Target-nontarget relations determine ‘bottom-up’ feature priming and ‘top-down’ contingent capture

Stefanie I. Becker¹ & Ulrich Ansorge²

¹ The University of Queensland, Australia.
² University of Vienna, Austria.
Architecture of Attention Theories:

*Visual Selection:* Function of top-down tuning + bottom-up factors (e.g., feature contrast).

**Top-down tuning consists in activating/inhibiting specific Feature Maps =**

Topographically organised neurons that respond to specific features

(e.g., Lee et al., 1999; Maunsell & Treue, 2006; Navalpakkam & Itti, 2007; Martinez-Trujillo & Treue, 2004; Scolari & Serences, 2010; Treisman & Sato, 1990; Wolfe, 1994)
**Architecture of Attention Theories:**

*Visual Selection:* Function of top-down tuning + bottom-up factors (e.g., feature contrast).

**Top-down tuning consists in activating/inhibiting specific Feature Maps =**

Topographically organised neurons that respond to specific features

**Problem:**

We need one topographical feature map for each feature that we can top-down select.

Top-down selection is fine-grained (e.g., Navalpakkam & Itti, 2007).

→ Ontological ghetto of feature maps

(e.g., Lee et al., 1999; Maunsell & Treue, 2006; Navalpakkam & Itti, 2007; Martinez-Trujillo & Treue, 2004; Scolari & Serences, 2010; Treisman & Sato, 1990; Wolfe, 1994)
Alternative: Relational Theory (Becker, 2010; JEP-Gen.)

- Attention not top-down tuned to particular feature values, but:
  - to *relationships* = properties that target has in relation to other irrelevant items that constitute the context (redder, larger, darker...).
  - e.g., when searching for a person with an orange shirt, it depends on the context whether attention is tuned ...

  to *redder* ... or *yellower*.

Advantage: Avoids ontological ghetto of feature maps.
Previous Studies & Current Study

- Some evidence for top-down tuning to relations:
  - Attention can be tuned to target-nontarget relationships, regardless of feature values of the target or the nontargets (Becker, 2008, *JEP-HPP*; Becker 2010, *JEP-Gen*).
  - When target has constant feature value & context is constant, we do not tune attention to target feature or against nontarget features, but to target-nontarget relations (Becker, Folk & Remington, 2010, *JEP-HPP*, & in press, *Psych Science*).

- BUT: No decisive evidence against feature-based theories of attention.

- To show that feature maps don’t exist, or that they don’t guide attention, it needs to be shown that we *cannot* tune attention to a specific feature or range of similar features.

- Never been tested.

- Present Study: *Can* we tune attention to a feature value?
Can we tune to a specific color?

Task: Make a fast eye movement to color target (specified) & respond to x/o. Colors of target and nontargets vary across trials; differently in two blocks:

**Target-Same Condition:**
*Can we tune to orange?*

Target always orange / relations vary: target redder or yellower.

**Relation-Same Condition:**
*Can we tune to redder?*

Target always redder / features vary: target or nontargets orange.
Repeat Trials As A Baseline

Colors can randomly repeat or change → *repetition trials serve as baseline.*
Central comparison: Repeat trials (n-1 same) vs. Change trials (n-1 diff).

**Target-Same Condition:**

```
  o  x  o  o  n-1 same
  o  +  o  x  n-1 diff
  o  x  o  x  n-1 same
  o  x  o  x  n-1 diff
```

**Relation-Same Condition:**

```
  x  +  o  o  n-1 diff
  o  x  o  x  n-1 same
  o  x  o  x  n-1 diff
  o  x  o  x  n-1 same
```
Predictions: Target-Same Condition

Colors can randomly repeat or change \(\rightarrow\) *repetition trials serve as baseline*. Central comparison: Repeat trials (n-1 same) vs. Change trials (n-1 diff).

**Target-Same Condition:**

**Feature-based:**

**Relational:**

![Graphs showing proportion of first fixation on target for n-1 same and n-1 diff conditions for feature-based and relational predictions.](image-url)
Predictions: Relation Same-Condition

Colors can randomly repeat or change → repetition trials serve as baseline.
Central comparison: Repeat trials (n-1 same) vs. Change trials (n-1 diff).

Relation-Same Condition:

Feature-based:

Prop. 1st Fix on Target

<table>
<thead>
<tr>
<th></th>
<th>n-1 same</th>
<th>n-1 diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Relational:

Prop. 1st Fix on Target

<table>
<thead>
<tr>
<th></th>
<th>n-1 same</th>
<th>n-1 diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results

Proportion of 1st Target Fixations

Target Same

Relation Same

Latencies: 242ms / 255ms **

235ms / 245ms ns
Can we tune to a particular shape?

Task: Make a fast eye movement to the shape target (specified prior to block).

**Target always the 10-spike star; Relation varies: target spikier vs. less spiky**

- **Target same**
  - n-1 diff
  - n-1 same
  - n-1 diff
  - n-1 same

**Target always less spiky; Features vary: target or nontarget can have 10 spikes**

- **Relation same**
  - n-1 diff
  - n-1 same
  - n-1 diff
  - n-1 same
Results

Proportion of 1st Target Fixations

Target Same

Relation Same

Latencies: 280ms / 279ms

278ms / 286ms
First Interpretation

No evidence that attention can be tuned to specific feature value.

(???)

→ Maybe feature-specific setting acquired later in block, through learning?
Feature-specific Tuning through Training?

No evidence that attention can be tuned to specific feature value. (???)

Proportion 1st Target Fixations in 1st, 2nd & 3rd 50 trials

1.) Color: Target Same

2.) Star Shapes
3.) Geometrical Shapes

4.) Element Shapes

6.) Size
Conclusions

- When target shape remains constant, we *can learn* to tune to particular target feature.

- But: More commonly relational search found; practice does not *always* lead to feature search & is never perfect.

- Fact that search was always initially relational and only later became feature-specific speaks against feature map concept.

- Important: Feature-based theories cannot explain good performance in Relation-Same condition where feature mapping was inconsistent.

→ Feature-based theories need to be modified, or replaced by relational account.
What guides visual attention & eye movements?

**Bottom-Up:**

- Salient Regions are **inspected** more frequently

- **Saliency defined as feature contrast in dimensions of:**
  - (a) Color
  - (b) Intensity / Brightness
  - (c) Orientation

(Itti, Koch & Niebur, 1998)
Aims: Clarify roles of
(1) Priming
(2) Saliency
(3) Novelty

in different tasks of
(A) scene recognition
(B) free viewing

Hypothesis: Whenever a task requires re-orienting within visual scenes, repeated visual input prioritized for selection.

Visual input that repeats across images of natural scenes should attract attention because it is diagnostic of whether a scene has been seen before.

1. In free viewing, re-inspection of familiar scenes may show bias for salient or novel regions.

2. BUT: In scene recognition, gaze will be biased to previously viewed regions, features or feature differences (priming).
Priming in Scenes

Method
Priming in Scenes

**Method**

- **Learning Block**
  - 90 encoding trials
- **Transfer Block**
  - 120 retrieval trials
<table>
<thead>
<tr>
<th>condition</th>
<th>identical</th>
<th>view change</th>
<th>mirrored</th>
<th>new scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>stimulus in learning block</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td>(not presented)</td>
</tr>
<tr>
<td>stimulus in transfer block</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>response</td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
</tr>
</tbody>
</table>
- Priming: Is there a recognition-specific gaze preference for repeated over salient or novel information?

- Two alternative tasks:

  (1) Recognition
  ("study the images for later recognition")

  (2) (repeated) free-viewing
  ("just look at the following images")
Primed: There is a recognition-specific gaze preference for repeated over salient or novel information.

- Recognition task:
  - Overall fixation duration is higher in repeated image regions.
  - Fixations are more frequent in repeated image regions.

- Free-viewing task:
  No differences between repeated and new image regions.
Primed in Scenes

Method

- Priming: Does feature priming (i.e. looked-at information) facilitate recognition?

  - Familiar images – fixated details

  ![Images of familiar scenes with fixated details]

  - Familiar images – details not fixated

  ![Images of familiar scenes without fixated details]
- Priming: Does feature priming (i.e. looked-at information) facilitate recognition?

- Unfamiliar images – salient details

- Unfamiliar images – randomly selected details
Priming: Looking at familiar cutouts improves scene recognition
- Priming: Looking at familiar cutouts improves scene recognition

![Graph showing discrimination performance (%) for different types of image detail. The graph compares fixated, not fixated, salient, and random types of image detail. The performance is shown for familiar and unfamiliar images, with a chance level of 50% represented by a dotted line.](image)
Scene Recognition depends on Priming

Main conclusion: Priming during natural scene viewing characteristic of recognition tasks.

Recognition model based on feature contrasts – why does it work?

Signed/directed feature contrasts provide invariant properties across varying illumination conditions.

Signed/directed feature contrasts are therefore helpful in orienting across views.
Summary: What do we know?

Strong evidence for relational tuning with inconsistent feature mappings =
Feature map theories cannot account for relational effects
→ Feature map theories have to be modified. Or replaced by relational theory.

On a broader scale

Neuronal coding is sparse or efficient and adapted to natural vision

Relational tuning is therefore likely: it yields invariance during natural vision,
e.g. invariance of local feature contrasts across different illumination
conditions and perspectives
Thank You!
Sponsored by

VIENNA SCIENCE
AND TECHNOLOGY FUND