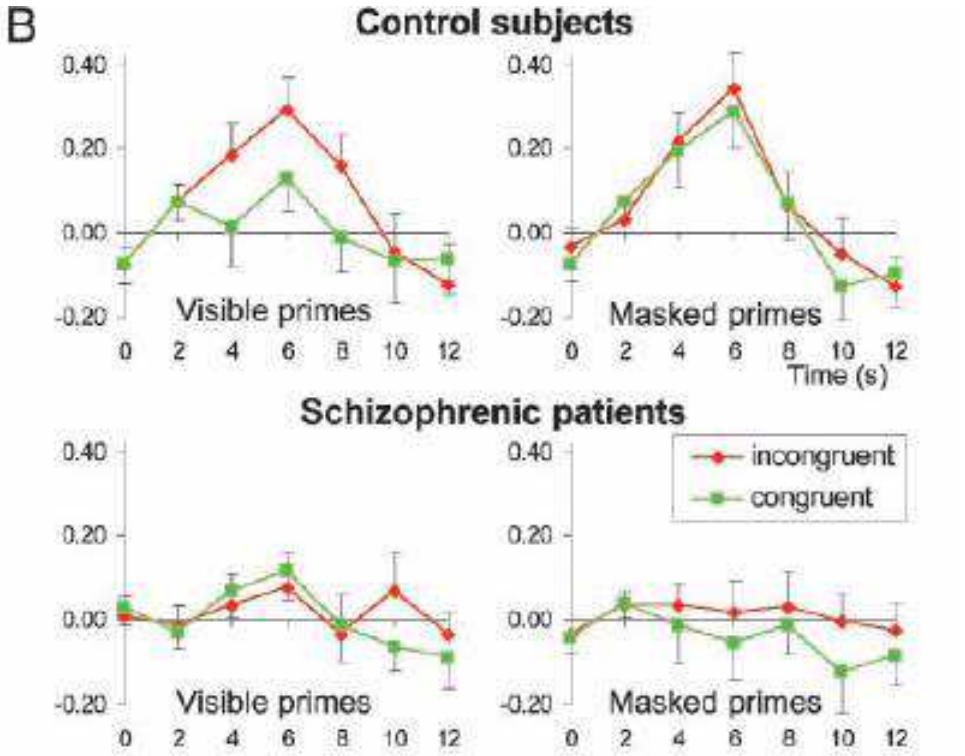


Fig. 4. Effect of conscious conflict in the anterior cingulate in controls and in patients. (*A Upper*) Congruity × visibility interaction in normal subjects showing greater activation in ACC and other brain regions on incongruent trials than on congruent trials, but only when the prime was unmasked. This effect was not found in patients, thus resulting in a triple-interaction group × congruity × visibility (*A Lower*). Curves show the mean percent signal change in the left ACC as a function of time (*B*), revealing a hypoactivation and an absence of conscious conflict effect in the patients.

Conflict regulation by ACC is possible only with awareness.

awareness -> conflict regulation ... or ... conflict regulation -> awareness

Global Workspace Theory

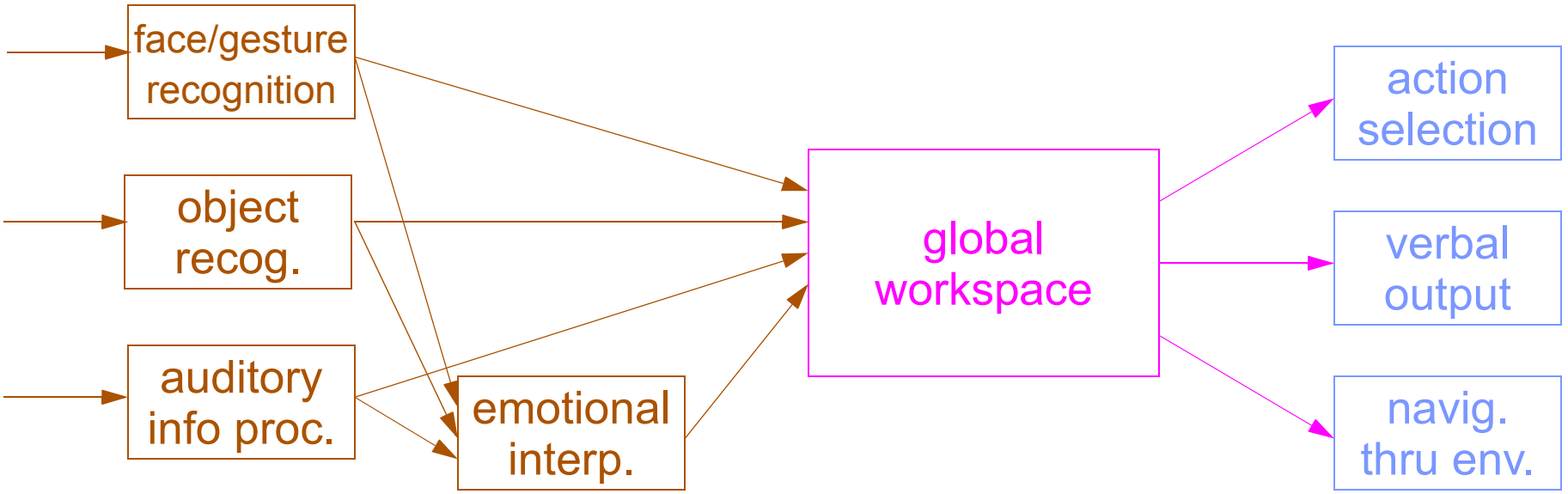
A type of *integration* theory (vs. quality of representation, etc.)

Global workspace facilitates widespread communication between otherwise independent brain functions.

Many specialized perceptual processing systems

Many specialized response processing systems

Global workspace serves to connect them



Related to blackboard models in AI

Global Workspace Theory

Global workspace serves as a means of coordination and control.

Central information exchange

Related to notion of working memory

visual working memory: tracking visual objects, imagery

verbal working memory: inner speech, used for holding on to task-related instructions

Data consistent with global workspace theory

- More activation when stimuli are conscious than unconscious (e.g., binocular fusion study). [consistent with many theories]
- Activation of frontal-parietal circuits when stimuli are conscious (binocular rivalry studies)
- More powerful learning mechanisms come into play when information is conscious (implicit/explicit learning contrast)
- Executive control requires awareness.

Global Workspace Theory

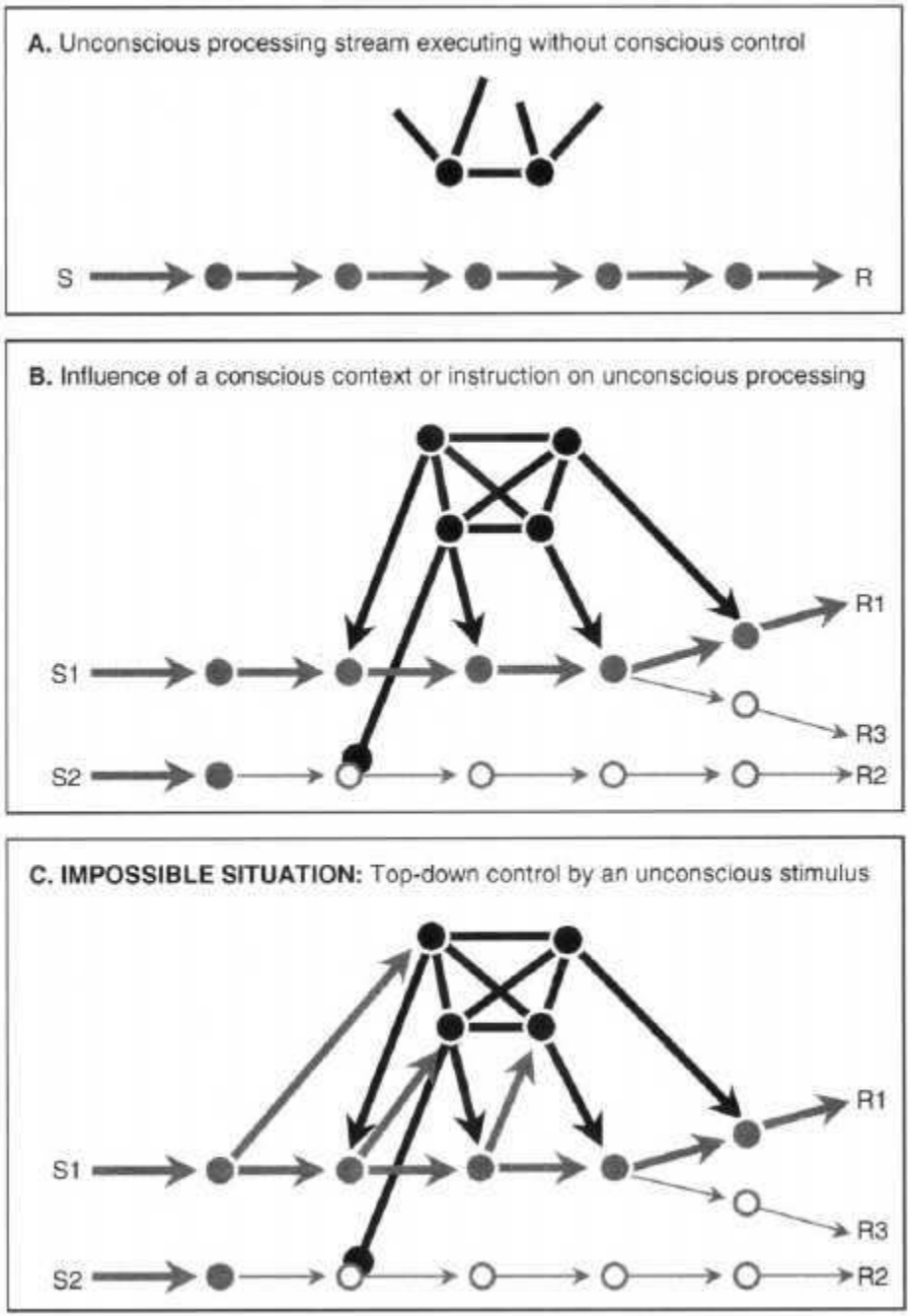
Criticisms

- Explains how access can take place, but not how content is selected
- More of a framework than a theory
- Many other views are consistent with the above data
- Allows for many different neural implementations: some variants allow distributed neural representation of workspace, others place workspace in frontal regions

Global Workspace Model (Dehaene et al.)

A state of *representational coherence* within a global workspace gives rise to both consciousness and cognitive control.

ACC and PFC are neural substrates of global workspace.



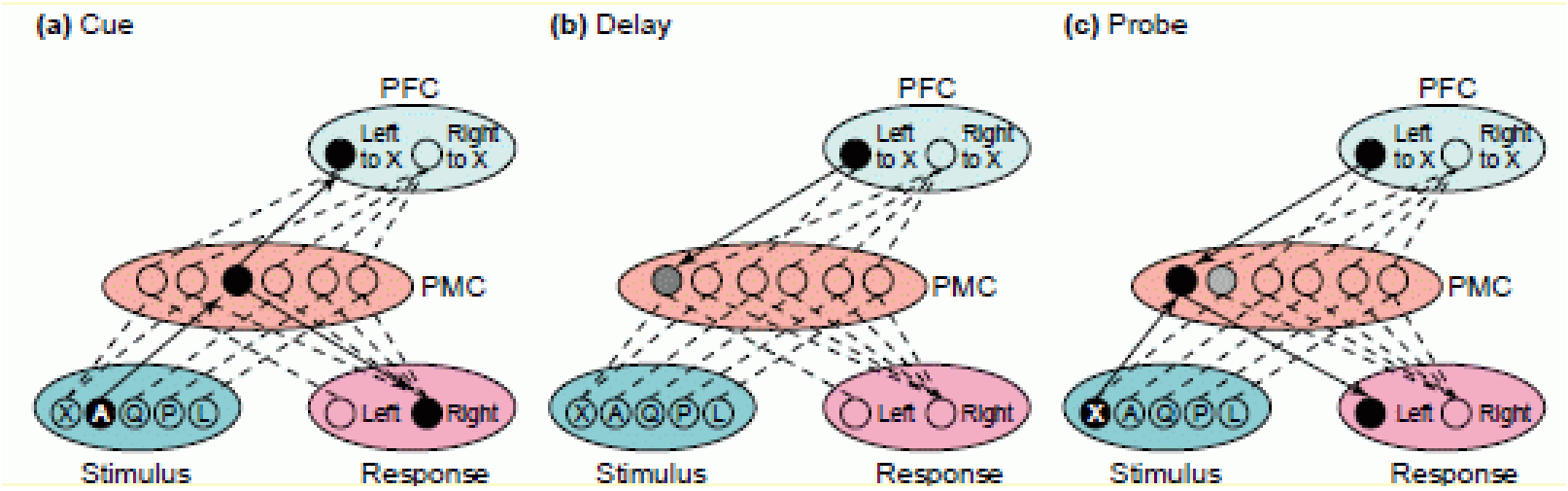
Global Workspace Model (Braver, Cohen, O'Reilly)

AX-CPT

push left button if A then X; right button otherwise

A Q P X A L A X L Q Q X A X

Connectionist model



Can be interpreted as a global workspace model

PFC serves as workspace

Can be interpreted as biased competition model

An Information Flow Perspective on Access Consciousness

Michael C. Mozer
Donald W. Mathis
Jennifer Finger
Michael Colagrosso

Institute of Cognitive Science and
Department of Computer Science
University of Colorado, Boulder

Cognitive Science
Fall 2008

Cognitive Architecture

Cortical computation can be characterized by coarse-scale, functionally specialized pathways.

E.g., visual word-form recognition

identification of semantic features of visual objects

auditory word recognition

computation of spatial relationships

construction of motor plans

Cognition requires coordination of multiple pathways.

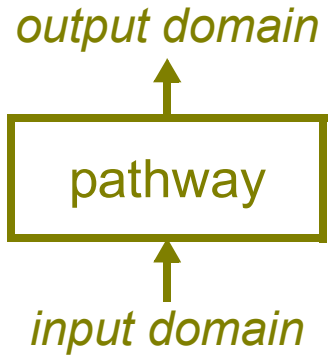
Pathways act as associative memories.

E.g., visual word recognition

input: visual features
output: letter strings

Past experience imposes well-formedness constraints on the output.

Letter strings must be consistent with English orthography.
Output is the best-fitting interpretation of the input.

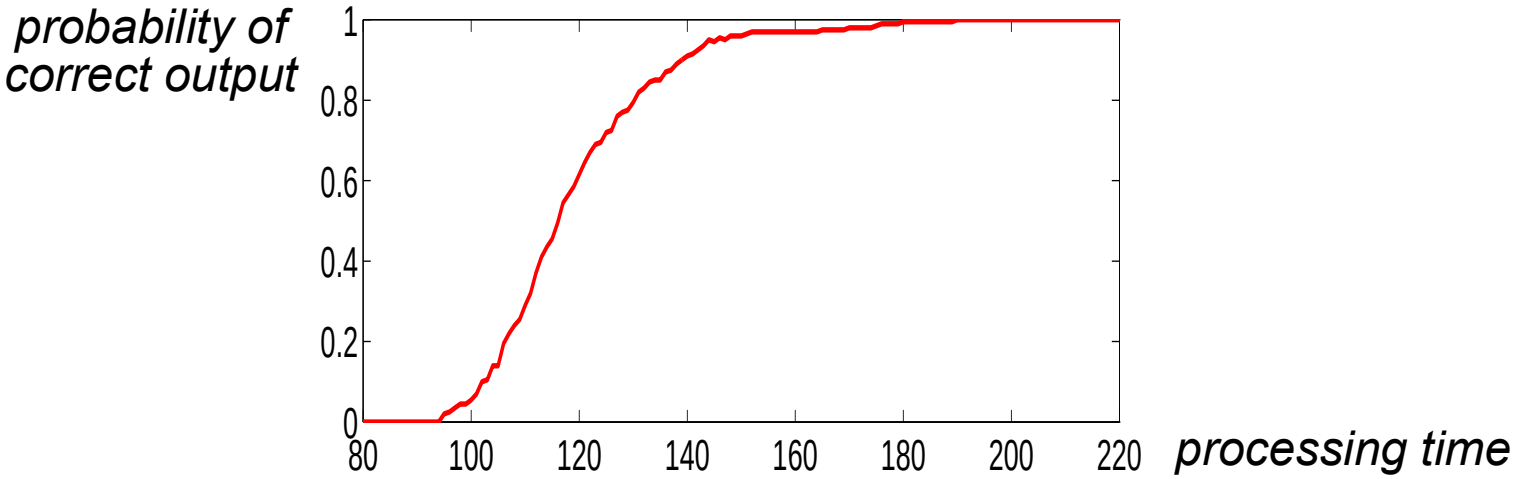


Cognitive Architecture

Pathway operation shows a speed-accuracy trade off.

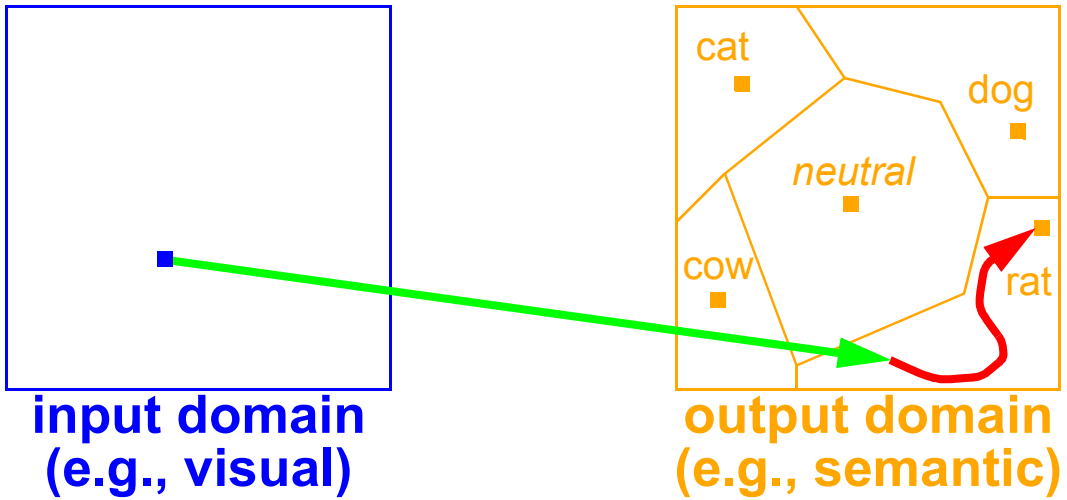
Initial output appears rapidly, but is inaccurate.

Pathway asymptotically converges on the best interpretation.

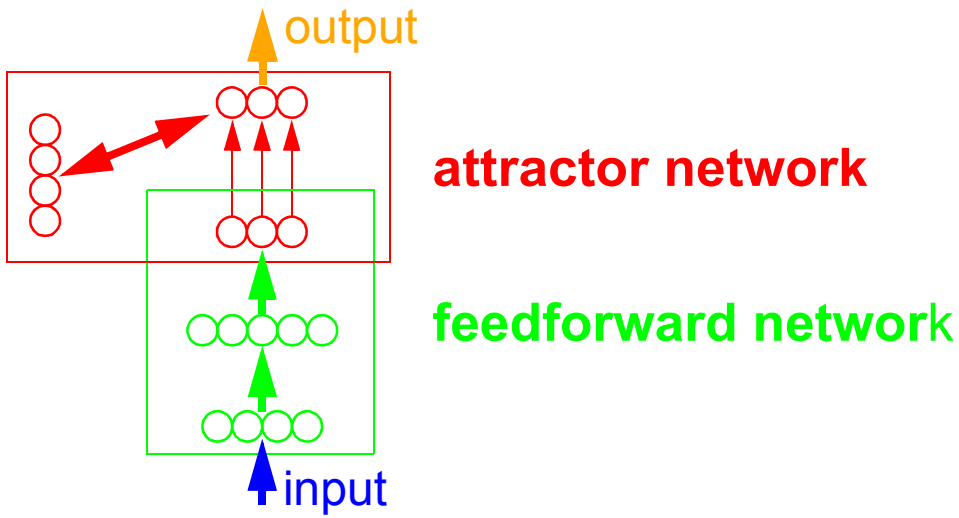


With each experience, pathway tends to produce its response more rapidly.

Pathway Operation



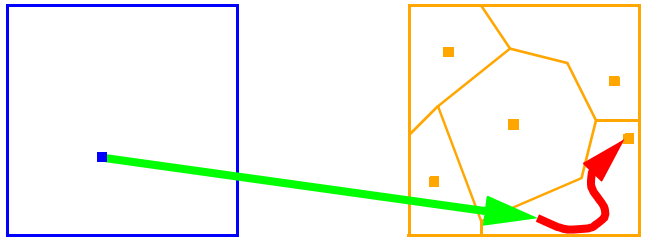
Implementation using two standard connectionist components:



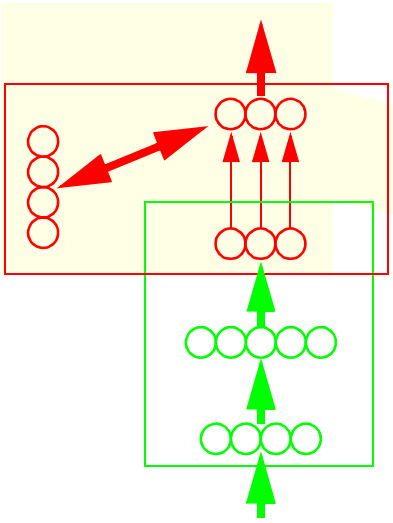
Pathway Implementation

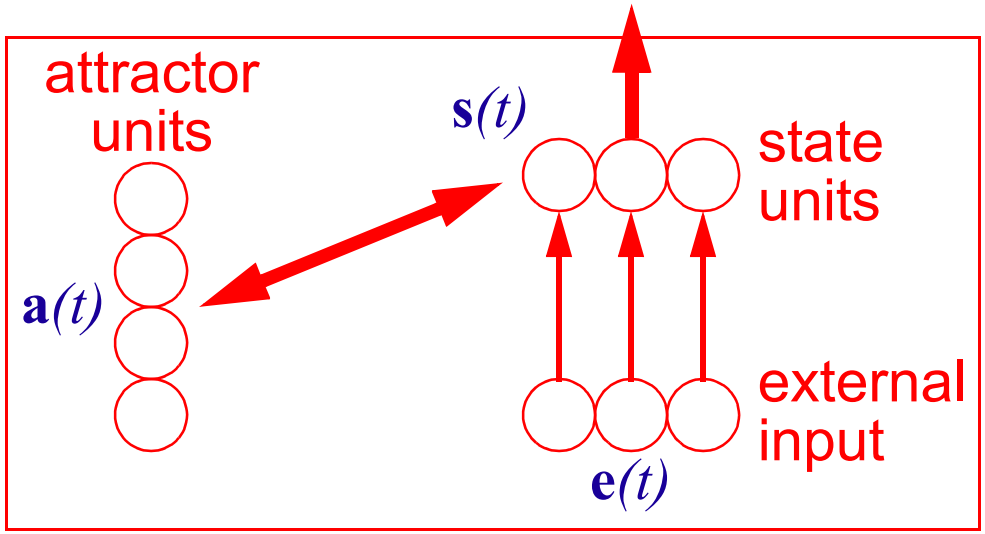
Speed-accuracy trade off due to

- inaccuracies in feedforward mapping
- noise in unit activities



Pathway Dynamics





attractor unit update equation:

$$\hat{a}_j(t) = \exp(-\|\mathbf{s}(t) - \mu_j\|^2 / \beta_j)$$

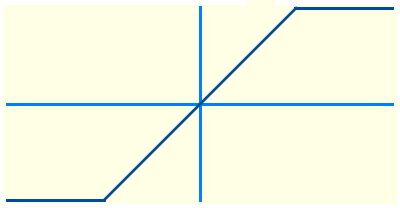
$$a_j(t) = \hat{a}_j(t) / \sum \hat{a}_i(t)$$

μ_j: center of attractor *j*

β_j: width of attractor *j*

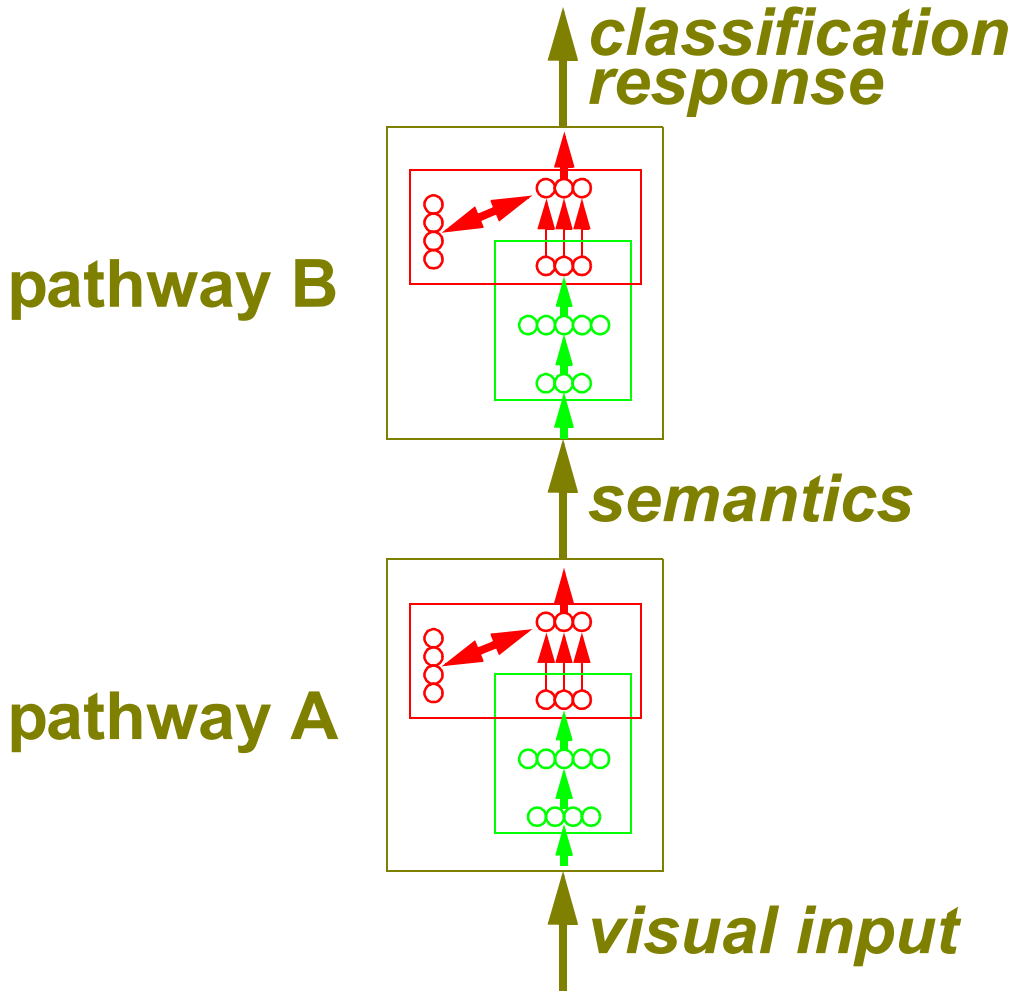
state unit update equation:

$$s_i(t) = h\left[s_i(t-1) + \omega_{ext} e_i(t) + \omega_{attr} \sum_j a_j(t-1)\mu_{ji} + \eta_i(t;\sigma) \right]$$



free
parameters

Cascaded Pathways

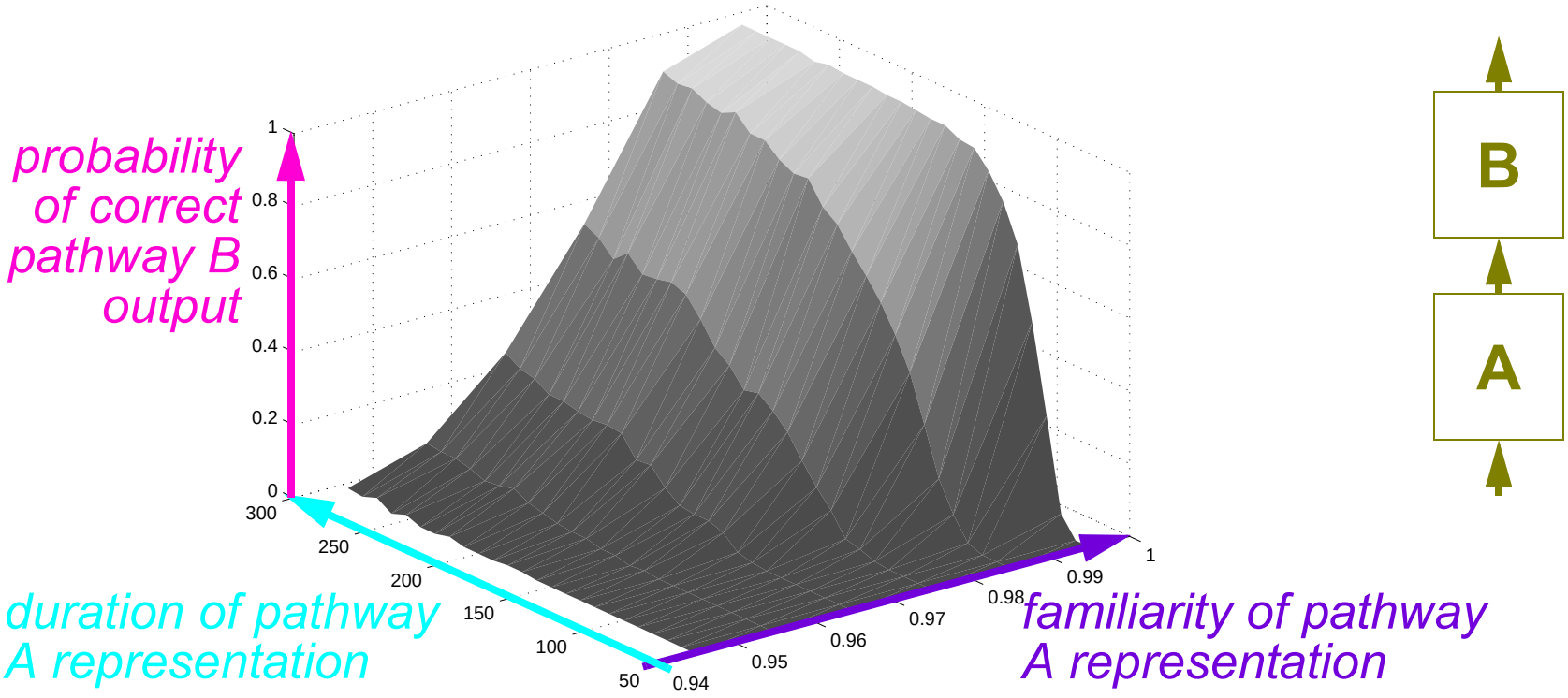


Sequential operation?

- biologically implausible
- computationally inefficient

Information Flow

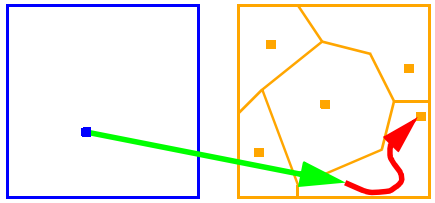
What sort of representation in the output of pathway A will support a response from pathway B?



Sufficient condition: familiar and temporally persistent representation

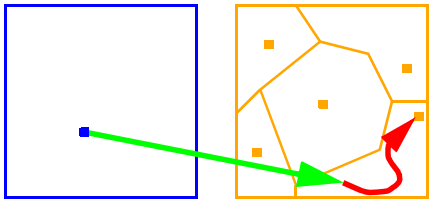
Information Flow

Familiar and persistent representation attained when pathway reaches attractor



Information Flow

Familiar and persistent representation attained when pathway reaches attractor

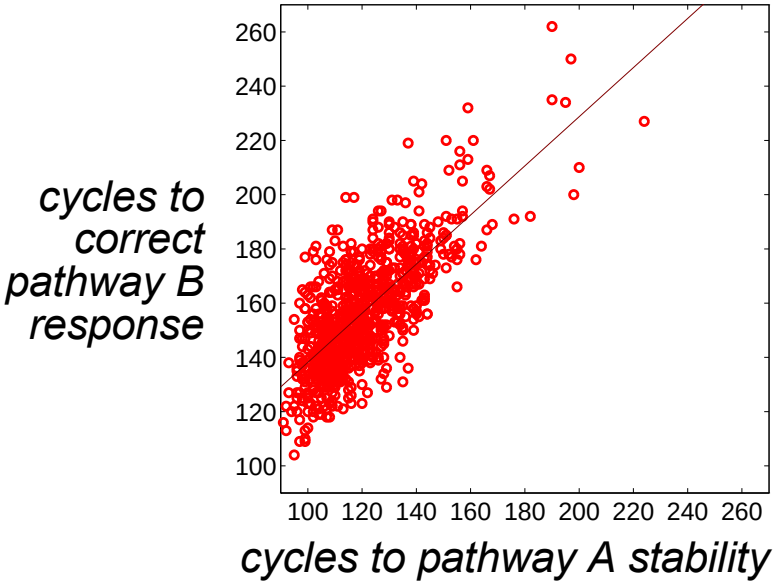


At attractor, representation is *stable*.

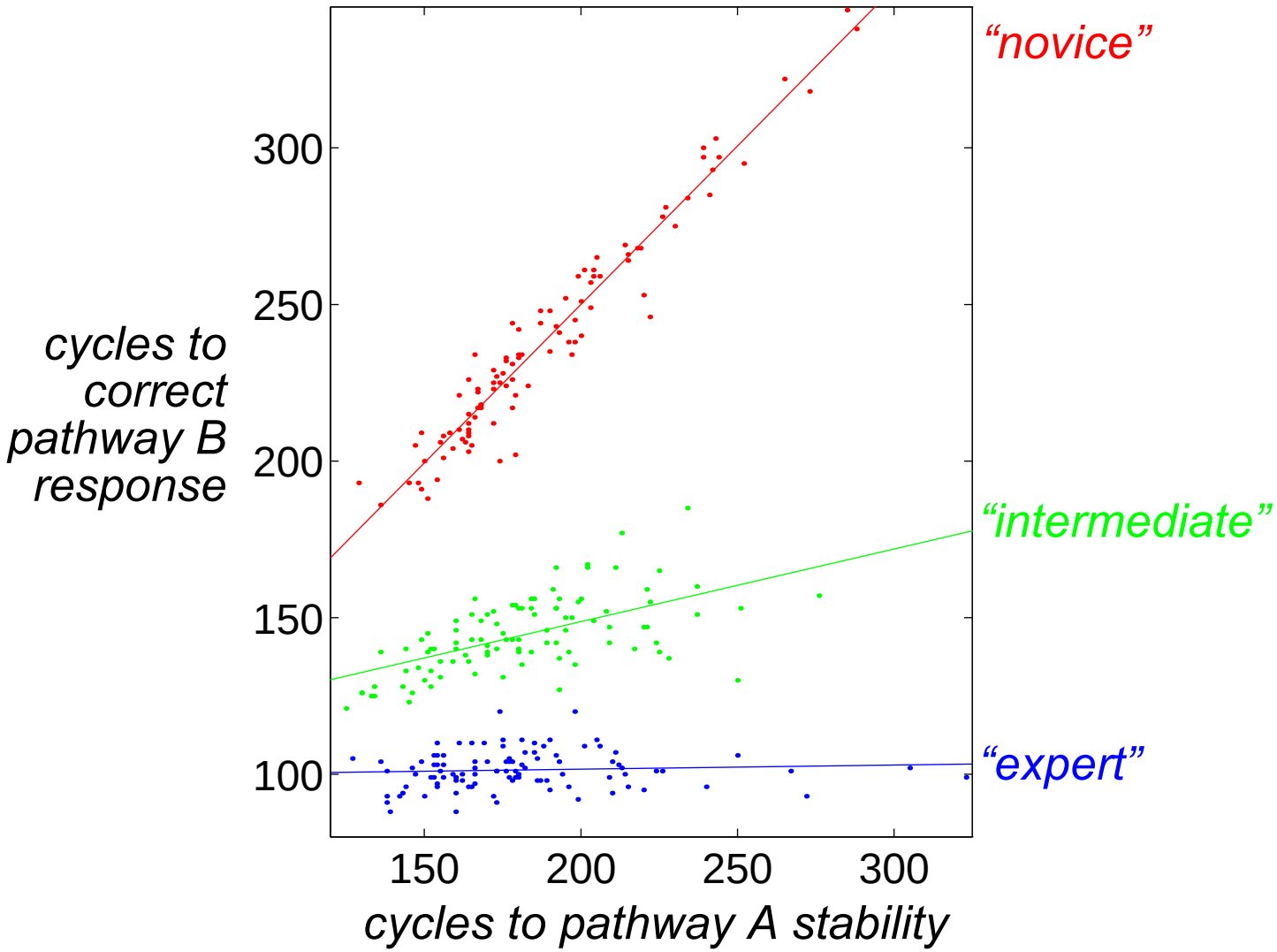
Stability defined to occur if $\bar{d}(t) < \zeta$,
where $\bar{d}(t) = \alpha \bar{d}(t-1) + (1-\alpha) d(t)$,

$d(t) = \|\mathbf{s}(t) - \mathbf{s}(t-1)\|$,

$\zeta = 0.25$, **and** $\alpha = 0.9$

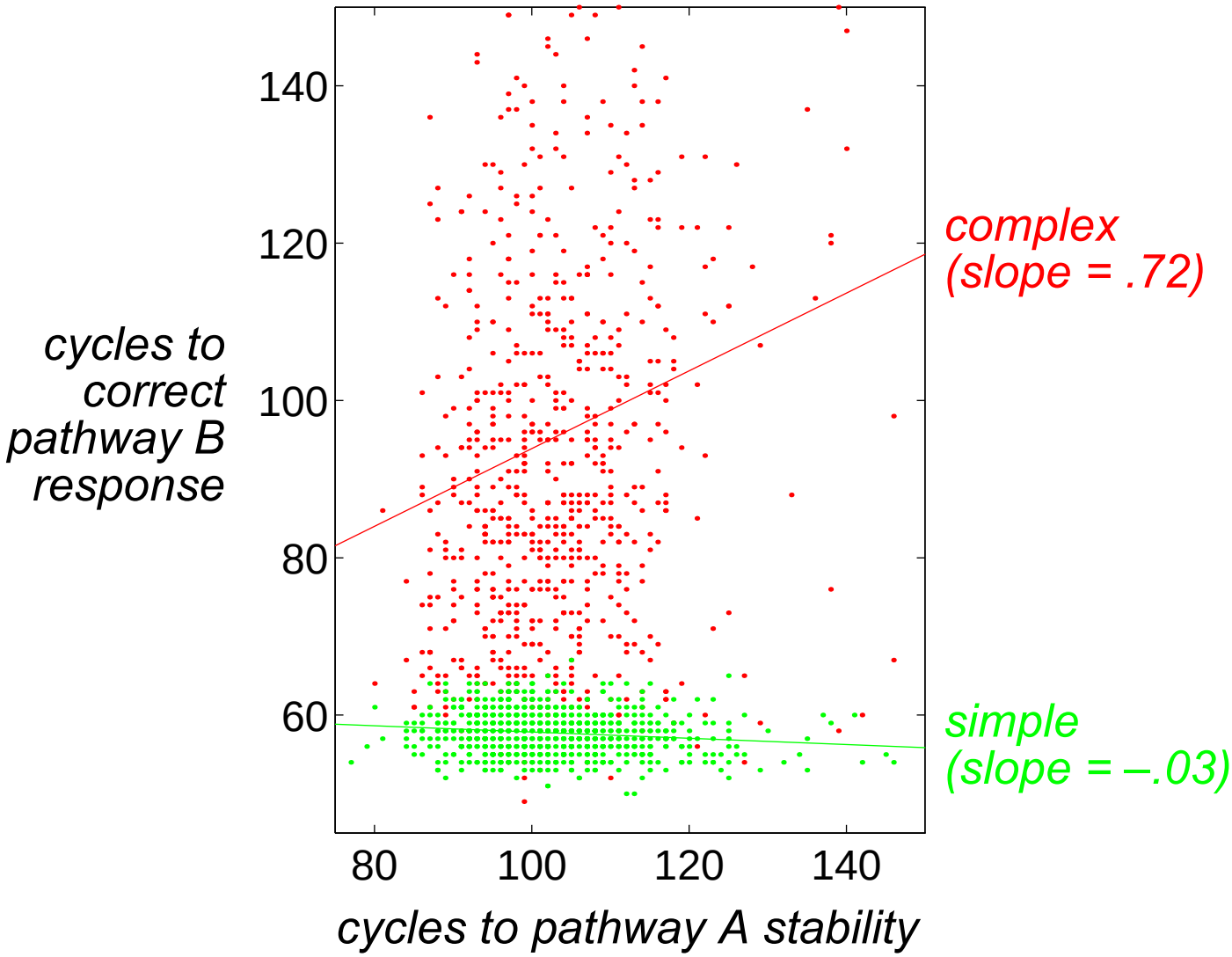


Level of Expertise



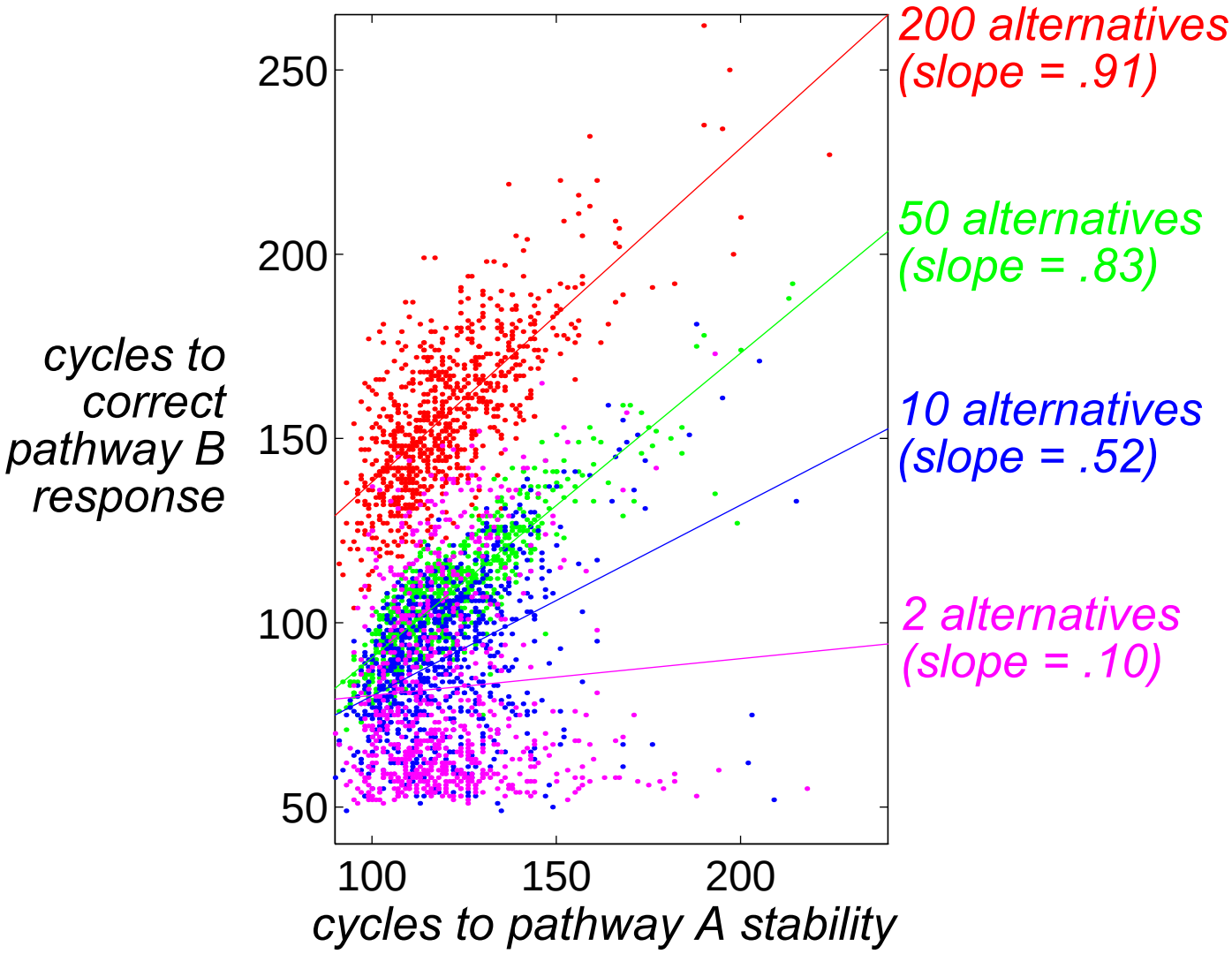
Stability required to initiate underlearned responses.

Task Difficulty



Stability required for tasks relying on finer granularity of information.

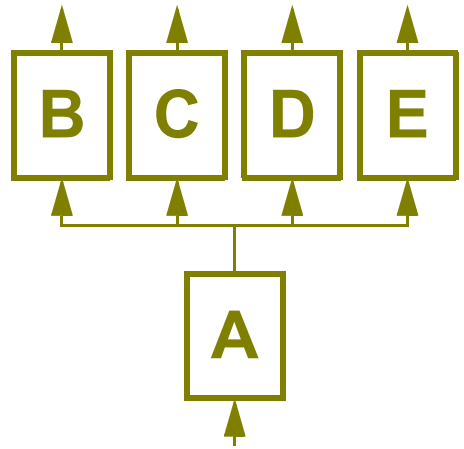
Number of Response Alternatives



Stability required for discriminations involving a large number of alternatives.

Role of Stability

Stability of a pathway's output is *sufficient* to induce the correct response from pathways to which it is connected.

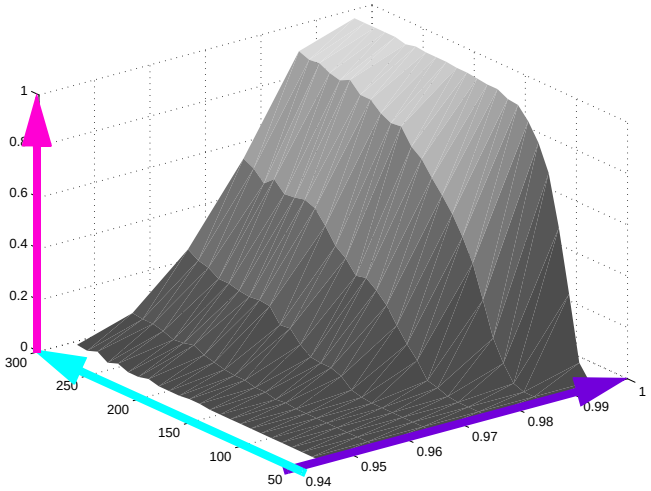


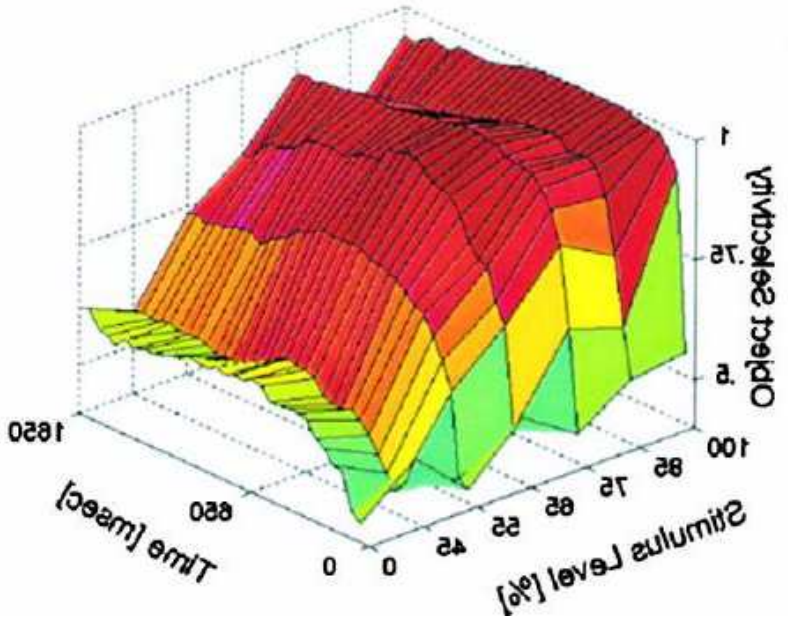
Stability of a pathway's output is *necessary* when

- decision making with limited domain expertise
- responses are arbitrary or complex
- responses involve discrimination among many alternatives

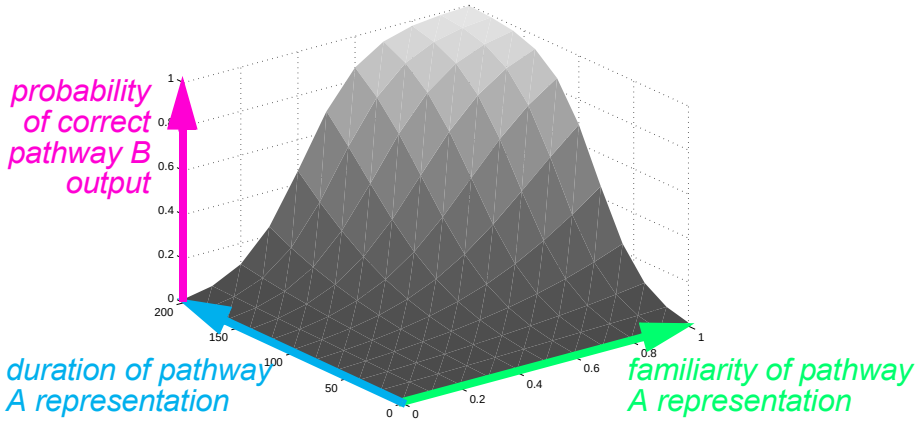
Stable representations are the most *accessible* — they allow for the flexible control of behavior (including verbal report).

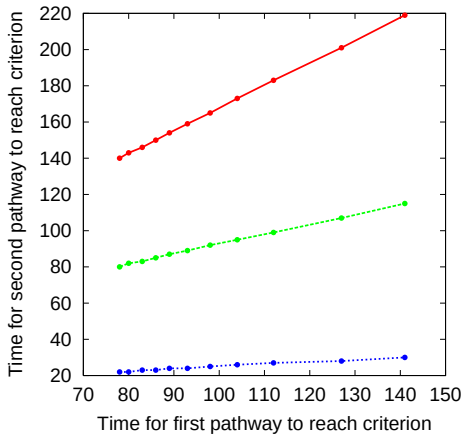
Neurobiological Correlate of Stability?

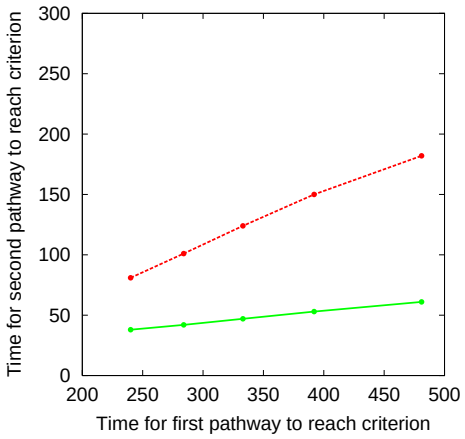


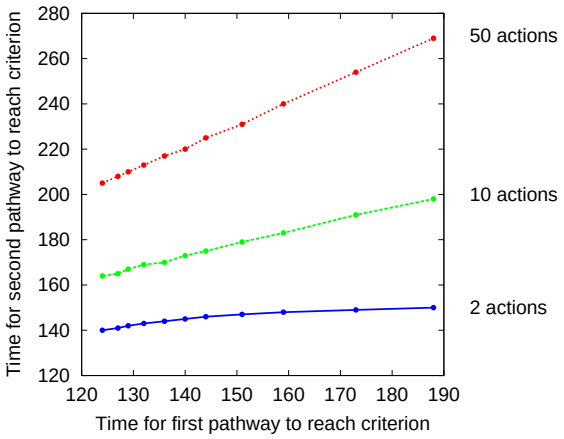


Similar Properties for Belief Net Implementation









Relation to Other Stories about Consciousness 1

Conscious states are interpretations

Attractor state is an interpretation.

Relation to Other Stories about Consciousness 2

Conscious states require high quality representations

Any attractor state is a high quality representation, in the sense that adding noise moves away from attractor and also lowers quality of a representation.

Probably makes more sense to focus on familiarity than quality:
Familiarity a natural way of explaining how pathway A can have an effect on pathway B.

Relation to Other Stories about Consciousness 3

Conscious states require temporal persistence

Relation to Other Stories about Consciousness 4

Conscious states require explicit representations

Relation to Other Stories about Consciousness 5

Conscious states are attractors

Relation to Other Stories about Consciousness 6

Consciousness entails global availability of information to to widespread brain sources

Baars, Dehaene & Naccache: global workspace framework

My story requires functional connectivity between perceptual and response pathways for access to occur.

Frontal areas could provide this functional connectivity, so global workspace ideas complement my story.

Relation to Other Stories about Consciousness 7

Attention is a prerequisite for consciousness

In my story, attention is not necessary, but it makes awareness more likely: Standard notion of attention is to amplify some representations. Attractor corresponding to amplified representation more likely to be chosen.

I.e., biased competition notion can apply to this sort of model as well