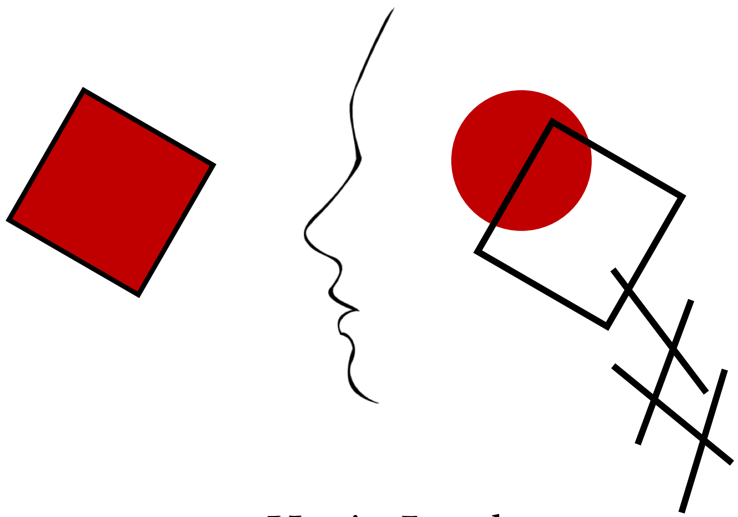
SCAN FOR SLIDES

Towards a Compositional Semantics of Perception



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

Philosophy of Psychology

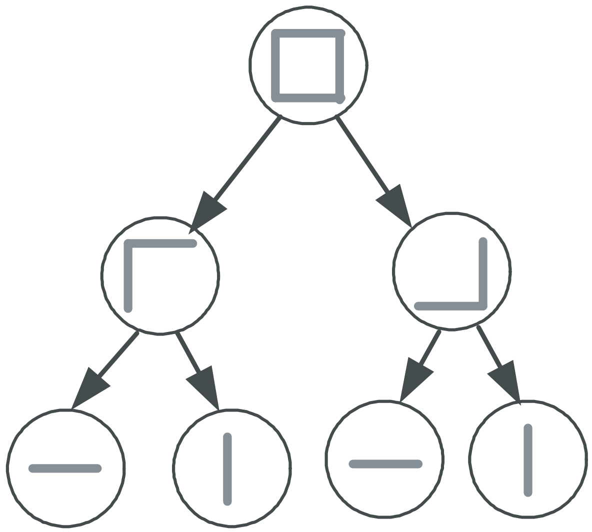
Philosophy of science

Empirically informed philosophy of mind

The Visual Code

“it is eminently reasonable to believe that the perceptual representation of a square includes the representation of lines as subparts. … perceptual representations are selectively organized data structures.”

-Stephen Palmer, “Hierarchical Structure in Perceptual Representation” (1978)



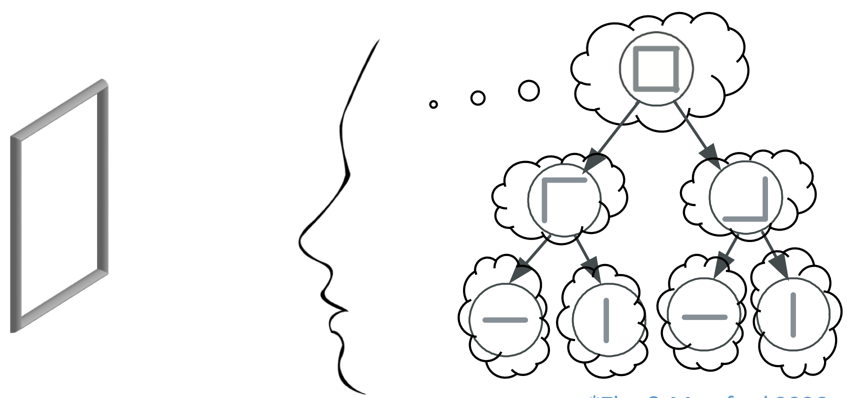
*Zhu & Mumford 2006

Three Foundational Principles

1. Perceptual states represent
2. Perceptual states have constituent structure
3. What a perceptual state represents is a function solely of what its constituents represent and how they are related.

It should be possible to develop empirically informed semantic theories of perception, as there are for language.

- What would such a theory look like?
- How could be empirically constrained?
- What would it tell us about the nature of perceptual representation?



*Zhu & Mumford 2006

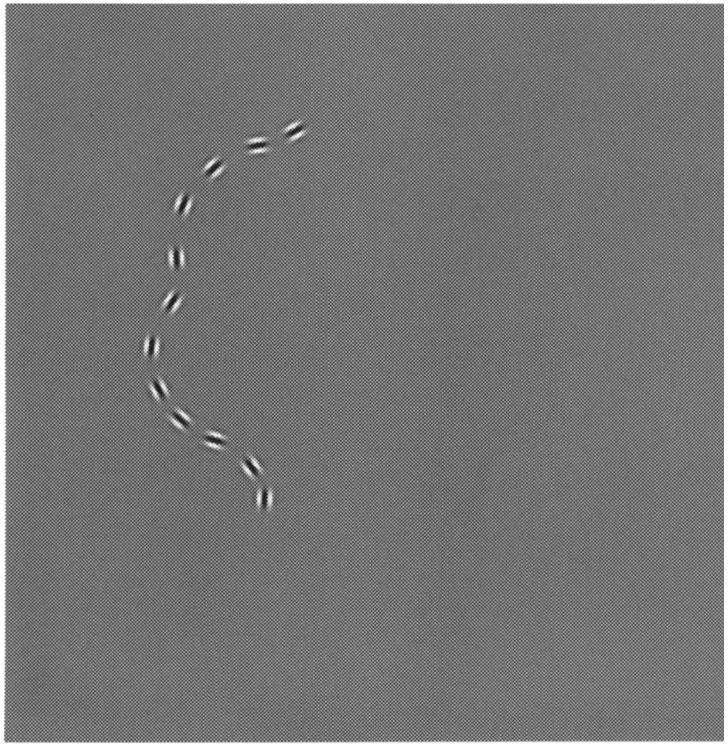
Outline

Background

1. The Semantics of Mental Representations

Case study: contour perception

2. Concatenating Fragments
3. Integrating Features
4. Integrating Contours
5. The Format of Contour Perception
6. The Whole and the Parts



1. **The Semantics of Mental Representations**
2. Concatenating Fragments
3. Integrating Features
4. Integrating Contours
5. The Format of Contour Perception
6. The Whole and the Parts

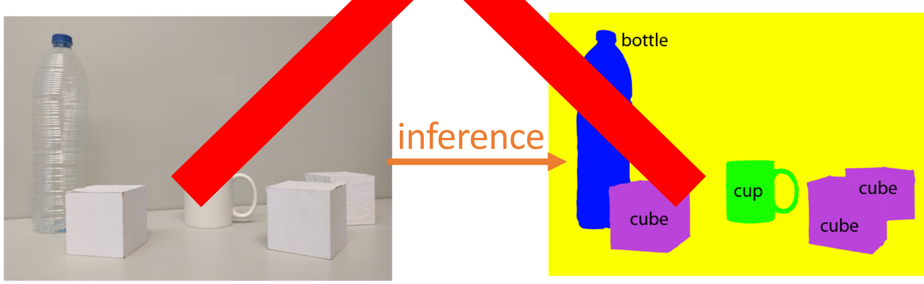
An Origin Story

*Frege 1923; Tarski 1936; Davidson 1967; Montague 1974; Partee 1975; Palmer 1978

Semantics – Two Conceptions

Cognitive Psychology/Computer Vision

- Mapping from input representations ([illegible] images, percepts) to co[illegible]ptual rep[illegible]entations/labels (*r*[illegible]*n*, *chef*, *laundr*[illegible]*t*)
- Relations a[illegible]ng r[illegible]esentations “inside the sy[illegible].”



Input: Image

Output: Conceptual labels

Philosophy, Logic, Linguistics

- A theory that specifies what contents or information representations carry.
- Mapping from representations to world.
  - Reference (to what individual or property does a term refer?)
  - Truth (under what conditions is the representation true?)

Representational System **World**

*Heim & Kratzer 1998

Compositionality

- The content of a complex representation is a function of the contents of its parts and the way they are combined.
- Exceptionally novel contents can thereby be "coded" efficiently in terms of more commonplace elements and forms of combination.

Representational System | **Semantics** | **World**

$\{x\text{: } x \text{ smokes}\}$

⟦"Ann smokes"⟧ is true iff 👩🏻 ∈ $\{x\text{: } x \text{ smokes}\}$

S

NP VP

N V

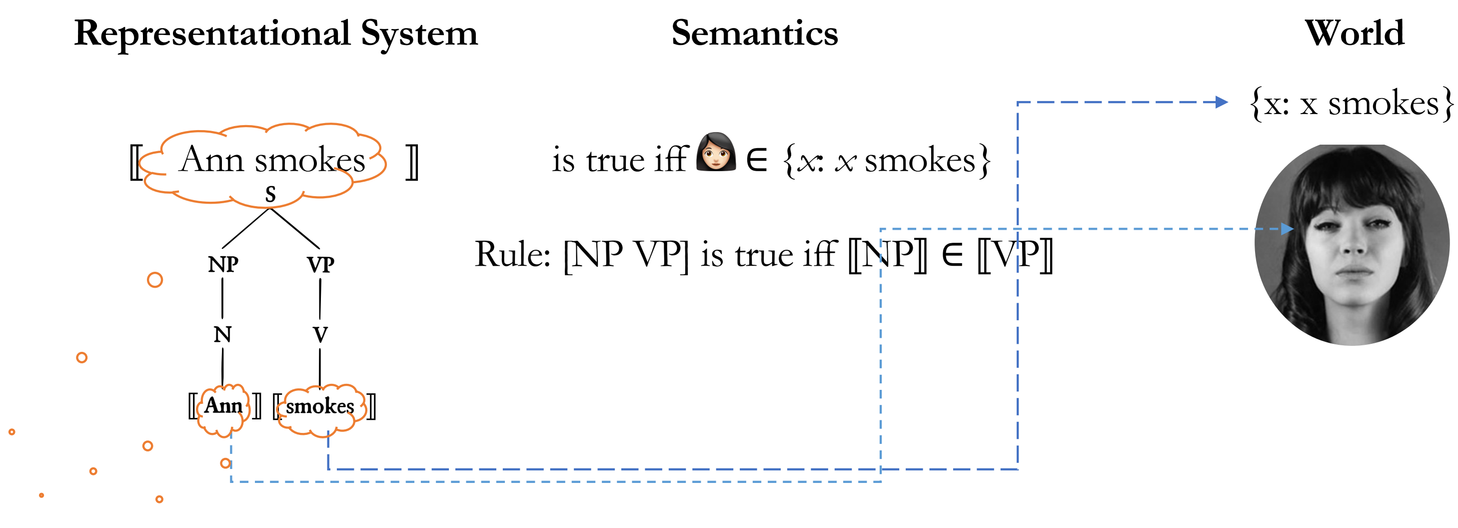
⟦ Ann ⟧ ⟦ smokes ⟧

Rule: [NP VP] is true iff ⟦NP⟧ ∈ ⟦VP⟧

Syntax

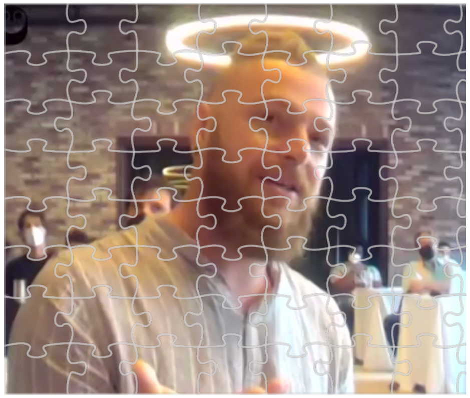
Semantics for Mental Representation

1. The Language of Thought Hypothesis: Just as natural and artificial languages have a syntax and semantics, computational psychology presupposes that so too do mental states involved in thinking, believing, desiring, and intending. (Fodor 1975)



Semantics for Mental Representation

1. The Language of Thought Hypothesis: Just as natural and artificial languages have a syntax and semantics, computational psychology presupposes that so too do mental states involved in thinking, believing, desiring, and intending. (Fodor 1975)
2. There are non-linguistic types of representation (images, pictures, etc.) that fall outside the scope of compositional syntax and semantic theorizing. (Kosslyn, Cummins)



"Images do not seem to have a syntax."

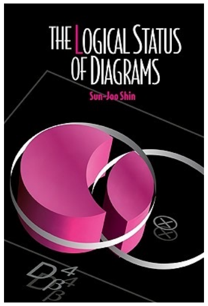
-Stephen Kosslyn, *Image and Mind* (1980)

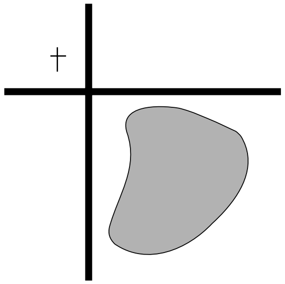
"Pictorial representations do not have a canonical structure. Their structure can be analyzed in many different ways (corresponding to the jigsaw puzzles that one can construct from it), but none of these can properly be described as giving the structure of the state of affair."

-José Bermúdez, "Two Arguments for the Language-Dependence of Thought" (2010)

Semantics for Mental Representation

1. The Language of Thought Hypothesis: Just as natural and artificial languages have a syntax and semantics, computational psychology presupposes that so too do mental states involved in thinking, believing, desiring, and intending. (Fodor 1975)
2. There are non-linguistic types of representation (images, pictures, etc.) that fall outside the scope of compositional syntax and semantic theorizing. (Kosslyn, Cummins)
3. There are non-linguistic types of representations (diagrams, maps) that have compositional syntax and semantics, but very different from language. (Shin, Casati & Varzi, Camp, Rescorla)





"A division between syntax and semantics is not intrinsic to symbolic systems only. For *any* representation system, whether it is symbolic or visual, we can discuss two levels, a syntactic level and a semantic level. This is what representation means."

-Sun-Joo Shin, *The Logical Status of Diagrams* (1994)

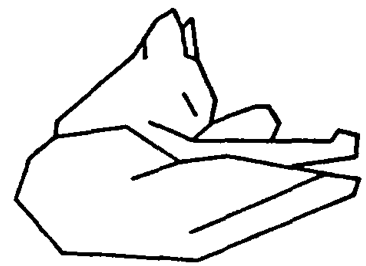
Semantics for Mental Representation

1. The Language of Thought Hypothesis: Just as natural and artificial languages have a syntax and semantics, computational psychology presupposes that so too do mental states involved in thinking, believing, desiring, and intending. (Fodor 1975)
2. There are non-linguistic types of representation (images, pictures, etc.) that fall outside the scope of compositional syntax and semantic theorizing. (Kosslyn, Cummins)
3. There are non-linguistic types of representations (diagrams, maps) that have compositional syntax and semantics, but very different from language. (Shin, Casati & Varzi, Camp, Rescorla)
4. Empirically grounded semantic theories for perceptual representations. (Matthen, Clark(e), Burge, Quilty-Dunn, Lande)

A Case Study: Contour Perception

“the human visual system represents contours and shapes in a piecewise manner. In other words, it segments contours and shapes into simpler ‘parts’ and organizes shape representations using these parts and their spatial relationships. Far from being arbitrary subsets, these perceptual parts are highly systematic, and segmented using predictable geometric ‘rules.’

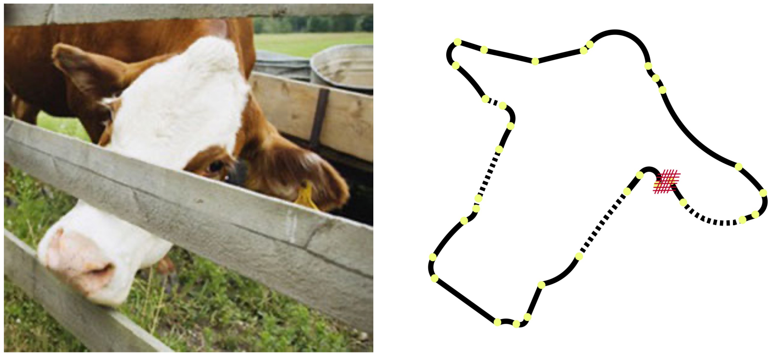
-Manish Singh, “Visual Representation of Contour and Shape” (2015)”



*Attneave 1954

“Just as linguists must work hard to discern the eligible phonemes and computational rules that ultimately lead to the well-formed syntax of a native speaker, so too must vision scientists carefully design experiments to figure out the features and compositional rules that govern [contour representations].

-Brian Keane, “Contour Interpolation: A Case Study in Modularity of Mind” (2018)”



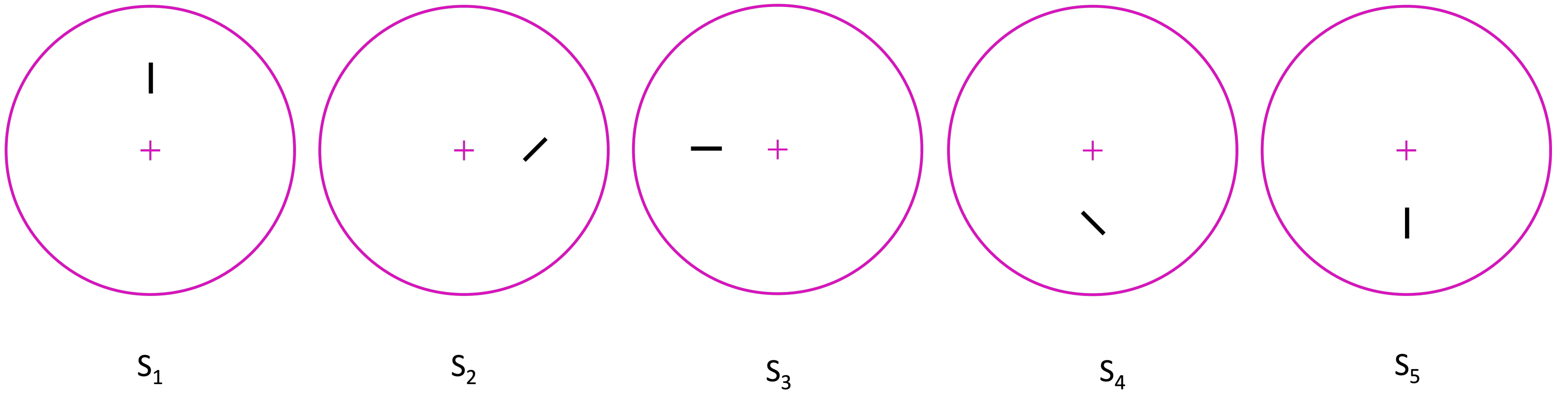
A B

*Kellman 2013

1. The Semantics of Mental Representations
2. **Concatenating Fragments**
3. Integrating Features
4. Integrating Contours
5. The Format of Contour Perception
6. The Whole and the Parts

A warm-up.

Accuracy Conditions



(S_1) If $\mathbf{S_1}$ is a representation of *x* in context δ, then it is perfectly accurate ⇔ (if and only if) *x* is a contour element oriented 0° w.r.t. gravity and located at top of scene in δ.

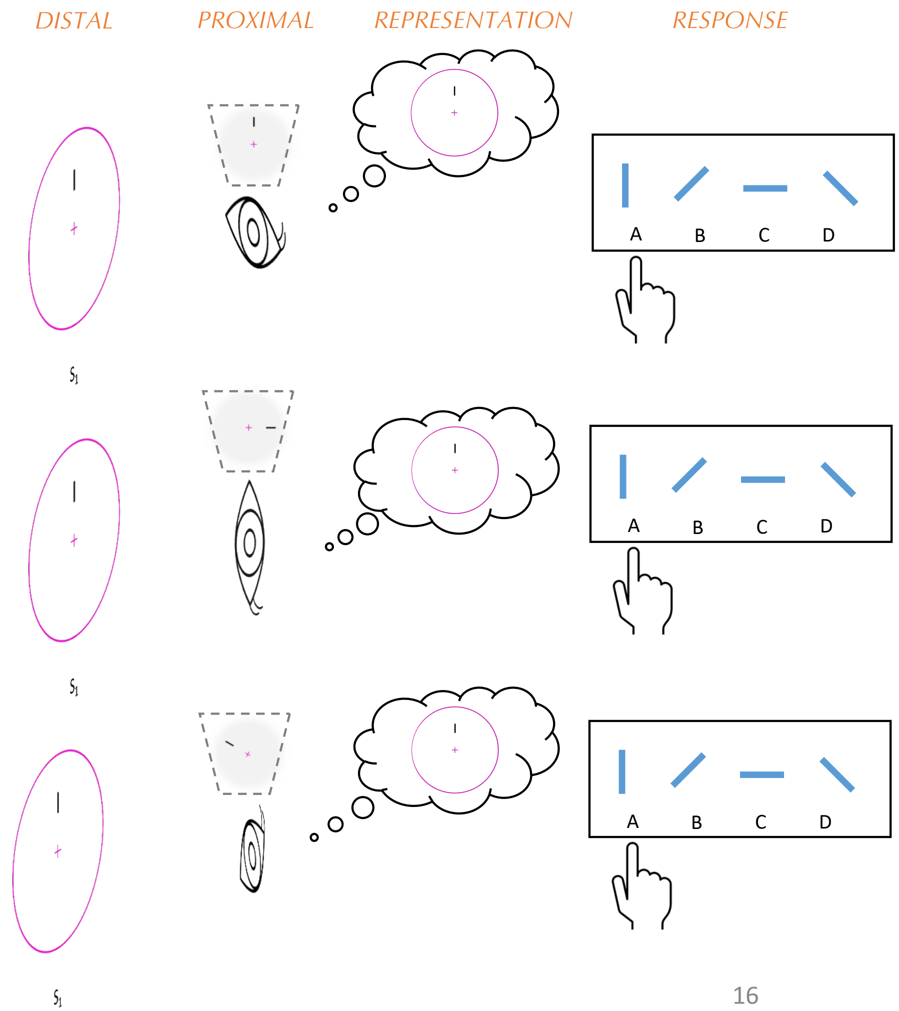
…

*Cf. Burge 2010

Accuracy Conditions: *Representation*

(S_1)	If $\mathbf{S_1}$ is a representation of x in context δ, then it is perfectly accurate $\Leftrightarrow$ x is a contour element oriented 0° w.r.t. gravity and located at top of scene in δ.

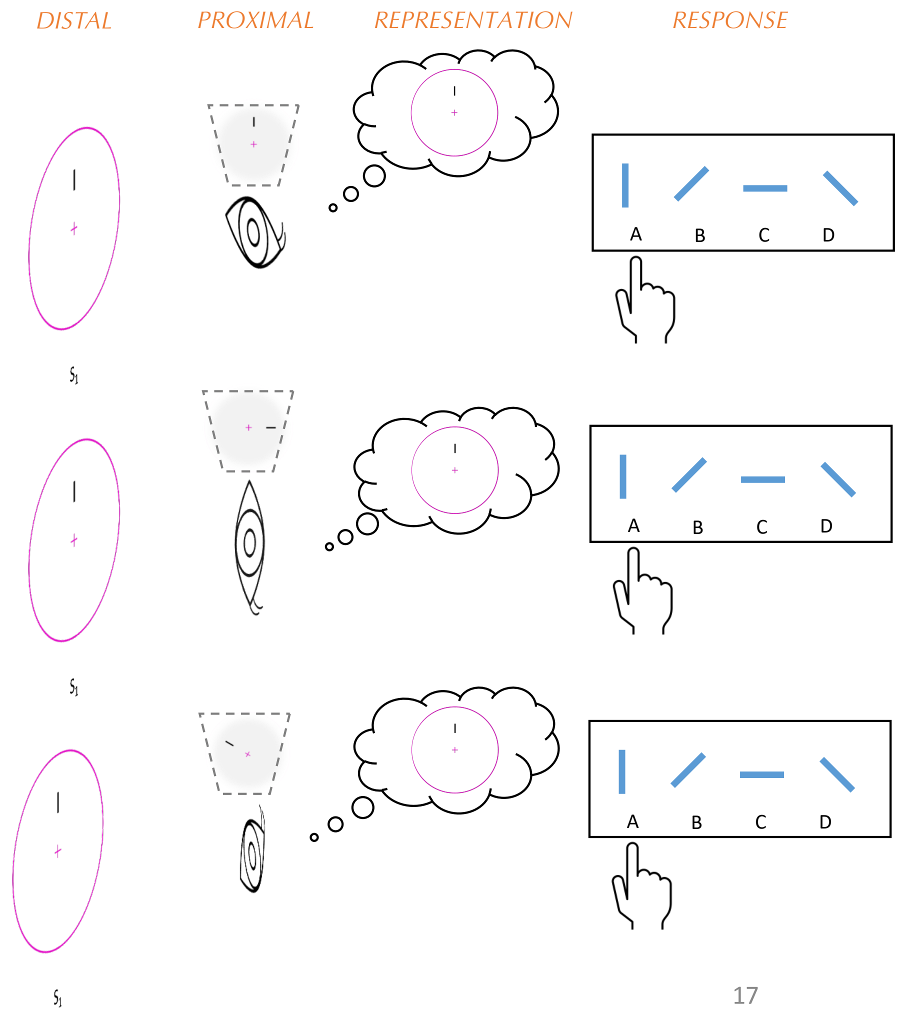
1. Orientation of the distal stimulus predicts task responses better than any independently individuated feature of the proximal stimulus.
2. Task responses cannot causally depend on the distal stimulus itself.
3. An explanation of task responses is that they causally depend on an internal representation or estimate of distal orientation, such that, other things equal, the distal orientation is predictive to the degree that the internal representation is accurate.



Accuracy Conditions: *Accuracy vs. Truth*

(S$_1$) If **S$_1$** is a representation of x in context δ, then it is perfectly accurate $\Leftrightarrow$ x is a contour element oriented 0° w.r.t. gravity and located at top of scene in δ.

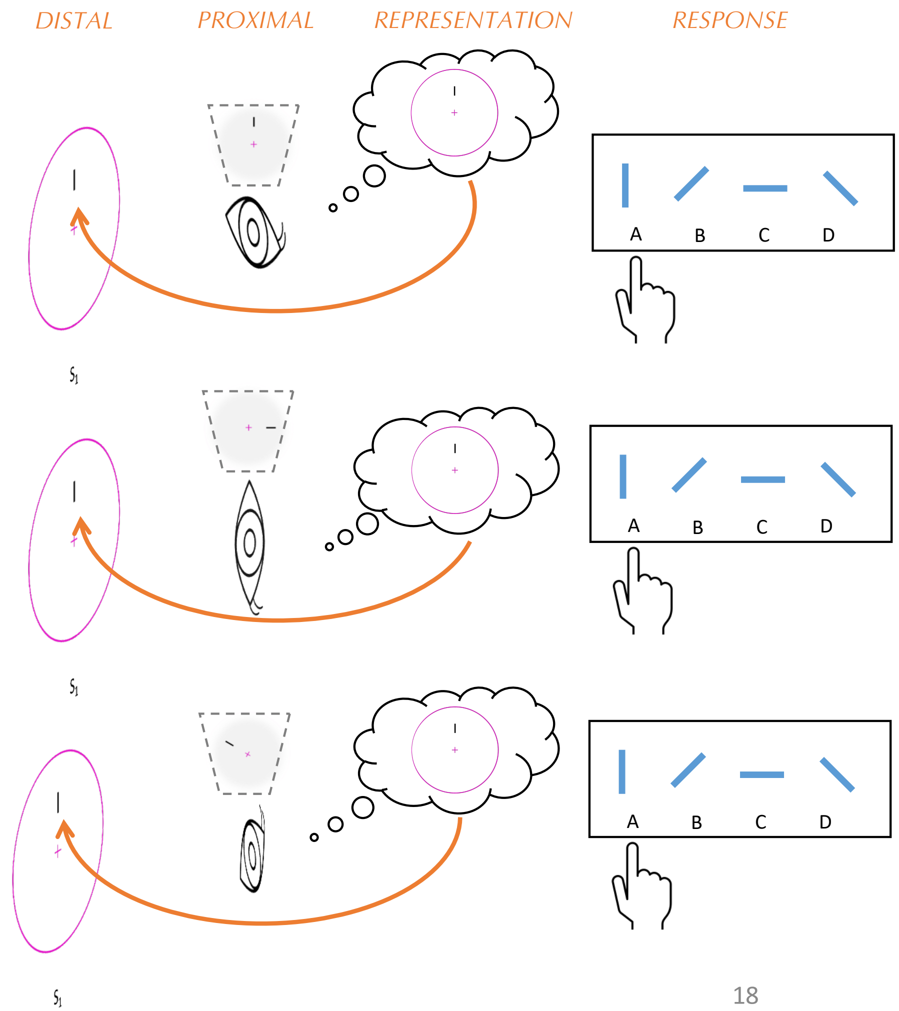
1. Accuracy comes in degrees
   - Distal orientation predicts response *to the degree that* the representation is accurate.
2. Accuracy does not require propositional truth
   - Non-propositional representations (e.g. images) can be accurate.
   - Sub-propositional representations (“That vertical column”) can be accurate.
   - True propositional representations (“That vertical column predates Caesar”) can be inaccurate.



*Cf. Kaplan 1989; Burge 2010; Recanati 2012; Lande 2021b

Accuracy Conditions: *Perceptual Reference*

(S_1) If **S_1** is a representation of x in context δ, then it is perfectly accurate $\Leftrightarrow$ x is a contour element oriented 0° w.r.t. gravity and located at top of scene in δ.

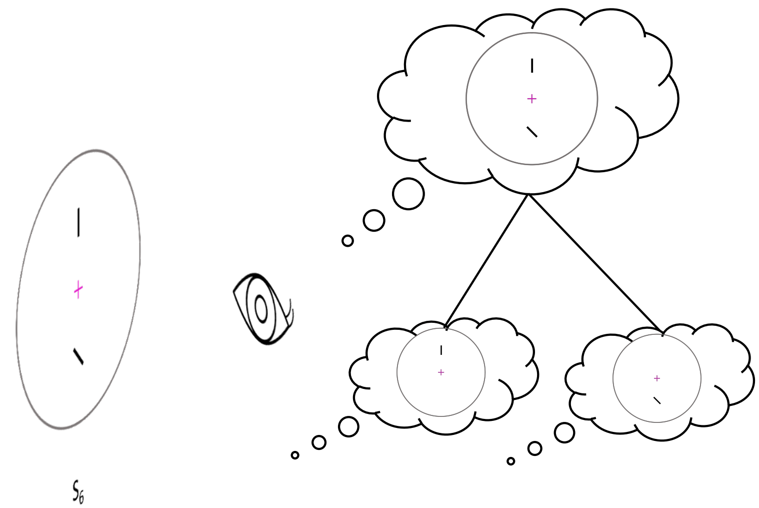


Accuracy is evaluated with respect to a target referent in the context in which the perceptual representation is formed.

Accuracy Conditions: *Structural Description*

(S_1)	If **{$\mathbf{e_a}$, $\mathbf{e_b}$}** is a representation of *x* in context δ, then it is perfectly accurate $\Leftrightarrow$ *x* is a contour element oriented 0° w.r.t. gravity and located at top of scene in δ.

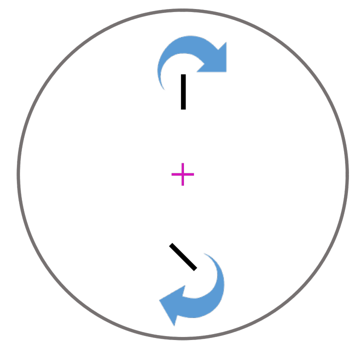
The name of the perceptual representation ("$\mathbf{S_6}$") can be replaced with a structural description ("**{$\mathbf{e_a}$, $\mathbf{e_b}$}**") of the representation.



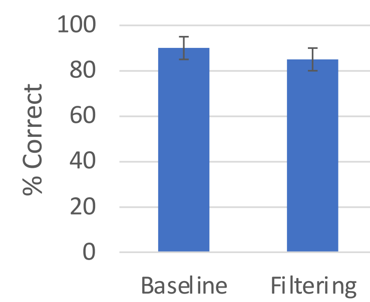
*Garner 1974; Ashby & Townsend 1986; Algom & Fitousi 2016

Structural Description

- Structural/syntactic descriptions function to explain the distribution of representations: how they can, cannot, and must co-occur (other things equal). (Lande 2021a)
- E.g. Garner's speeded classification task
  - **Task:** discriminate feature (e.g. orientation of top element).
  - **Baseline condition:** vary target feature while holding non-target feature constant across trials.
  - **Filtering condition:** vary both features.
  - If filtering performance doesn't differ from baseline performance, then variance in response to target feature is independent of variance in response to non-target feature.
  - Explanation: The orientations of the two elements are coded by separate variables, with distinct variances: $\mathbf{S_6 = \{e_a, e_b\}}$



S_6



Semantics of *Concatenation*

REPS = set of representations in system

Atoms

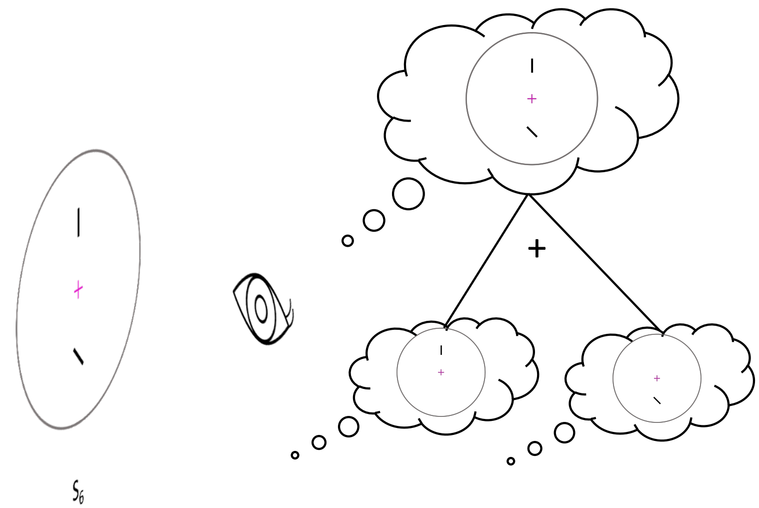
1. Syntax: $\{\mathbf{e}_{t,0°}, \mathbf{e}_{t,45°}, \ldots \mathbf{e}_{r,0°}, \ldots, \mathbf{e}_{l,135°}\} \subseteq$ REPS
2. Semantics:
    1. If $\mathbf{e}_{t,0°}$ is a representation of x in δ, then it is perfectly accurate $\leftrightarrow$ x is a contour located at top of scene and oriented 0° w.r.t. to scene in δ.
    2. …

Concatenation +

1. Syntax: $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n \in$ REPS, where $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n\}$,

    IF $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n \in$ REPS.

1. Semantics: $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n$ is perfectly accurate $\Leftrightarrow$ for every constituent $\boldsymbol{\alpha}_i$, $\boldsymbol{\alpha}_i$ is perfectly accurate in δ.



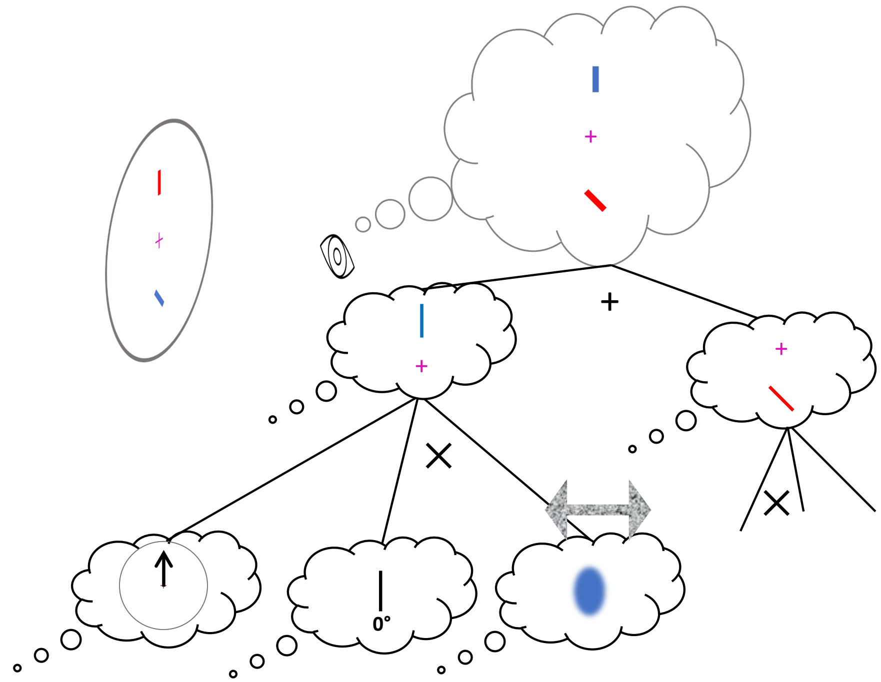
1. The Semantics of Mental Representations
2. Concatenating Fragments
3. **Integrating Features**
4. Integrating Contours
5. The Format of Contour Perception
6. The Whole and the Parts

Sub-atomic primitives.

*Treisman 1982; Matthen 2005; Clarke 2021

Structural Description
Illusory Conjunctions

- **Illusory conjunction:** Correct about features, incorrect about how they are co-instantiated.
- Explanation:
  - Representation of an item *x* as located at top, vertical, and blue consists of constituents
    - Representation of an item *y* as at top
    - Representation of an item *z* as vertical
    - Representation of an item *w* as blue
  - The import of combination is: the constituents must be accurate of the same item *(x=y=z=w)*.
  - Noise can lead to same feature representations being combined in different ways.
- Not all constituents are individually **well-formed** in the sense that they can be tokened in isolation.
  - Orientation and position receive distinct constituents.
  - But cannot represent orientation without combining with a position representation (see Quinlan 2003). These representations are **sub-atomic primitives**.
  - Constituents are not all self-standing building blocks
- No commitment to what the "binding mechanism" is (e.g. no commitment to FIT).



*cf. Clark 2004; Matthen 2014

Semantics of *Feature Integration*

Primitives

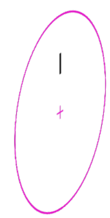
- Syntax:
  1. Position representations: POS = {**t, r, b, l**} ⊆ REPS
  2. Feature representations: FEAT = {**0, 45, 90, 145**} ⊆ REPS
- Semantics:
  1. Where μ is an structure-preserving mapping from psych dimensions to world dimensions (cf. Beck 2019):
     a. If **ρ** ∈ POS is a representation of *x* in δ, then it is perfectly accurate ⇔ *x* is located at position μ(**ρ**) in δ.
     b. If **ϴ** ∈ FEAT is a representation of *x* in δ, then it is perfectly accurate ⇔ *x* has orientation μ(**ϴ**)° in δ.

Feature Integration ×

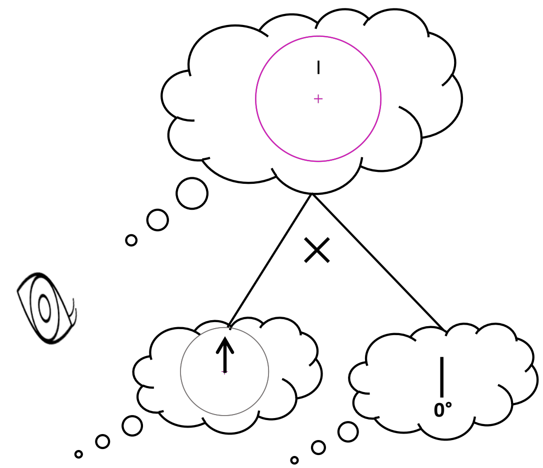
CONTOUR = set of contour representations (CONTOUR ⊆ REPS)

- Syntax: $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n$ ∈ CONTOUR, where $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n$ has the form, ⟨ $\boldsymbol{\alpha}_1$ … , $\boldsymbol{\alpha}_n$ ⟩, ⇔

  $\boldsymbol{\alpha}_1$ is a position representation ($\boldsymbol{\alpha}_1$ ∈ POS) and $\boldsymbol{\alpha}_{n>1}$ are feature representations ($\boldsymbol{\alpha}_{n>1}$ ∈ FEAT)
- Semantics: If $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n$ ∈ CONTOUR is a representation of *x* in δ, then it is perfectly accurate ⇔
  a. All its constituents $\boldsymbol{\alpha}_i$ are perfectly accurate in δ, and
  b. All its constituents represent the same thing: For all *y*, if there is a constituent $\boldsymbol{\alpha}_i$ that is a representation of *y* in δ, then *x*=*y*.



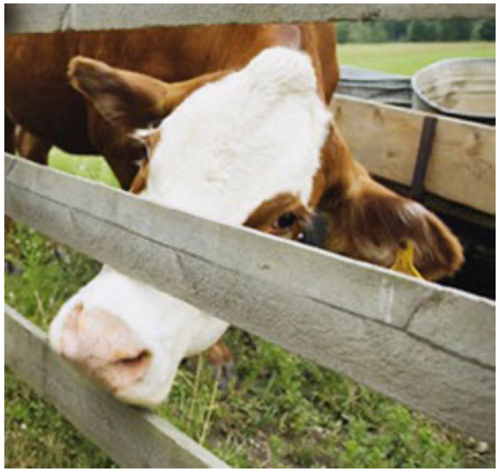
ς_1



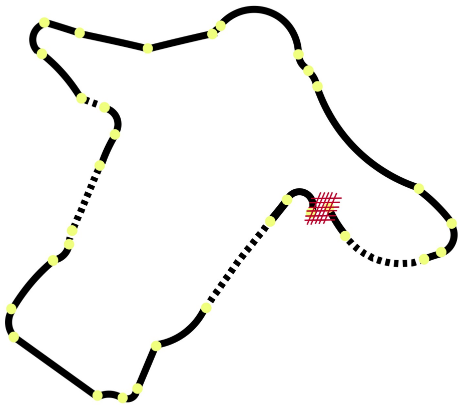
1. The Semantics of Mental Representations
2. Concatenating Fragments
3. Integrating Features
4. **Integrating Contours**
5. The Whole and the Parts
6. The Format of Contour Perception

From fragments to objects.

Structural Description



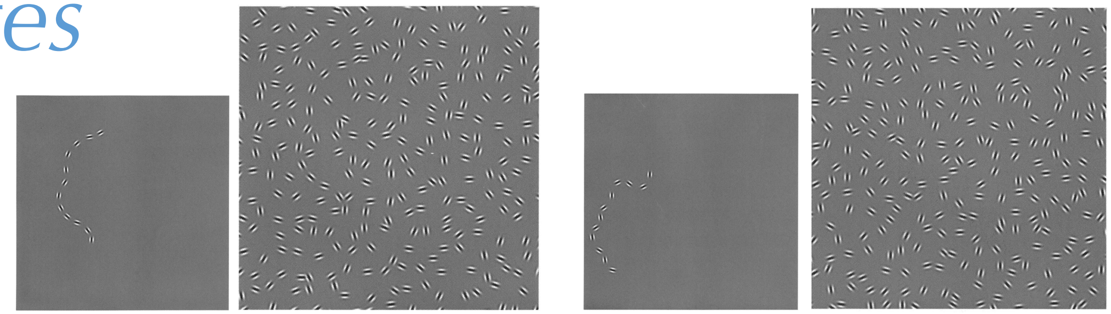
A



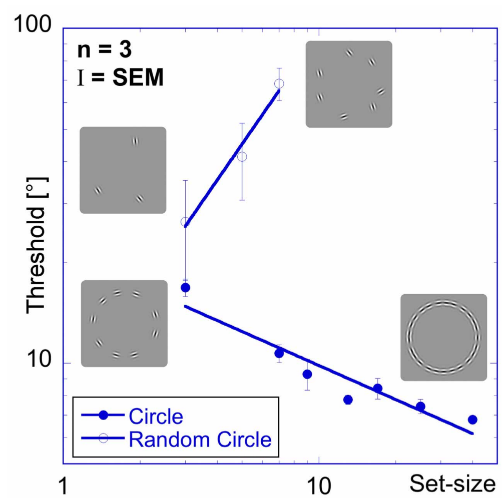
B

Structural Description:
Grouping Advantages

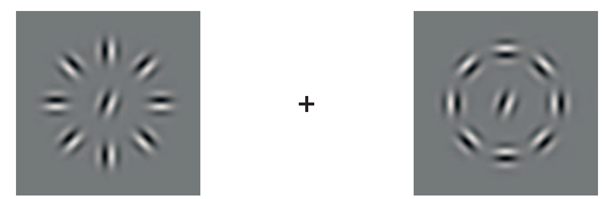
1. Less noisy, more accurate, faster to represent features within-contour than across-contour.



*Field, Hess, & Hayes 1993



*Kempgens, Loffler, & Orbach 2013

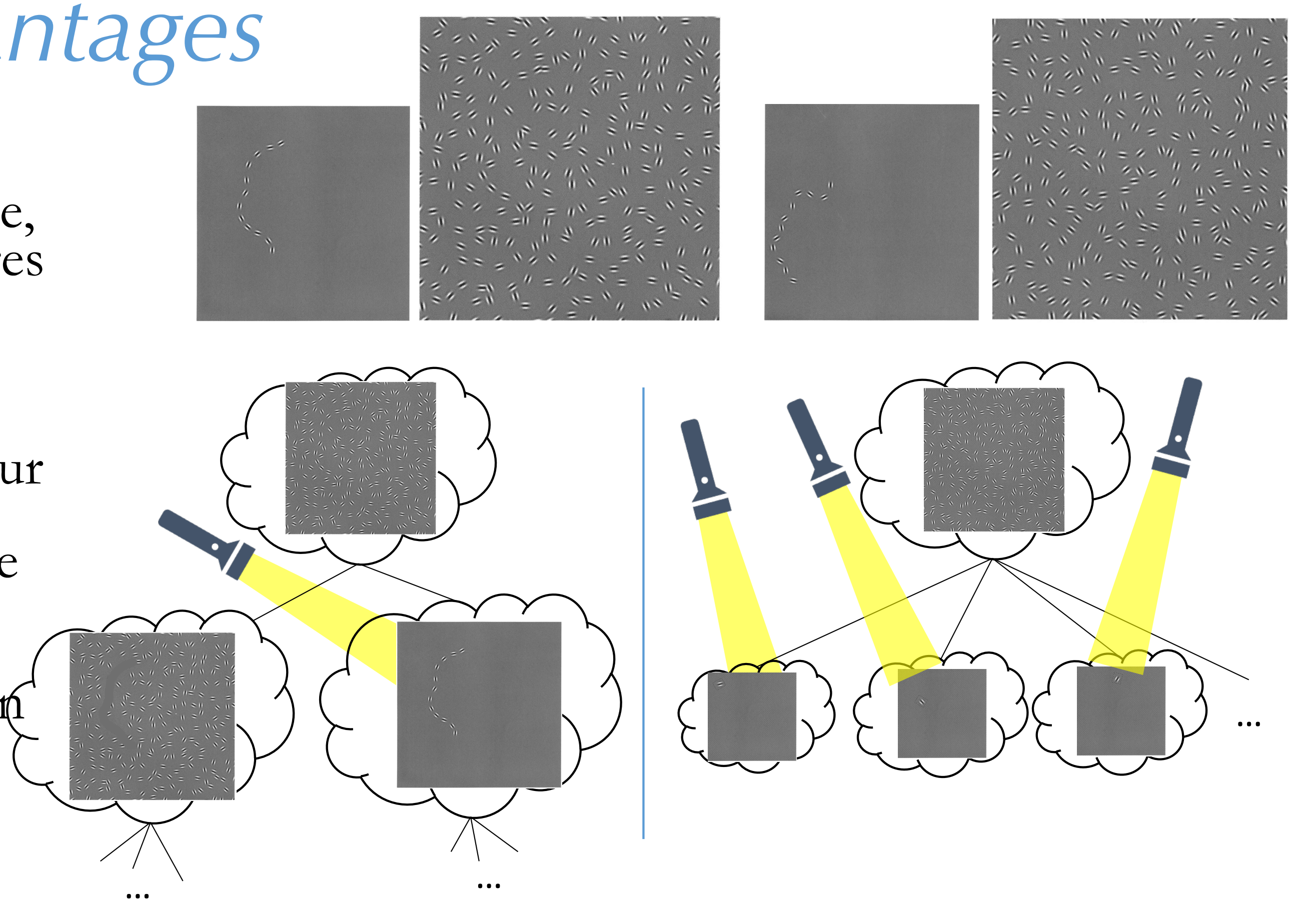


*Herzog et al., 2015; Livne & Sagi 2007

*Cf. Palmer 1977

Structural Description: *Grouping Advantages*

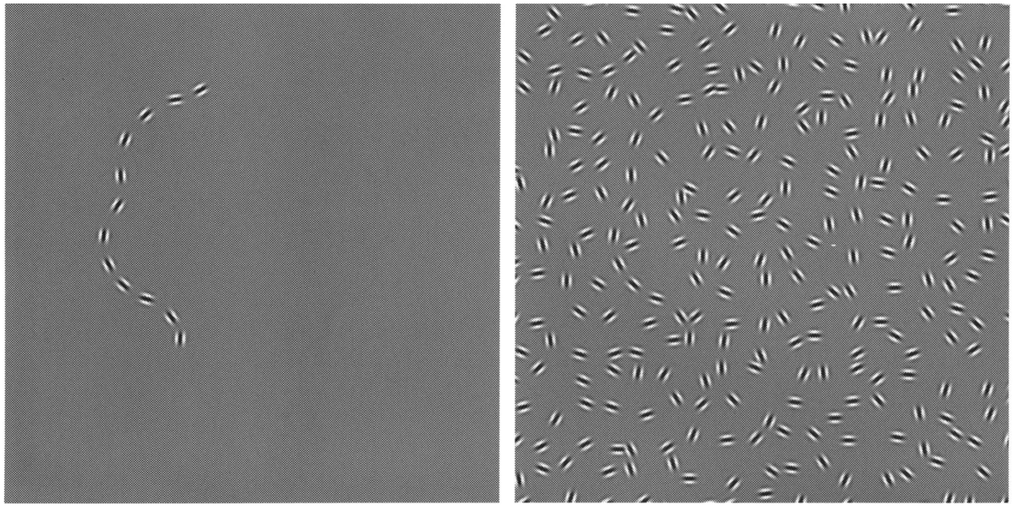
1. Less noisy, more accurate, faster to represent features within-contour than across-contour.
2. Explanation: representations of contour segments are combined into a unit, narrowing the search space and disentangling from variance in representation of other elements.



*Field et al. 1993; Baker, Garrigan, & Kellman 2020

Structural Description: *No whole without the parts*

1. Representation of contour requires representations of segments.
2. Explanation: representation of contour is composed from representations of segments.



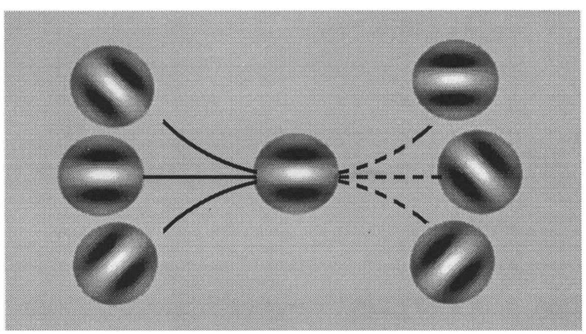
Structural Description: *Good continuation*

<u>Local binding via **association field**:</u>

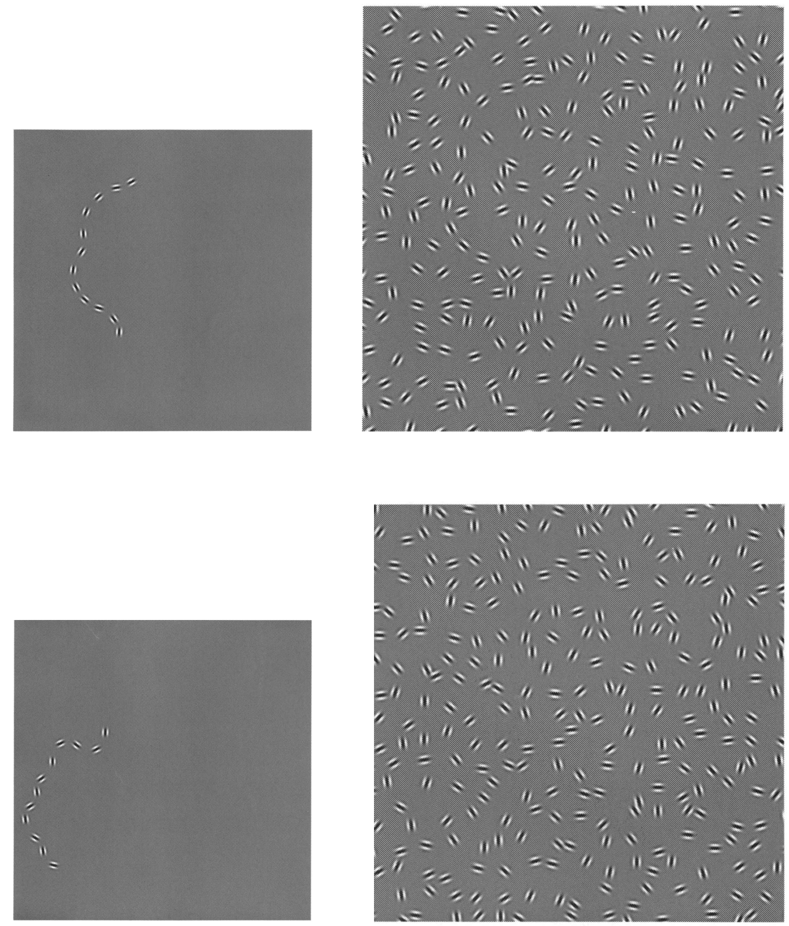
Representations of segments can combine only if

- Perceived positions and orientations of the fragments are aligned approximately along a smooth, non-inflected curve.

Can combine | Cannot combine



*Field et al. 1993



Structural Description: *Good continuation*

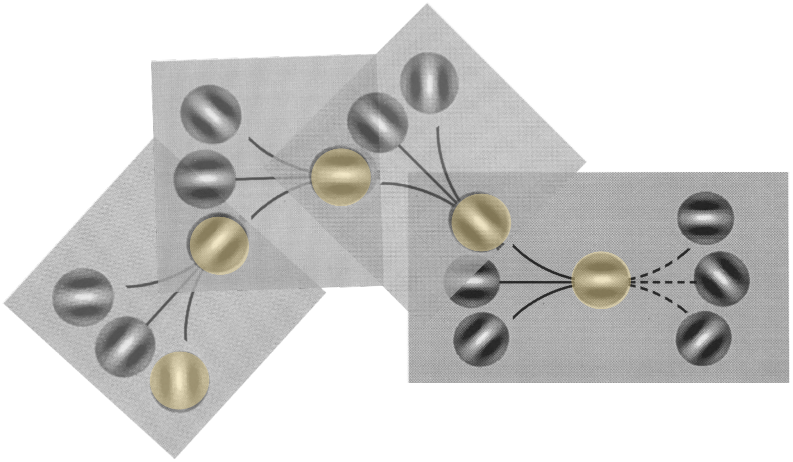
<u>Local binding via **association field**:</u>

Representations of segments can combine only if

- Perceived positions and orientations of the fragments are aligned approximately along a smooth, non-inflected curve.

<u>Global binding:</u>

- "a contour must be represented not as a set but as a sequence of local elements" (Elder & Goldberg 2002: 348).
- Markov property: A sequence of segment representations can combine if and only if each representation is combinable with the next in the sequence (Elder & Goldberg 2002).



*Elder & Goldberg 2002

Structure vs. metaphysics, semantics, inference

Not just metaphysics

- The point is not that coherent shapes necessarily consist of aligned segments.
- Rather: *representations* of coherent shapes can only be of contours with aligned segments.

Not just semantics

- The point is not just that a shape representation is accurate only if the segments of the represented contour are aligned.
- Rather: the degree to which shape representations can occur at all (whether accurately or inaccurately) depends on the alignment of the represented contours.

Not just process

- The point is not *merely* that alignment is an input cue or processing heuristic for deriving shape representations from the retinal image.
- A systematic constraint on the possibility of contour representations as a function of how their constituents are related.

Semantics of *Contour Integration*

Contour Integration ○

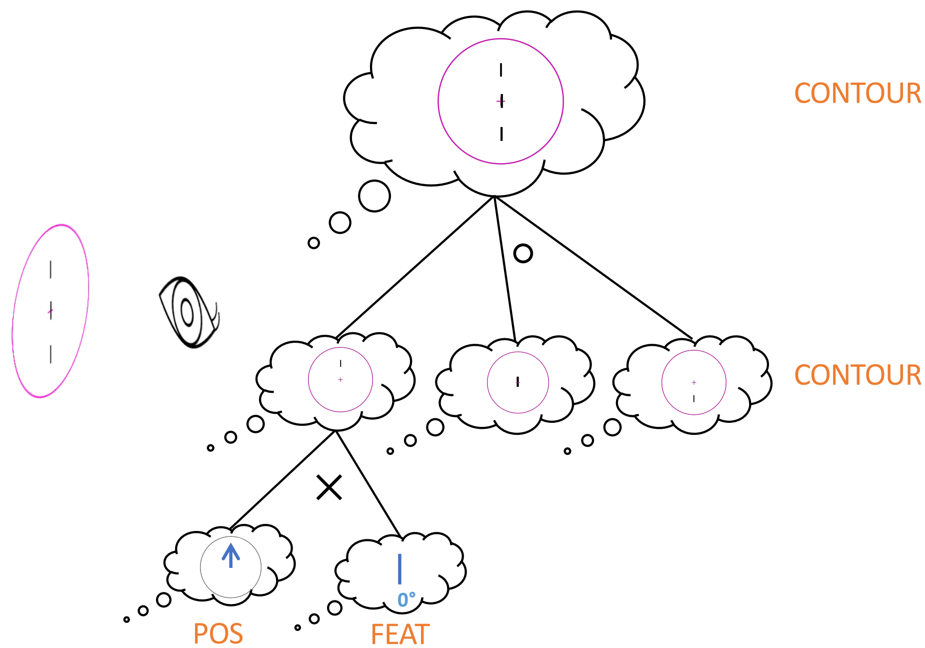
- Syntax†: $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n \in$ CONTOUR, where $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)$, $\Leftrightarrow$

  for all constituents $\boldsymbol{\alpha}_i$

  1. $\boldsymbol{\alpha}_i \in$ CONTOUR (all are contour representations)
  2. $\boldsymbol{\alpha}_1 \circ \boldsymbol{\alpha}_{i+1} \in$ CONTOUR (adjacent pairs are combinable), where
  3. $\boldsymbol{\alpha}_1 \circ \boldsymbol{\alpha}_{i+1} \in$ CONTOUR $\leftrightarrow$ *align*$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_{i+1})$ (pairwise combinability corresponds to alignment of represented positions and orientations)

- Semantics: If $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n \in$ CONTOUR is a representation of x in δ, then it is perfectly accurate $\Leftrightarrow$

  1. x is a contour, and
  2. Every constituent $\boldsymbol{\alpha}_i$ is perfectly accurate in δ, and
  3. For all y, if there is a constituent $\boldsymbol{\alpha}_i$ that is a representation of y in δ, then x is part of y.



†Syntactic rules may be *probabilistic* (Zhu & Mumford 2006). [But will save discussion for Q&A]

Semantics for a Fragment of Perception

REPS = set of all representations

Primitives:

- Syntax:
  1. Position representations: POS = $\{\mathbf{t}, \mathbf{r}, \mathbf{b}, \mathbf{l}\} \subseteq$ REPS
  2. Feature representations: FEAT = $\{\mathbf{0}, \mathbf{45}, \mathbf{90}, \mathbf{145}\} \subseteq$ REPS
- Semantics: Where μ is a structure-preserving mapping
  1. If $\boldsymbol{\rho} \in$ POS is a representation of x in δ, then it is perfectly accurate $\Leftrightarrow$ x is located at position $\mu(\boldsymbol{\rho})$ in δ.
  2. If $\boldsymbol{\theta} \in$ FEAT is a representation of x in δ, then it is perfectly accurate $\Leftrightarrow$ x has orientation $\mu(\boldsymbol{\theta})°$ in δ.

Feature Integration:

CONTOUR $\subseteq$ REPS

- Syntax: $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n \in$ CONTOUR, where $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n$ has the form, $\langle \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n \rangle$, $\Leftrightarrow$ $\boldsymbol{\alpha}_1$ is a position representation ($\boldsymbol{\alpha}_1 \in$ POS) and $\boldsymbol{\alpha}_{n>1}$ are feature representations ($\boldsymbol{\alpha}_{n>1} \in$ FEAT)
- Semantics: If $\boldsymbol{\alpha}_1 \times \ldots \times \boldsymbol{\alpha}_n \in$ CONTOUR is a representation of x in δ, then it is perfectly accurate $\Leftrightarrow$
  a. x is a contour
  b. All its constituents $\boldsymbol{\alpha}_i$ are perfectly accurate in δ, and
  c. For all y, if a constituent $\boldsymbol{\alpha}_i$ is a representation of y in δ, then x=y.

Contour Integration:

- Syntax: $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n \in$ CONTOUR, where $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)$, $\Leftrightarrow$ for every constituent $\boldsymbol{\alpha}_i$:
  1. $\boldsymbol{\alpha}_i \in$ CONTOUR (i.e. all constituents are contour representations)
  2. $\boldsymbol{\alpha}_1 \circ \boldsymbol{\alpha}_{i+1} \in$ CONTOUR (i.e. adjacent pairs of constituents are combinable), where
  3. $\boldsymbol{\alpha}_1 \circ \boldsymbol{\alpha}_{i+1} \in$ CONTOUR $\leftrightarrow$ *align*($\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_{i+1}$) (i.e. pairwise combinability corresponds to alignment of represented positions and orientations)
- Semantics: If $\boldsymbol{\alpha}_1 \circ \ldots \circ \boldsymbol{\alpha}_n \in$ CONTOUR is a representation of x in δ, then it is perfectly accurate $\Leftrightarrow$ for every constituent $\boldsymbol{\alpha}_i$:
  1. $\boldsymbol{\alpha}_i$ is perfectly accurate in δ, and
  2. If there is a y such that $\boldsymbol{\alpha}_i$ is a representation of y in δ, then x is part of y.

Concatenation:

- Syntax: $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n \in$ REP, where $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n\}$, IF

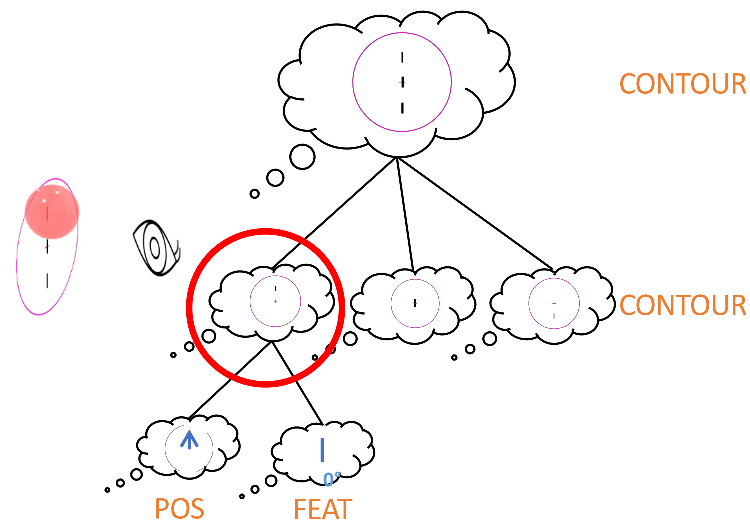
  $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n \in$ CONTOUR.
- Semantics: $\boldsymbol{\alpha}_1 + \ldots + \boldsymbol{\alpha}_n$ is perfectly accurate $\Leftrightarrow$ for every constituent $\boldsymbol{\alpha}_i$, $\boldsymbol{\alpha}_i$ is perfectly accurate in δ.

1. The Semantics of Mental Representations
2. Concatenating Fragments
3. Integrating Features
4. Integrating Contours
5. **The Format of Contour Perception**
6. The Whole and the Parts

The grammar of contours is unlike the grammar of languages.

Aspects of Iconicity

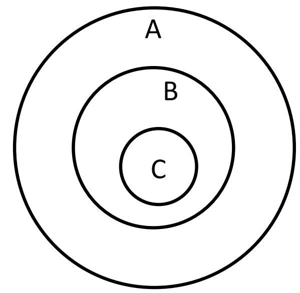
- **Mirroring** (Block 1983; Kulvicki 2014; Burge 2018)
  - Representational part-whole → worldly part-whole



*https://www.youtube.com/watch?v=E7dYXVLPd6Y

Aspects of Iconicity

- **Mirroring** (Block 1983; Kulvicki 2014; Burge 2018)
  - Representational part-whole → worldly part-whole
- **Constraint projection** (Shimojima 2001; Palmer 1978)
  - Combinatorial constraint → world constraints

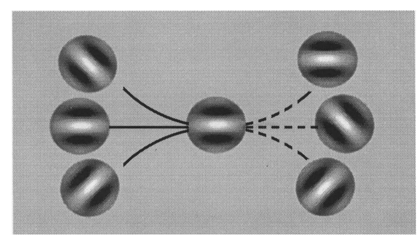


$C \subseteq B$
$B \subseteq A$
$C \subseteq A$

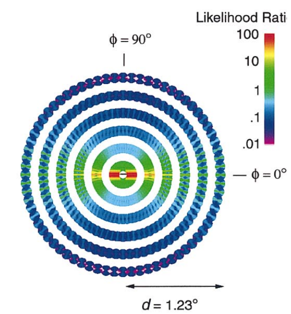
Representational Constraint
If C is contained in B and B is contained in A, then C is contained in A

World Constraint
If $C \subseteq B$, and $B \subseteq A$, then $C \subseteq A$

Representational Constraint



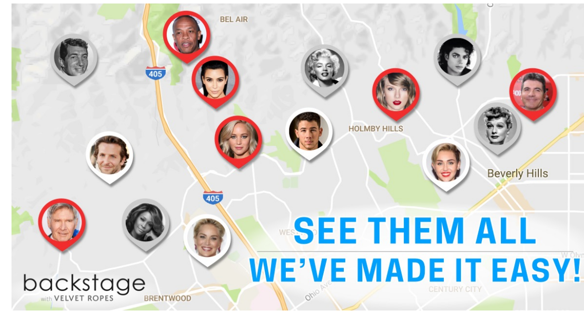
World Constraint

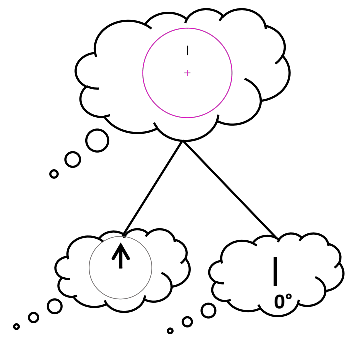


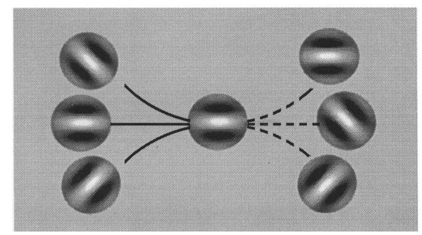
Geisler et al. 2002; cf. Elder & Goldberg 2002

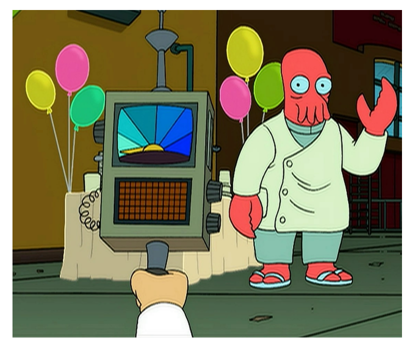
Aspects of Iconicity

- **Mirroring** (Block 1983; Kulvicki 2014; Burge 2018)
  - Representational part-whole → worldly part-whole
- **Constraint projection** (Shimojima 2001; Palmer 1978)
  - Combinatorial constraint → world constraints
- **Map-like** (Camp 2007; Rescorla 2009; Matthen 2014; Clarke 2021)
  - Features must be co-assigned with position.
  - Combination depends on representations of spatial properties.
- **Analog primitives** (Maley 2010; Beck 2019)
  - Primitives are related in psychological similarity space that corresponds to similarity of represented features









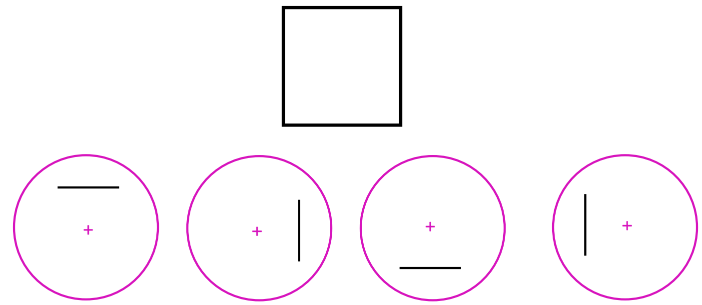
1. The Semantics of Mental Representations
2. Concatenating Fragments
3. Integrating Features
4. Integrating Contours
5. The Format of Contour Perception
6. **The Whole and the Parts**

"But… this can't possibly work, can it?!"

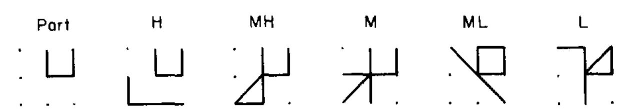
*Cf. Pomerantz & Cragin 2016; Hoffman & Richards 1984; Attneave 1968

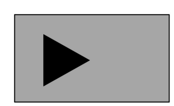
The whole is other than the sum of its parts

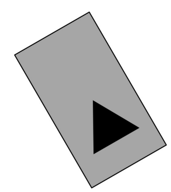
- Emergent contents: global features attributed to whole but not its parts, or attributed to an item in virtue of its being perceived as part of a whole.
  - Closure
  - Parthood
- Configural effects: the way an individual element is perceived depends on the way other elements are perceived.



> "There are wholes, the behavior of which is not determined by that of their individual elements, but where the part-processes are themselves determined by the intrinsic nature of the whole."
>
> -Max Wertheimer, "Gestalt Theory" (1924)

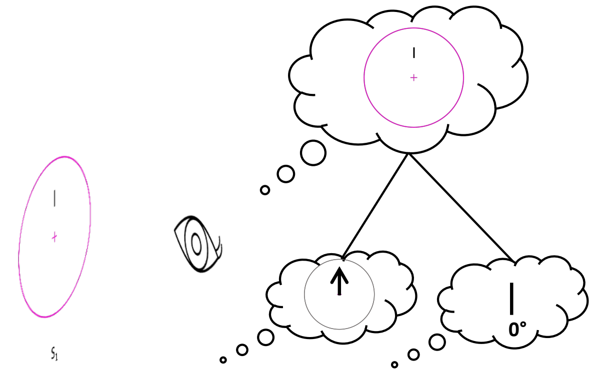


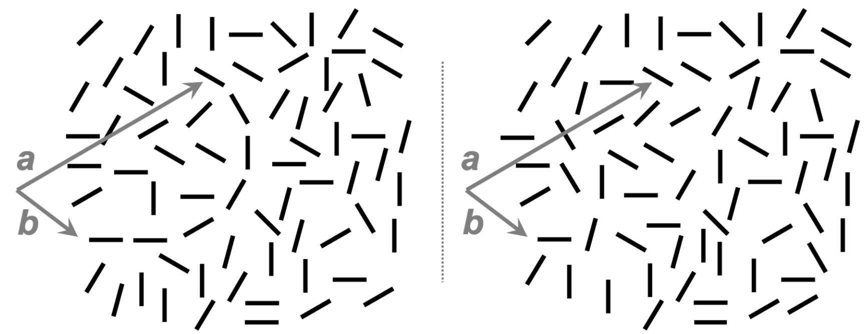


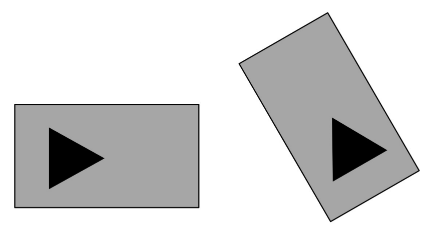


Explaining holistic effects

- **Emergent contents** through semantic composition
  - Feature integration introduces content about co-instantiation.
  - Contour integration introduces content about parthood.
- **Configural effects**
  - *Structural dependencies*: **a** and **b** combine only if sequence of pairwise combinable representations connects them.
  - *Semantic content*: representation of orientation *relative to x*.
- **Recurrent processing**
  - Representation has a compositional code $\nRightarrow$ representation is produced in purely bottom-up, feedforward, piecemeal way.
  - E.g., "analysis-by-synthesis" models in which bottom-up "proposals" about constituents are revised in light of scene context, given prior models (Yuille 2006).

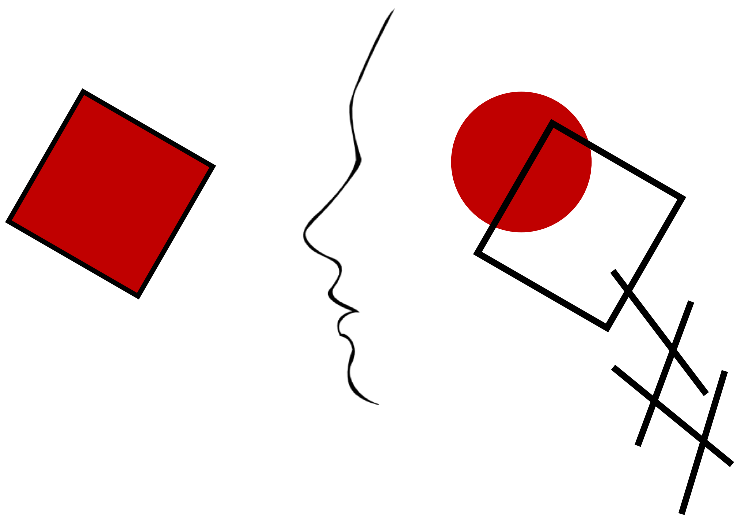




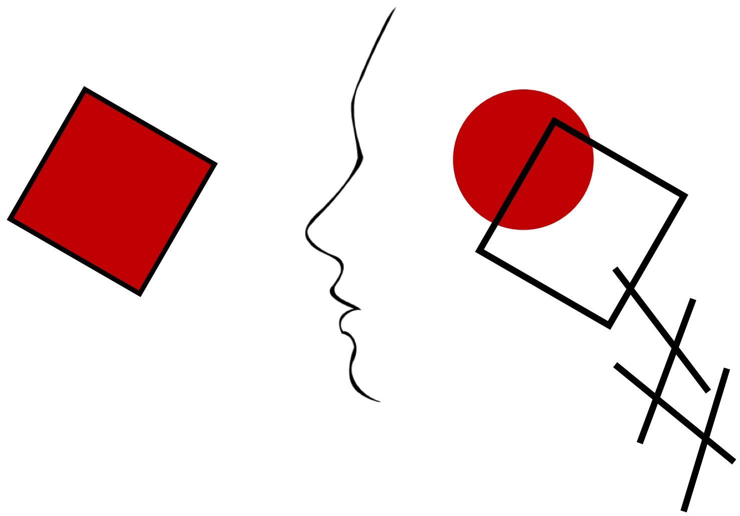


Summary

1. A fully explicit compositional syntax and semantics can be developed for visual contour representations.
   a) Concatenation
   b) Feature integration
   c) Contour integration
2. Structural and semantic claims are empirically constrained.
   a) Semantic claims explain why distal feature predicts responses better than independently identifiable proximal features.
   b) Structural claims explain distribution of representations.
3. The syntax and semantics of contour representations is quite different from that of languages, and has core aspects of iconicity.
4. The compositional program is consistent with the Gestalt maxim that the whole is other than the sum of the parts.



Thank you!



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

Rosa's Comments

1. "Could any findings in neuroscience show that in fact we were wrong about whether these things are represented in particular ways?"
2. "What is the goal of your systematic semantic analysis? How exactly does having the formalism help (and who does it help)?"
3. "Do we need compositionality for generativity?"
4. "Your semantic theory seems to require that visual representations are discrete"
5. "How does your theory help to illustrate how semantic theories can have empirical content?"

1. Neuroscience

- Neuroscientific results can inform structural and semantic claims. But doubt that we're in a position to make strong inferences from single experiments or paradigms in neuro. We're still figuring out what kinds of inferences are cogent.
  - E.g., "If you can't decode the contour from neural activity, then there must not be content about contours."
    - Assuming our assumptions about the decoder reflect how the brain works.
    - Assuming we're looking in the right place, at the right level.
    - Assuming we're using the right stimulus set.
    - …
  - E.g., "Representations of the segments are integrated only if the neural correlates are co-active."
    - Assuming we've identified the relevant sense of "co-active" (Synchrony? Response enhancement? Labeled lines?)
- There's a long line of research in neurophys. that is consistent with the contour integration research (Hubel & Wiesel, Pasupathy, etc.)
  - In this line, there is somewhat smooth crossing between levels, both in experiment and in computational modeling.
  - Vaziri and Pasupathy 2009: "during the same period in which computational vision became nonstructural, biological vision was discovered to be emphatically structural."
- But a lot of the recent work on DCNNs (e.g. DiCarlo lab) is challenging that kind of cross-scale consilience. It may be that the neurophysiology winds up looking so different from the psychophysically grounded models that there's no productive way to cross scales btwn psychophysics inspired structural models and DCNN models of neurophys. In which case…
  - But maybe not—growing recognition of need to introduce structured representations into DCNNs, even at level of contour perception?

2. What's the goal?

- **Philosophy of representation**: semantics provides an explicit, systematic theory of representation – i.e. how contents are structurally encoded.
  - The psych theories I'm interested in are *representational* theories—in that they invoke structured reps—but they rarely give a *theory of representation*. They may not need to for their aims.
  - I'm interested in the nature of representation/intentionality. How is it possible? What are the varieties of representation in nature? How should we conceptualize how research on different perceptual capacities will come together to account for content about rich, complex scenes?
- **Philosophy of science:**
  - Explicating aspects of scientific models, explanations, and inferences and how they fit together.
  - What sorts of tests probe content? What sorts of tests probe structure? Implicit knowledge about what tests to use and why hasn't been systematically explicated.
- Philosophy of perception: Epistemology:
  - How does perceptual warrant factor and transmit?
- Content-grounding:
  - A "context principle" (cf. Frege): any account of what makes a perceptual representation represent what it does must take into account how that state systematically contributes to other perceptual representations.
- Benefits to science?
  - Might be useful as part of the scientist's broader perspective on what they're doing and how everything might hang together (some nice examples of this: Phil Kellman and Nick Baker, Martin Arguin, Jacob Feldman, Odelia Schwartz, Song-Chun Zhu).

3. Generativity

- I'm sympathetic to general arguments from productivity/systematicity.
- But the conclusion is highly underspecified—that the system has a compositional scheme, but not what the scheme is like.
- I often find debates over these arguments boring and uninspiring.
- What's interesting to me is to see whether we can get any purchase on the work of actually describing the scheme in an empirically constrained way.

4. Discreteness

- I assume the primitives are analog and maybe continuous within a dimension.
- But I think the number of dimensions is discrete and finite. Perceptual feature space is not infinite-dimensional, no less continuum-dimensional.
- Poses an interesting question: why does a productivity argument require finite dimensions but not finite values along those dimensions?

5. Empirical content

- Uncharitable gloss on a lot of past philosophical lit: filled with metaphor (the "shape" of a representation), descriptions that are only hypothetically meaningful ("Mentalese"), worship of FOL, ambitions to give grand canonical form of perception, empty appeals to neuro (syntax is a "neural property"), and in the better cases one-off appeals to experimental findings.
- Can be forgiven for thinking that there's not much empirical content to the discussions. But even those who think that there is are trying to explicate where that content lies.
- My view: There are more or less systematic empirical constraints, some targeting structural descriptions and others targeting semantic attributions.
  - These constraints are more often implicit than explicit in psychological lit
- So it's not just a matter of being sensitive to experimental results, but trying to sort out in what ways the theory can/should be sensitive to results.