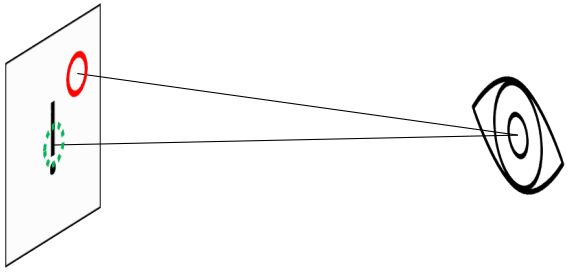


Gazing at the Interface

Kevin Lande
Department of Philosophy / Centre for Vision Research
York University
lande@yorku.ca / www.kevinlande.com

CANADA FIRST RESEARCH EXCELLENCE FUND

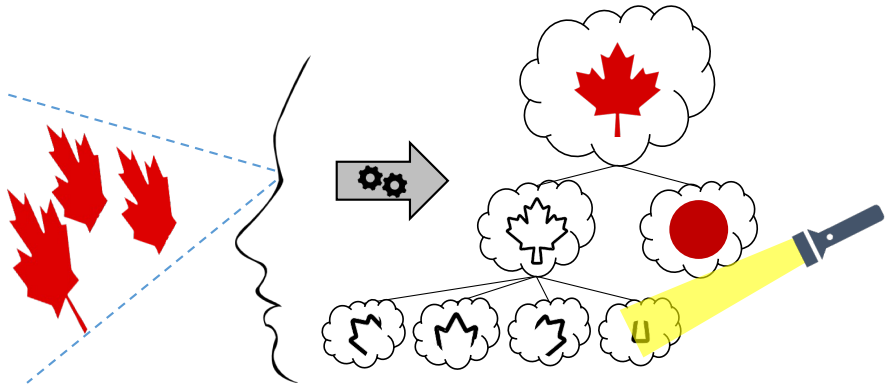
Centre for Vision Research

VISTA VISION: SCIENCE TO APPLICATIONS
YORK UNIVERSITY

Computational Theory of Mind

Mental capacities for perceiving, thinking, remembering, planning, and acting constitutively depend on mental representations that

1. Have representational or informational content
2. Encode that content in a structure/code/format
3. Enter into structure-sensitive, content-respecting causal relations.



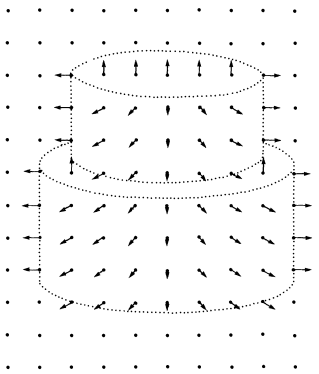
A Plurality of Codes

Psychologists posit different formats for different capacities.

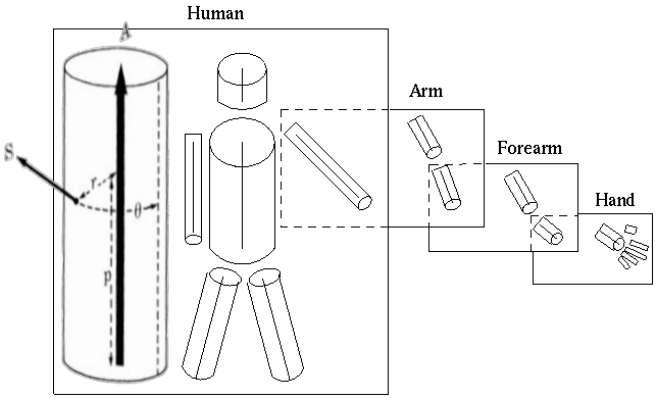
Different formats:

Normally, if Format A ≠ Format B, then representations in A don't function to compose with representations in B.

2.5D Sketch



3D Model



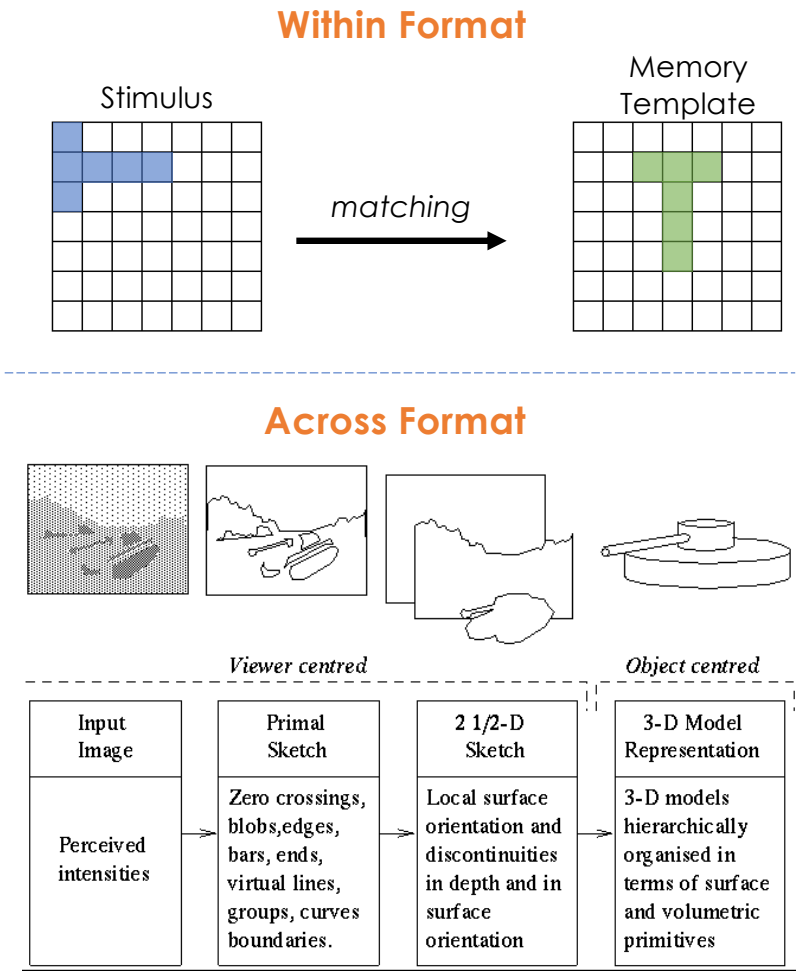
Marr 1982

Code Switching

Computational operations can be:

1. Within format
2. Across format

"Interface"



Marr 1982, Ullman 1996

Everyone's Problem

Perception

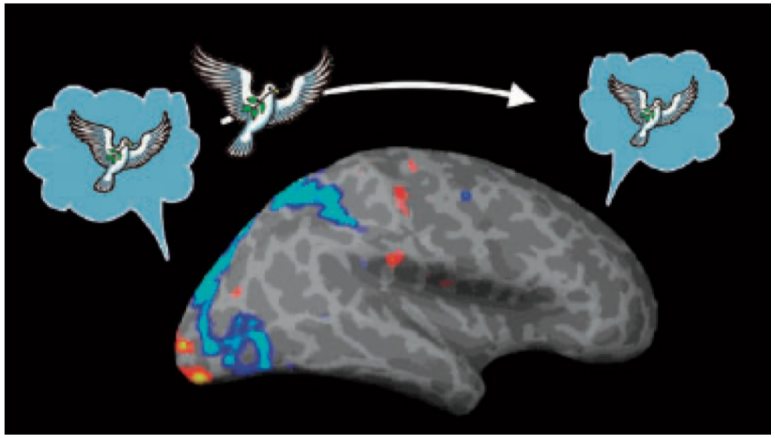
"How does vision communicate with other modules in the brain?" (Cavanagh 2021)

Imagery:

"[The mental imagery system has] to face the problem of providing a seamless interface between its form and the form used in reasoning, since both vision and imagery do play a role in reasoning." (Pylyshyn 2003)

Language:

"all conditions are interface conditions; and a linguistic expression is the optimal realization of such interface conditions." (Chomsky 1995)

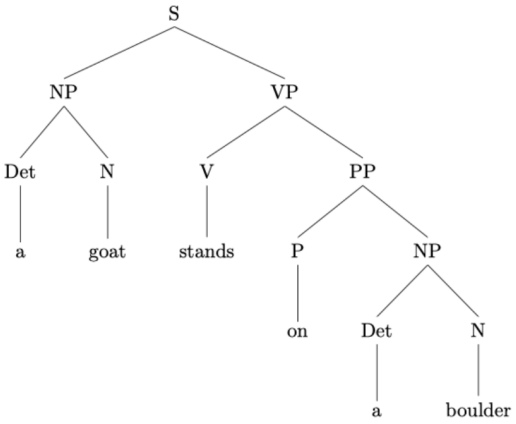


Everyone's Problem

How can a representation R^A, in Format A, and a representation R^B, in Format B, be related such that:

a) **[Functional Coordination]** R^A can be part of a cause of R^B, or *vice versa*, and

b) **[Semantic Coordination]** R^A and R^B are semantically related (e.g. one entails the other, makes the other more probable, they represent situations that are constitutively related, etc.)



Goals

1. What sort of explanatory costs are incurred when positing interfaces? Do they resist computational explanation?

   No.

2. What do such interfaces tend to have in common?

   They are rarely translational.

3. A dilemma for the perception/cognition interface.

Do interfaces resist computational explanation?

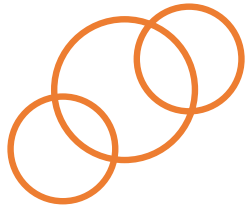
Computing across the interface?

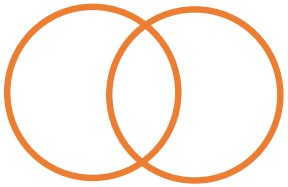
1) **Translation:** "for there to be reliably content-respecting causal processes linking intentions with motor representations there would have to be some process of translation." (Butterfill and Sinigaglia 2014)
   - "The heart of the problem… is one of *translation*: in order for us to talk about what we see, information provided by the visual system must be translated into a form compatible with the information used by the language system." (Jackendoff 1987)
   - "Part of the problem here is simply to understand how non-conceptual representations are 'translated' into conceptual representations." (Heck 2007)

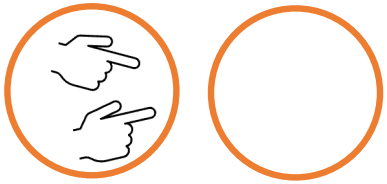
2) **Ignorance:** "The difficulty is that nothing at all is known about this hypothetical translation between intention and motor representation, nor about how it might be achieved, nor even about how it might be investigated."
   - "Another shortcoming of amodal symbol systems is their failure to provide a satisfactory account of the transduction process that maps perceptual states into amodal symbols." (Barsalou 1999)

Other Solutions

- Common Format or Interlingua (e.g. Ferretti and Caiani 2018, Quilty-Dunn, Porot, Mandelbaum 2023)
  - **But:** Evidence for common formats is unclear.
- Mixed Format (e.g. Shepard 2019)
  - **But:** What are the principles of composition and inference? How do we distinguish between mixed format and common format?
- Deference (e.g. Butterfill and Sinigaglia 2014)
  - **But:** Deference either *requires* independent functional coordination (Evans, M&P) or *permits* the lack of functional coordination (Burge, Bach).



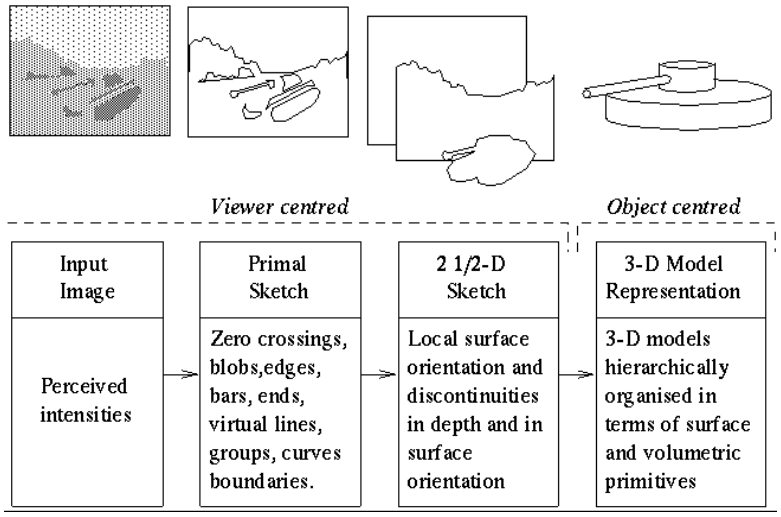




Reconsidering computation across the interface

1. Reject **Translation**
   - Understood as informational equivalence—"mere recoding".

2. Reject **Ignorance**
   - We have models of many interfaces…
3. What do existing models of interface computations have in common?

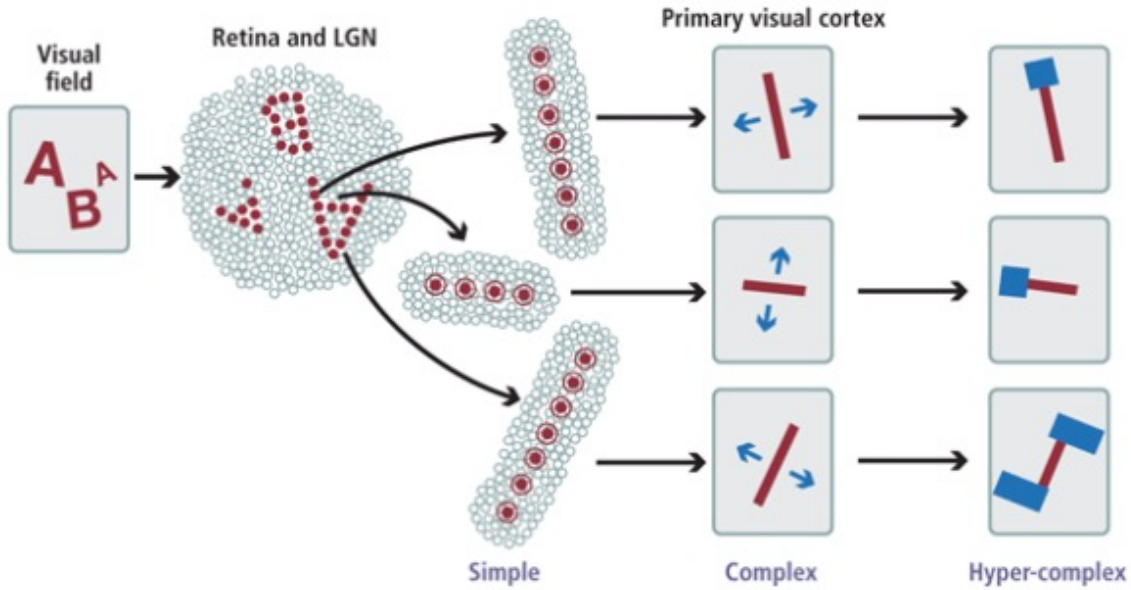


Models of interfaces

A) Vision: Spots to Shapes

Spots to Local Orientation

- Retinal ganglion cells: light and dark contrast
- Ganglion cells converge onto simple cells in V1, which thereby respond to oriented patterns of contrast.



Hubel and Wiesel 1968

Orientation to Contour

The model recovers shape information in five stages:

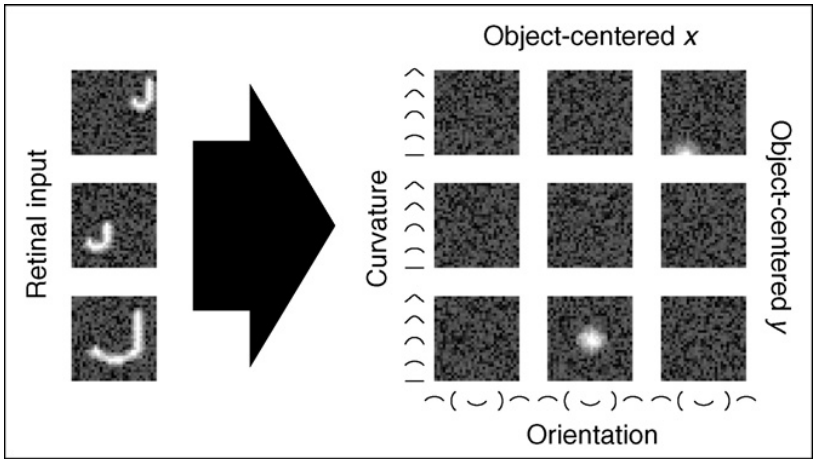
(1) contour information is recovered using oriented filters,

(2) object center is recovered using higher-order filters that respond at the center of concentric contours **(A)**,

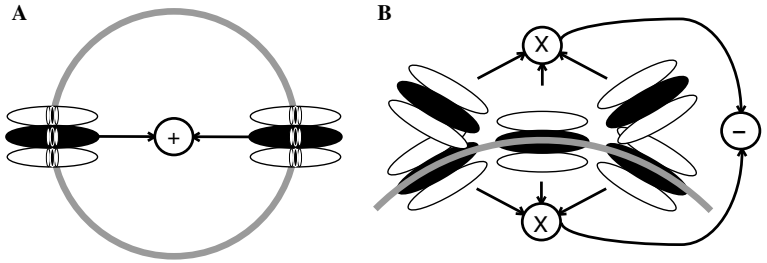
(3) the number and average radii of objects is recovered using the contour energy in one direction from the object center,

(4) local curvature signals are recovered around the contour using a few curvature mechanisms tuned to different degrees of curvature **(B)**, and

(5) shape is represented as curvature signal strength as a function of orientation around the object's center.



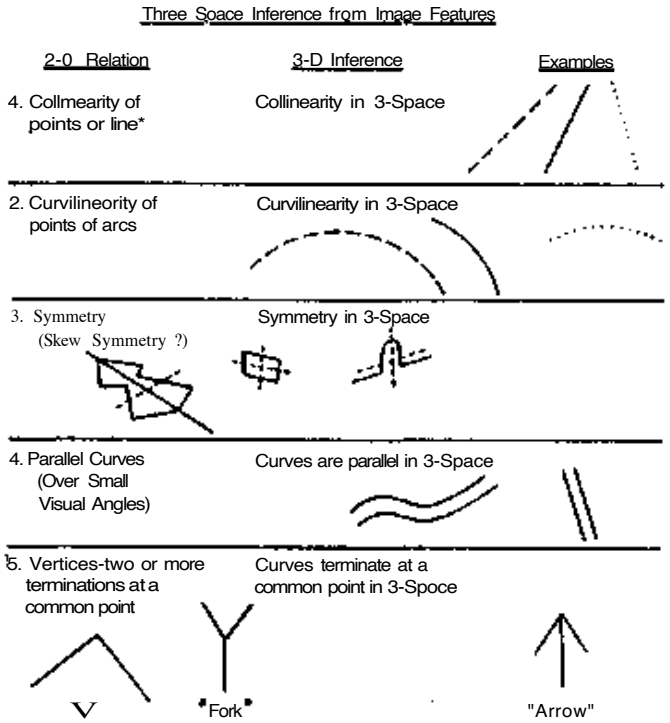
Connor, Brincat, and Pasupathy 2007



Poirier and Wilson 2007

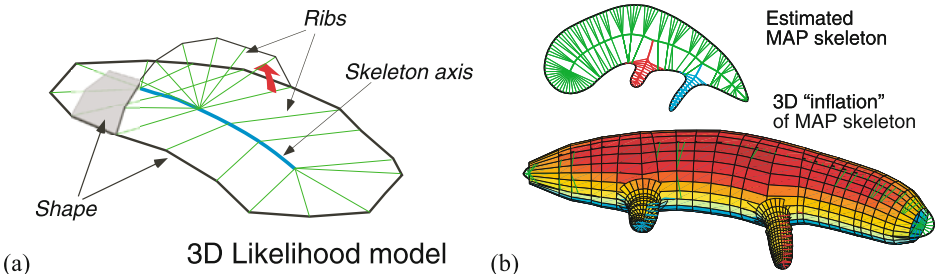
Contour to Solid Shape

Natural Constraints



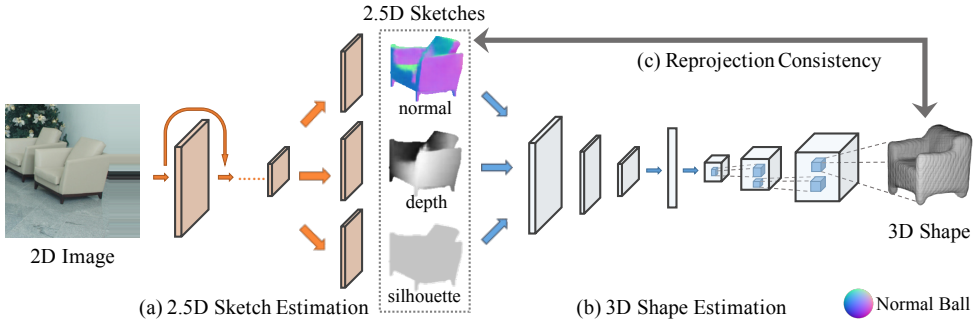
Biederman 1987

Generative Model



Feldman, Singh, Briscoe 2013

DCNNs



Wu et al. 2017

Visual Interfaces, In General

The visual hierarchy contains a series of interfaces.

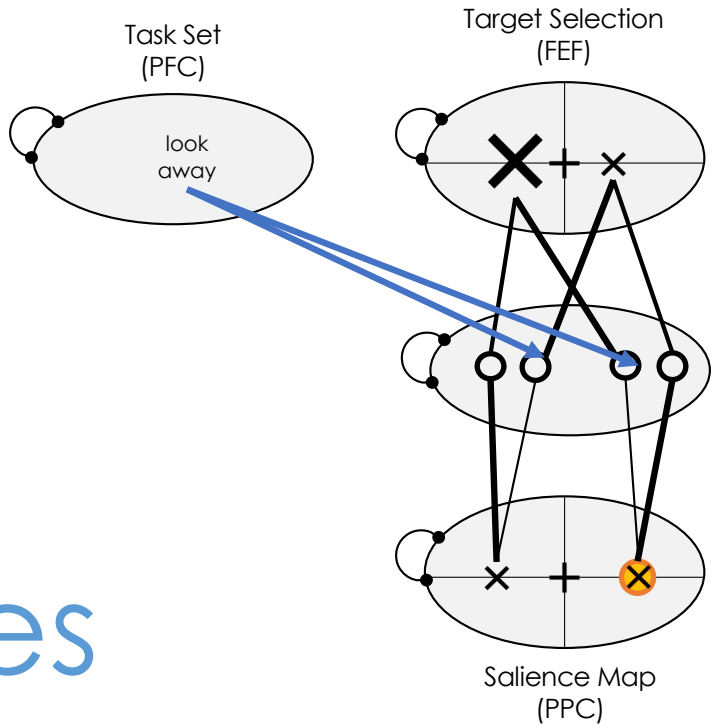
- **Functional coordination:**
  - Certain transitions are primitive operations.
    - Mechanistic explanation @ neurobiological level (inhibition and excitation).
    - Ecological explanation: These operations reflect natural constraints and/or conditional probabilities
  - Structure-sensitive: correspondence between combinatorial features in different formats
    - E.g. skeletal structure of 3D representation depends on vertex-structure in the contour representation.
  - Adaptable (modulated by attention, expectation)

Visual Interfaces, In General

- **Semantic coordination**:
  - Meta-semantics: The contents of the basic representations depend on
    - functional relationships between those basic constituents and other representations (functional coordination), and
    - The relationships between the representations and the environment.
    - Anti-individualism; not deference or demonstration.
  - Compositionality: contents of complex representations (in either format) constitutively depend on the contents and relations between their basic constituents
- **Mere recoding is rare**: Transitions across formats in the visual hierarchy are typically ampliative.

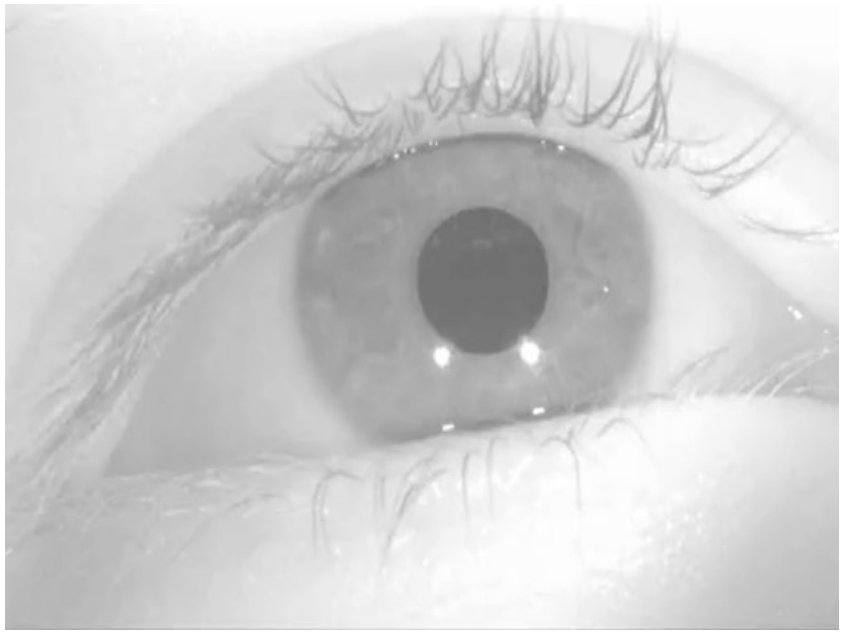
Models of interfaces

B) Oculomotor Control



Eye Movements

- **"Saccade":** a quick, jerky eye movement from one fixation point to another.

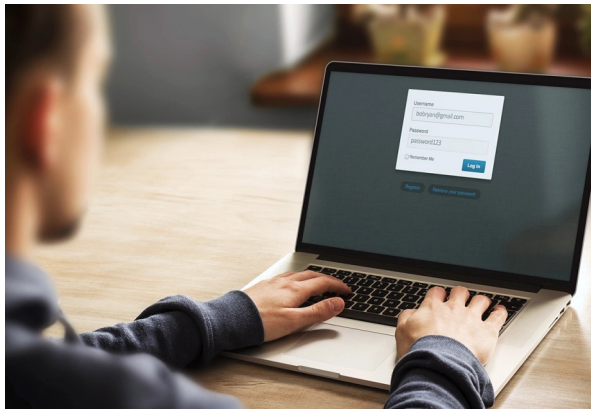


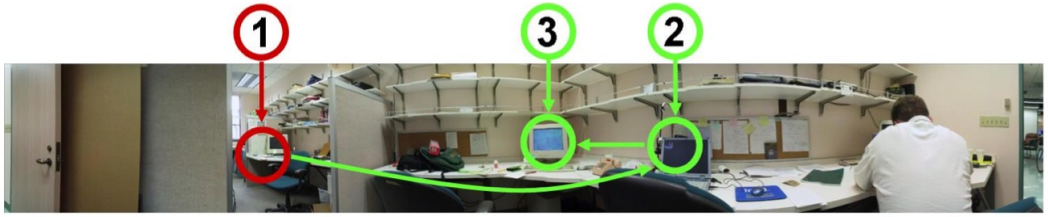
Eye Movements

Schall 2002: "we know more about and have easier access to every stage of the production of visually guided eye movements than we do for limb or vocal movements.

Several lines of evidence indicate that the knowledge gained about the cognitive control of eye movements can generalize to other systems and more complex behaviors…

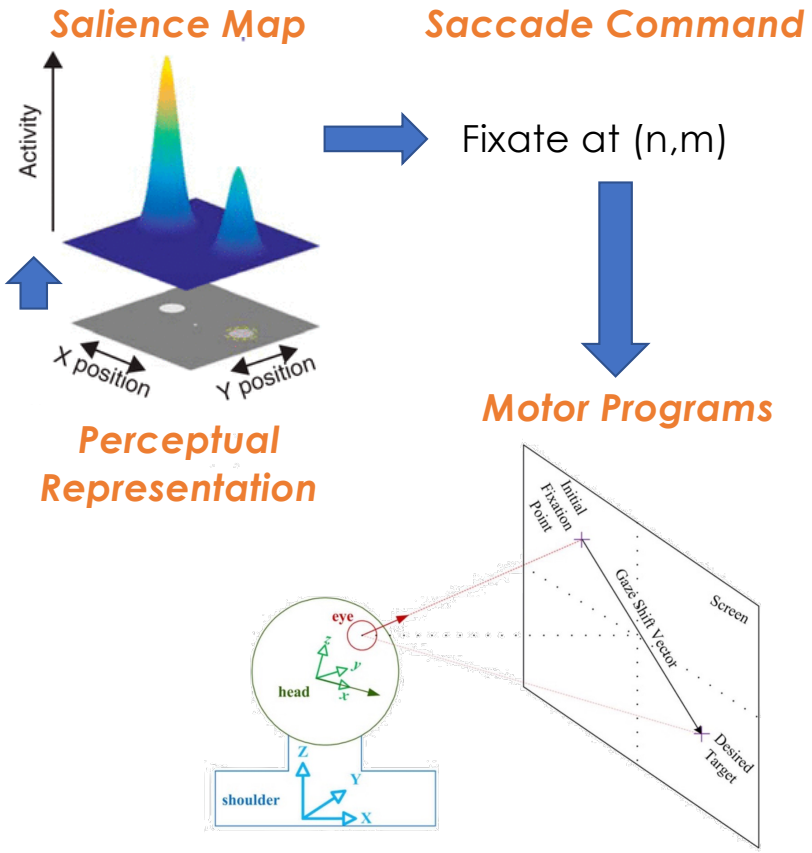
the most general aspects of the voluntary control of behavior seem to be independent of effector."





Reflexive Saccades

- Motor goal is represented in multiple formats,
  - **Visual representations** (3D, eye-centered, object-centered)
  - **Salience map** (PPC) → **Saccade command** (superior colliculus): 2D, eye-centered
  - **Eye motor program** (brainstem): 3D, head-centered vector
  - **Head motor program** (brainstem): 3D, shoulder-centered vector

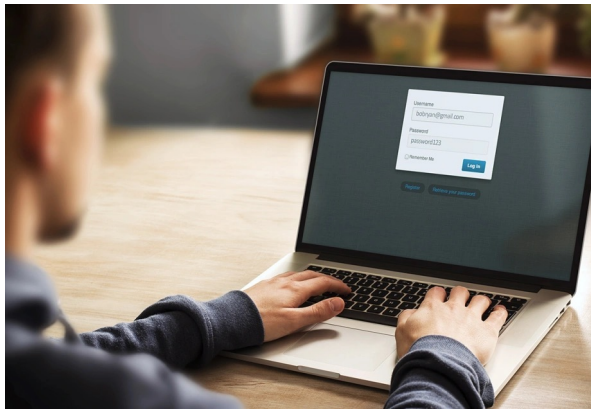


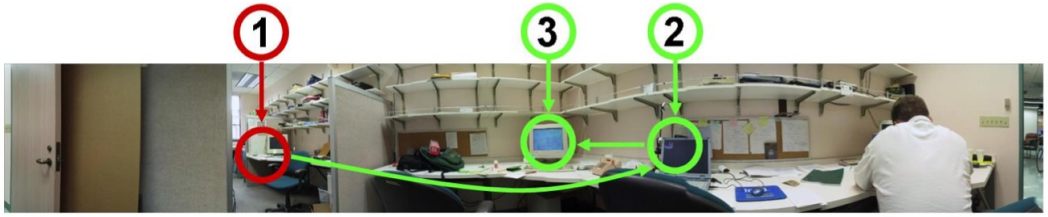
Sajad, Sadeh, Daemi and Crawford 2015, Crawford 2020

Intentional Eye Movements

While often reflexive, saccades sometimes are intentional actions.

- **No-Go**: Don't Look!
- **Anti-saccade**: Look away!
- **Planning**: Look there first, then there, then there.

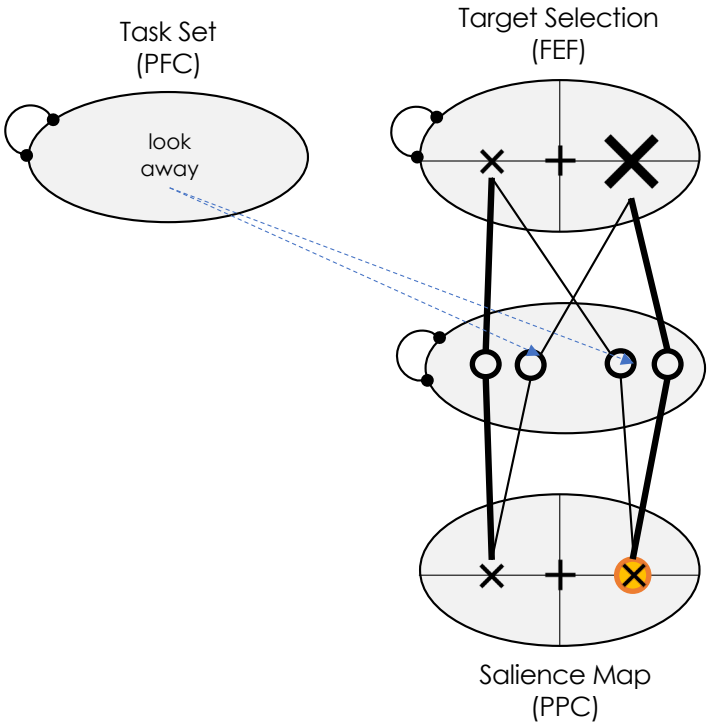




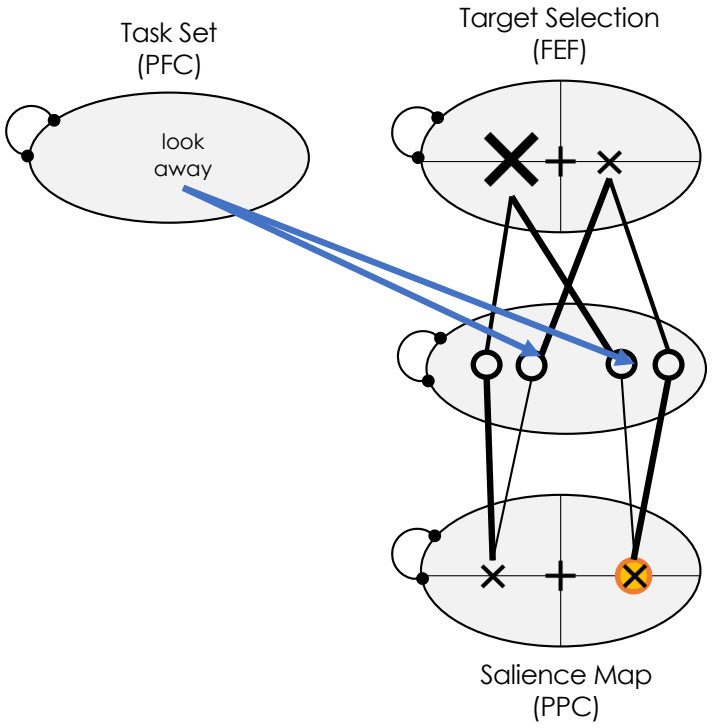
Schall, Palmeri, Logan 2017, Munoz and Everling 2004, Zingale and Kowler 1987

Biased Competition Model

Default: Look



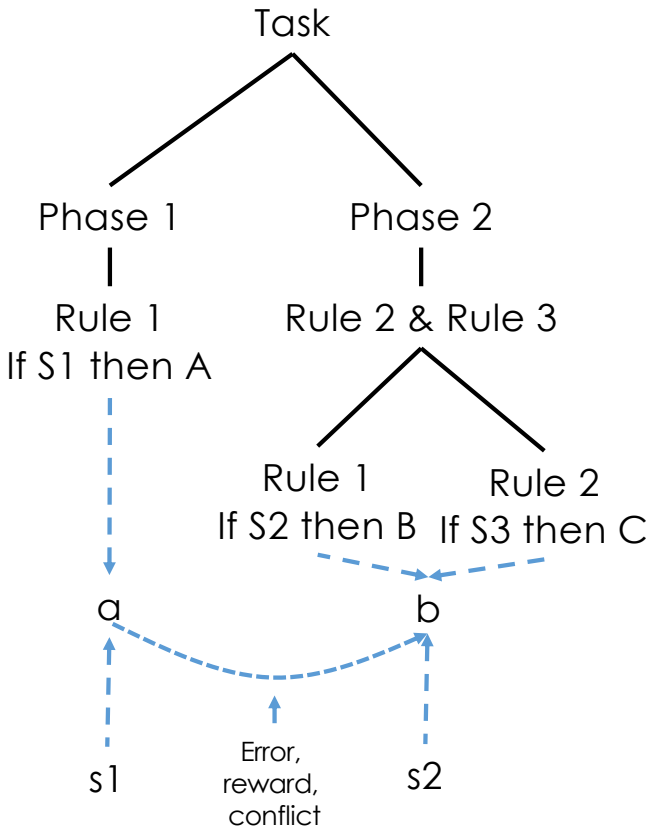
Task: Look away



Miller and Cohen 2001, Heinzle, Hepp and Martin 2007, Wiecki and Frank 2013, Schall 2013, Lo and Wang 2016, Buehler 2021

Structure Matters

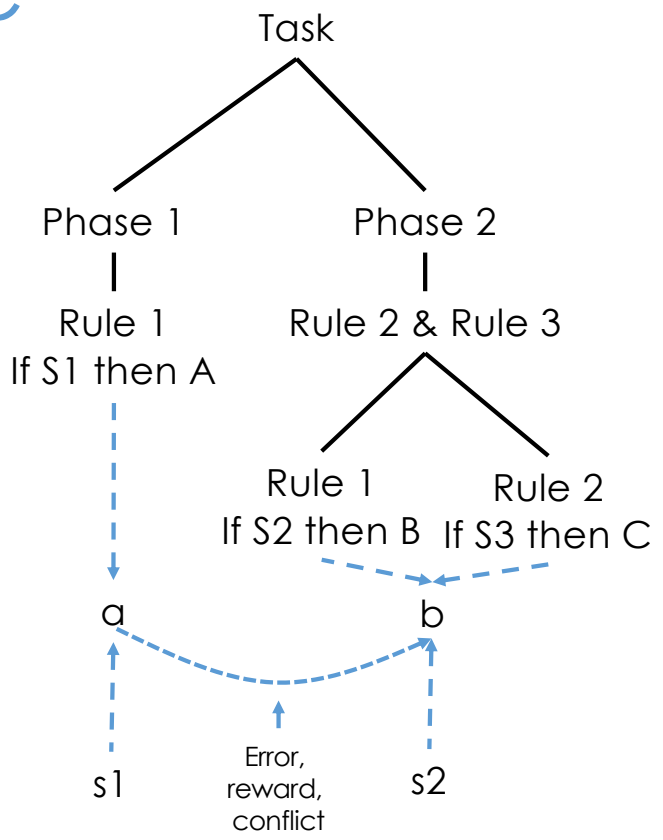
- Format of task representations (PFC)
  - Compound task rules and abstract task sequences are represented compositionally
  - abstract from visual and motor details.
- Sequencing of motor goals (FEF, SC) depends on structured task rules and planning.
- Control is modulated as a function of error, reward, and conflict (SEF, ACC).



Fuji and Greybiel 2003, Sakai 2008, Sigala 2008, Reverberi, Görgen, and Haynes 2012

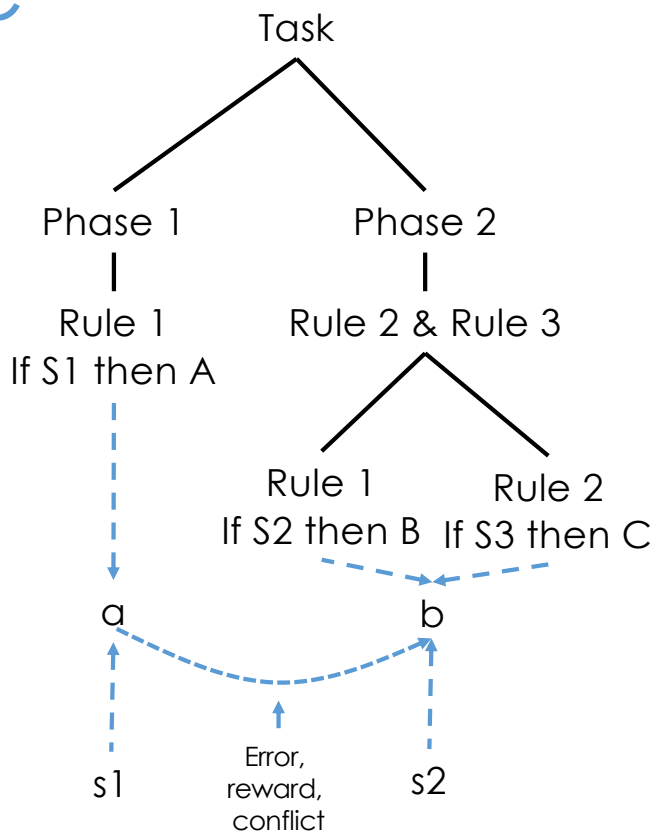
The Oculomotor Interface

- **Functional coordination:**
  - Basic task representations ("look away") modulate visuomotor mappings (vector inversion)
  - Structure-sensitive: sequencing and selection of saccade goals sensitive to complex task representation.
  - Adaptable: selection of saccade goal is sensitive to error, reward, conflict.



The Oculomotor Interface

- **Semantic coordination**:
  - Meta-semantics: Content of basic task rep constitutively depends on
    a) The visuomotor mappings that it modulates, and
    b) The environmental patterns that reinforce the modulatory connections.
  - Compositionality: Complex task reps composed from more basic task reps.



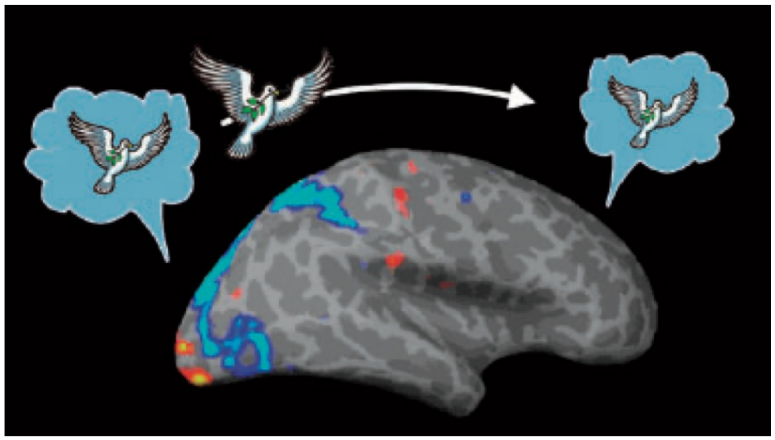
Properties of Interfaces

Common Properties of Interfaces

- **Functional coordination**
  - At basic level, primitive operations--innately specified, or acquired and updated through learning.
  - Structure-sensitive.
  - Often adaptable/modulated across contexts
- **Semantic coordination**
  - Semantic coordination between basic atoms and structural features in Format A and those in Format B is explained meta-semantically.
  - …Given this, semantic coordination between *complex* representations is explained compositionally.
  - Information equivalence (mere recoding) is rare. The brain abhors redundancy.

Is this really an "explanation"?

- Why does it feel like every proposal "begs the question"?
  - E.g., in response to M&P: "one wonders how, after all, action concepts and motor schemata non-accidentally link up in action control" (Shepard 2017).
  - In many proposals, the bridge across the interface seems just to be ***association*** (possibly with fancy things on either side).
- *Any* account (even of a single format) will have to posit primitive operations (Block 1983), for which
  - Functional coordination does not have a further functional analysis at psychological level,
  - And semantic coordination can only be explained meta-semantically.
- But, not the end of the story: structure-sensitive and adaptable, and so non-trivial at psychological level.



A dilemma for conceptualization

A dilemma for conceptualization

- Typically understand as (selective) translation of perceptual representation into propositional format
  - "every concept in a perceptual belief *conceptualizes* a perceptual attributive in the underlying perceptual state…. Every conceptual attributive in a perceptual belief must have a counterpart perceptual attributive in the perceptual state from which the belief (and the conceptual attributive) is formed. The concept must indicate and attribute the same attribute that the perceptual attributive does…" (Burge 2020).
  - "A minimal perceptual judgment conceptualizes each representational aspect of a perception and no more" (Block 2023).
- Dilemma: either
  - The perceptual-conceptual interface is akin to others and skews toward non-translational, ampliative transitions (Westfall 2020).
  - Or this interface is special, which calls for explanation.

In sum

1. What sort of explanatory costs are incurred when positing interfaces? Do they resist computational explanation?

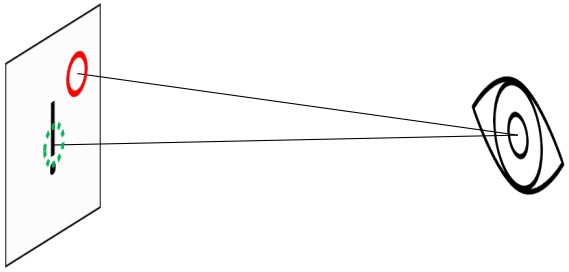
   No.

2. What do such interfaces tend to have in common?

   They are rarely translational.

3. A dilemma for the perception/cognition interface.

SCAN FOR SLIDES



Thanks!

Kevin Lande
Department of Philosophy / Centre for Vision Research
York University
lande@yorku.ca / www.kevinlande.com

CANADA FIRST RESEARCH EXCELLENCE FUND

Centre for Vision Research

VISTA VISION: SCIENCE TO APPLICATIONS
YORK UNIVERSITY