Perception: Psychophysics and Modeling

12 | Visual Attention

Felix Wichmann

Neural Information Processing Group
Eberhard Karls Universität Tübingen

Reading

Goldstein, E. B. (2007). Visual Attention. In: Sensation and Perception, ch. 6, pp. 133-152. Wadsworth Publishing, Wadsworth, CA, USA, 8th edition.

Allport, D.A. (1993). Attention and Control. Have we been asking the wrong questions? A critical review of twenty-five years. In David E. Meyer & Sylvan Kornblum (eds.), *Attention and Performance XIV*. The MIT Press. 183-218.

Petersen, S. E. and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35:73–89.

Posner, M. I. (2012). Attentional networks and consciousness. *Frontiers in Psychology*, 3:1–4.

Rosenholtz, R., Huang, J., and Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology*, 3(13):1–15.

Limits of object recognition

The human visual system is very good at object recognition. However, we cannot recognise multiple objects simultaneously.

If we suppose that object recognition is a primary goal of the visual system, then additional mechanisms are required for it to be generally useful in complex scenes:

- Scene perception (gist)
- Attention (selection mechanisms more generally)
- Saliency (overt attention in scenes)

''Everyone knows what attention is.''

— William James (1890)

''It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatter-brained state which in French is called *distraction*, and *Zerstreutheit* in German.''

... the many meanings of attention ...

- *Attention* is the mental process involved in attending to other objects.
- *Attention seeking*, behavior which is intended to attract attention from others
- *Attention span*, the amount of time a person can concentrate on a single activity
- *Selective attention*, the process by which only a subset of stimuli received by our sensory organs are selected to enter the consciousness
- *Attentional bias*, a type of cognitive bias
- *Attention-deficit hyperactivity disorder*, a developmental disorder
- *Adult attention-deficit disorder*, attention-deficit hyperactivity disorder present in adulthood
- *Continuous partial attention*, is constantly paying attention to different things, but only partially
- *Inattentional blindness*, is the phenomenon of not being able to see things (as of a cognitive filtering process) that are actually there.

from: http://en.wikipedia.org/wiki/Attention_(disambiguation)

... and further ...

Attention is the behavioral and cognitive process of selectively concentrating on a discrete aspect of information, whether deemed subjective or objective, while ignoring other perceivable information. Attention has also been referred to as the allocation of limited processing resources.[1]

Attention remains a major area of investigation within education, psychology, neuroscience, cognitive neuroscience, and neuropsychology. Areas of active investigation involve determining the source of the sensory cues and signals that generate attention, the effects of these sensory cues and signals on the tuning properties of sensory neurons, and the relationship between attention and other behavioral and cognitive processes like working memory and vigilance. A relatively new body of research, which expands upon earlier research within neuropsychology, is investigating the diagnostic symptoms associated with traumatic brain injuries and their effects on attention. Attention also has variational differences among differing cultures.[2]

... all pretty confused ... is attention something that needs explaining, or does attention itself explain certain behaviours? Is attention a "resultant" or a "force?"

[1],[2] from: http://en.wikipedia.org/wiki/Attention

Overview

Terminology

Selection in Space

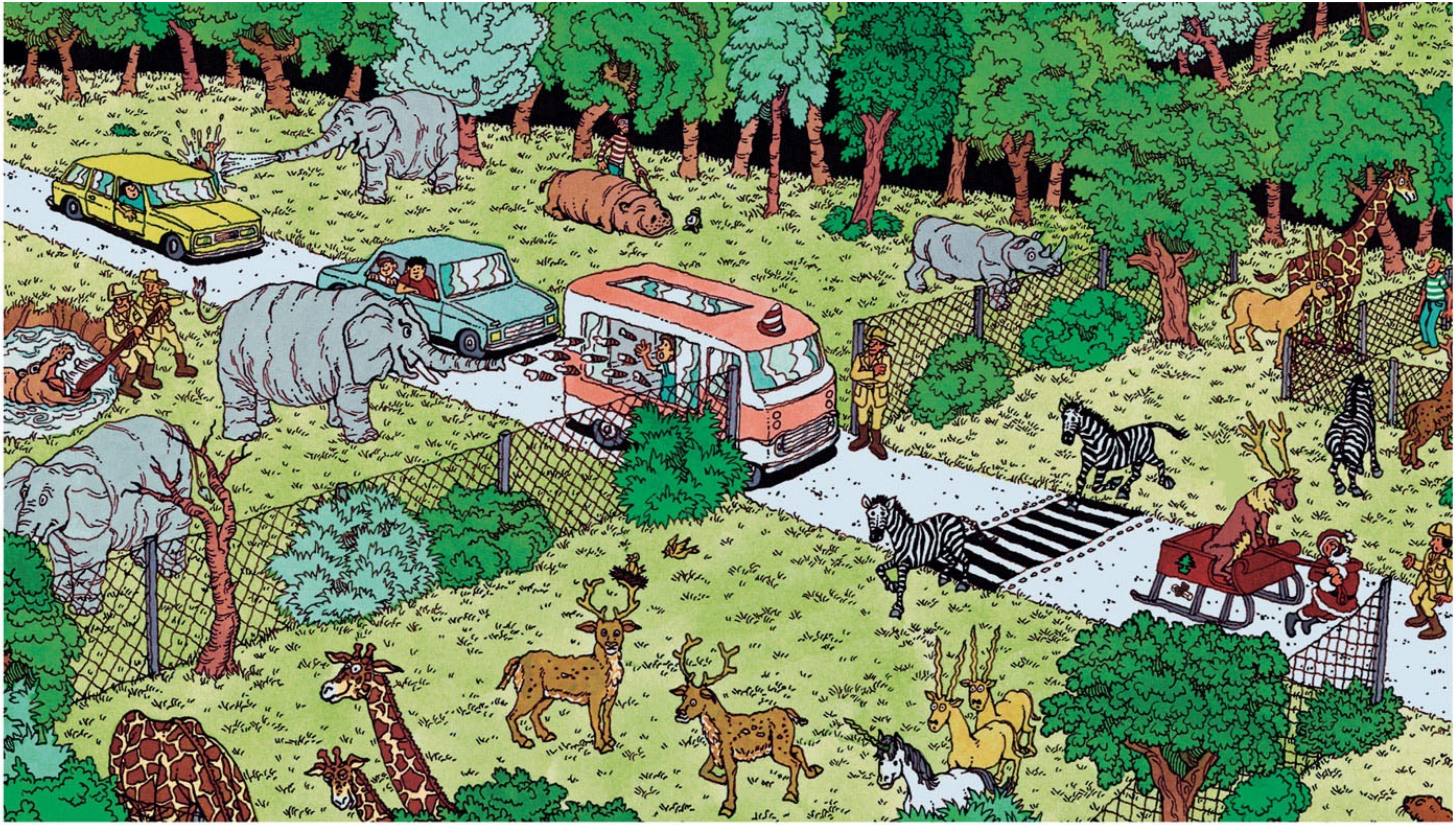
Visual Search

Terminology

Attention: Any of the very large set of selective processes in the brain.

To deal with the impossibility of handling all inputs at once, the nervous system has evolved mechanisms that are able to restrict processing to a subset of things, places, ideas, or moments in time.

Selective attention: The form of attention involved when processing is restricted to a subset of the possible stimuli.



***SENSATION & PERCEPTION 4e*, Figure 7.2**

Varieties of Attention

External: Attending to stimuli in the world

Internal: Attending to one line of thought over another or selecting one response over another

Overt: Directing a sense organ toward a stimulus, like turning your eyes or your head

Covert: Attending without moving eyes / head / body

Divided: Splitting attention between two different stimuli

Sustained: Continuously monitoring some stimulus

We simply cannot read the left-hand and right-hand sentences at the same time (limited divided attention)

The letters are large enough to resolve them while looking at the "x" in the middle, we cannot read the left-hand and the right-hand side simultaneously.

(Note: You can read the words on both sides of the "x" but you cannot comprehend the two sentences simultaneously!)

These	x	Is
letters	x	it
are	x	time
big	x	for
and	x	a
easy	x	quick
to	x	snack
read.	x	yet?

***SENSATION & PERCEPTION* 4e, Figure 7.1**

Selective attention: alerting and orienting

Alerting: Maintaining vigilance, readiness to respond, arousal, attention to when a target will occur.

Orienting: Selecting a location or modality to attend.

Alerting and orienting are thought to be subserved by separate brain networks.

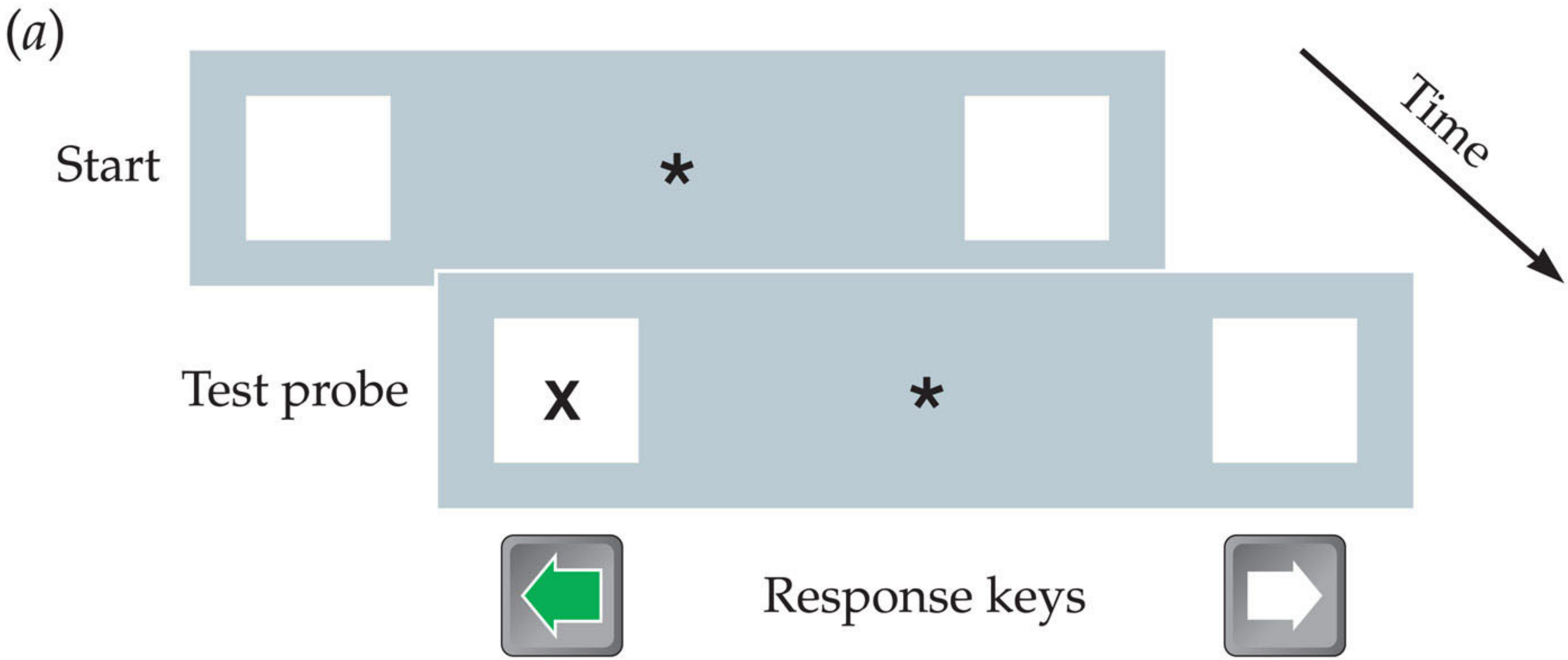
Brain regions involved in orienting to visual stimuli are also involved in orienting to stimuli in other modalities.

Orienting to a location can provide better information access from all modalities at that location, not only from the expected modality (Driver et al., 2004)

The most common subject of study: selection in space

Selective attention to spatial locations (orienting) is one of the most studied domains of attention.

Classic paradigm: Posner's cueing paradigm. Detect a simple target as quickly as possible.



SENSATION & PERCEPTION 4e, Figure 7.3 (Part 1)

The most common subject of study: selection in space

Reaction time (RT): A measure of the time from the onset of a stimulus to a response.

Cue: A stimulus that might indicate where (or what) a subsequent stimulus will be.

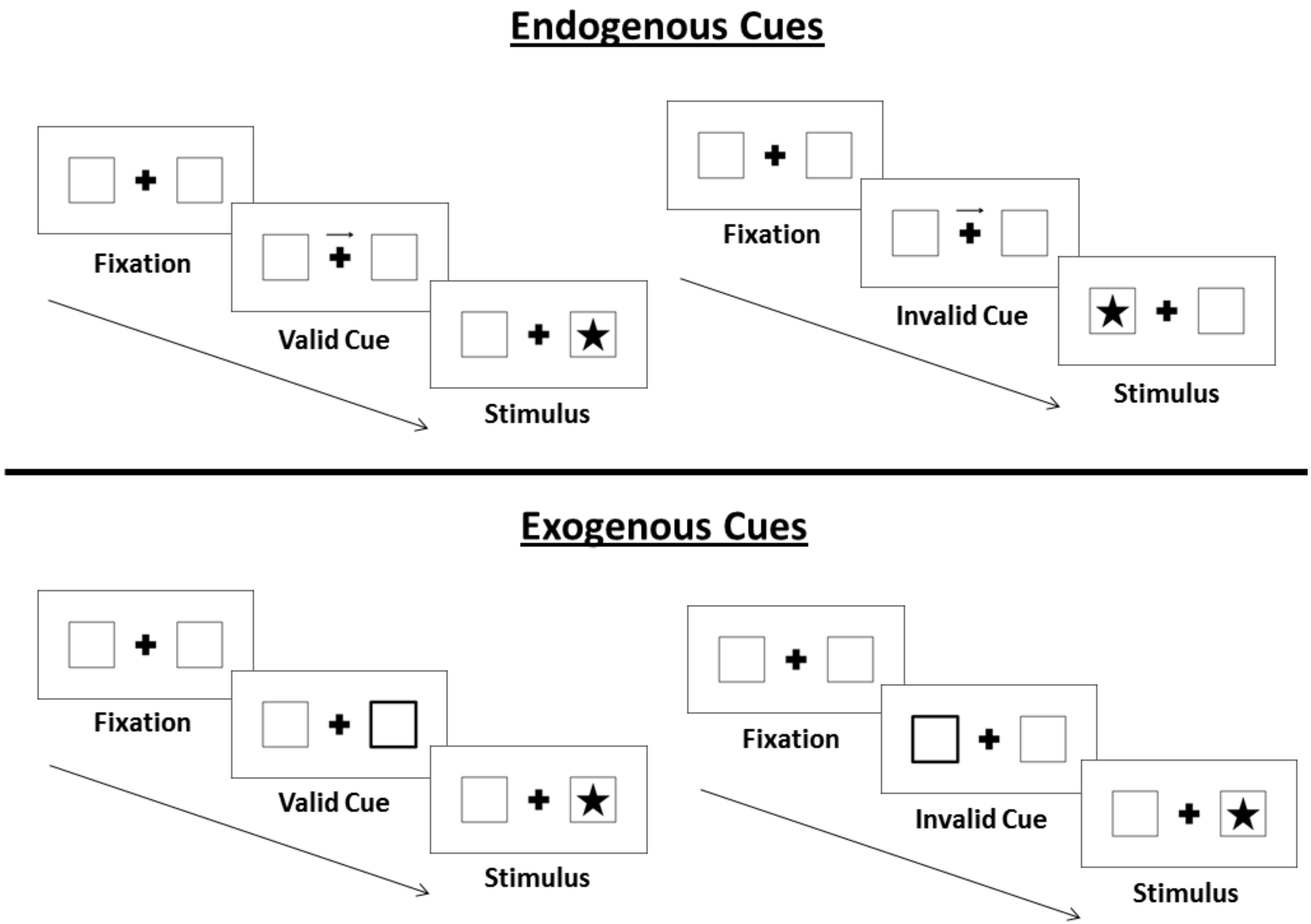
Cues can be valid (correct information), invalid (incorrect), or neutral (uninformative).

Exogenous cue: In directing attention, an exogenous cue is located out (exo) at the desired final destination of attention.

Endogenous cue: In directing attention, an endogenous cue is located in (endo) or near the current location of attention. An instruction that can be voluntarily obeyed.

Stimulus onset asynchrony (SOA): The time between the onset of one stimulus and the onset of another.

Posner cueing paradigm



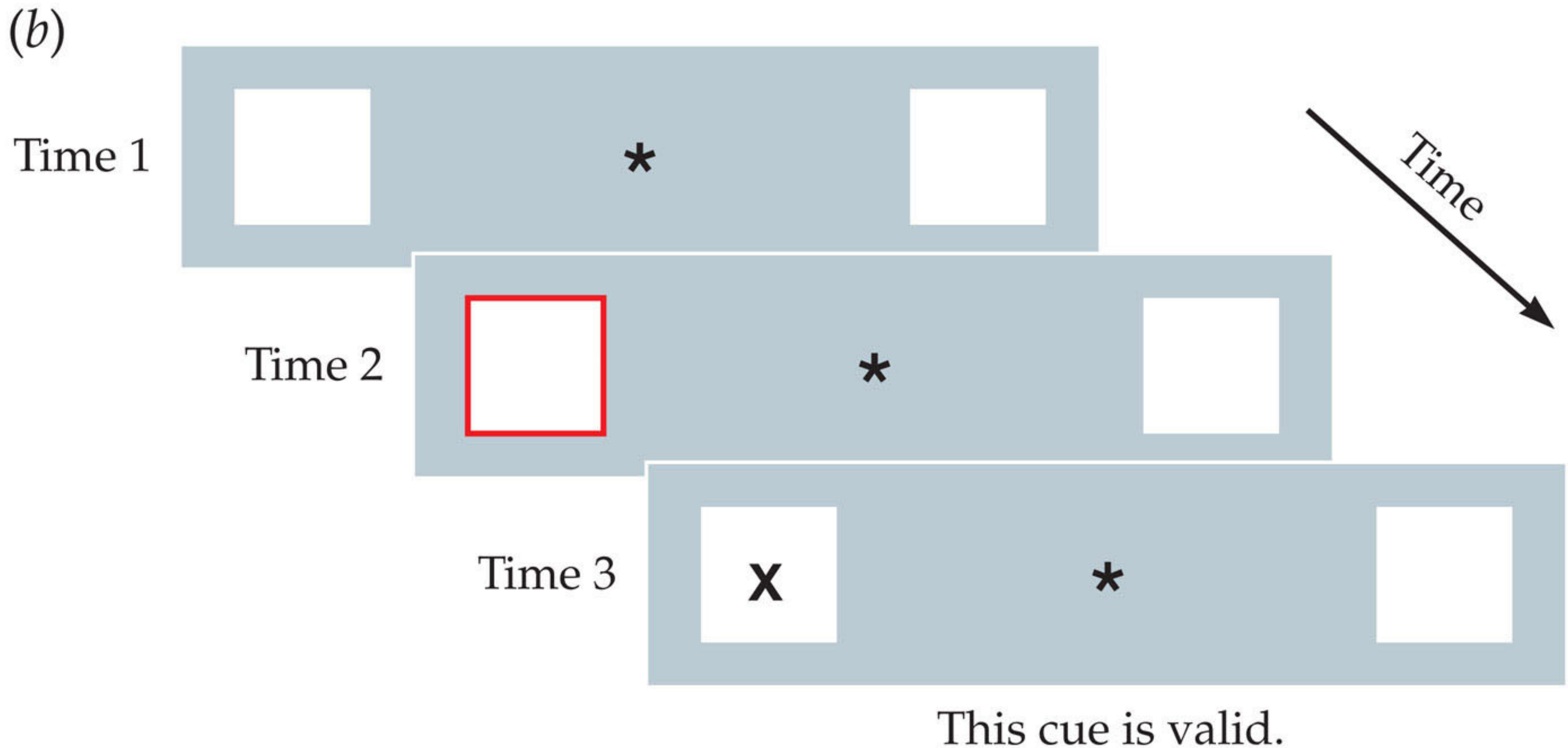
http://en.wikipedia.org/wiki/Posner_cueing_task

Selection in space

RTs are shorter on valid cue trials

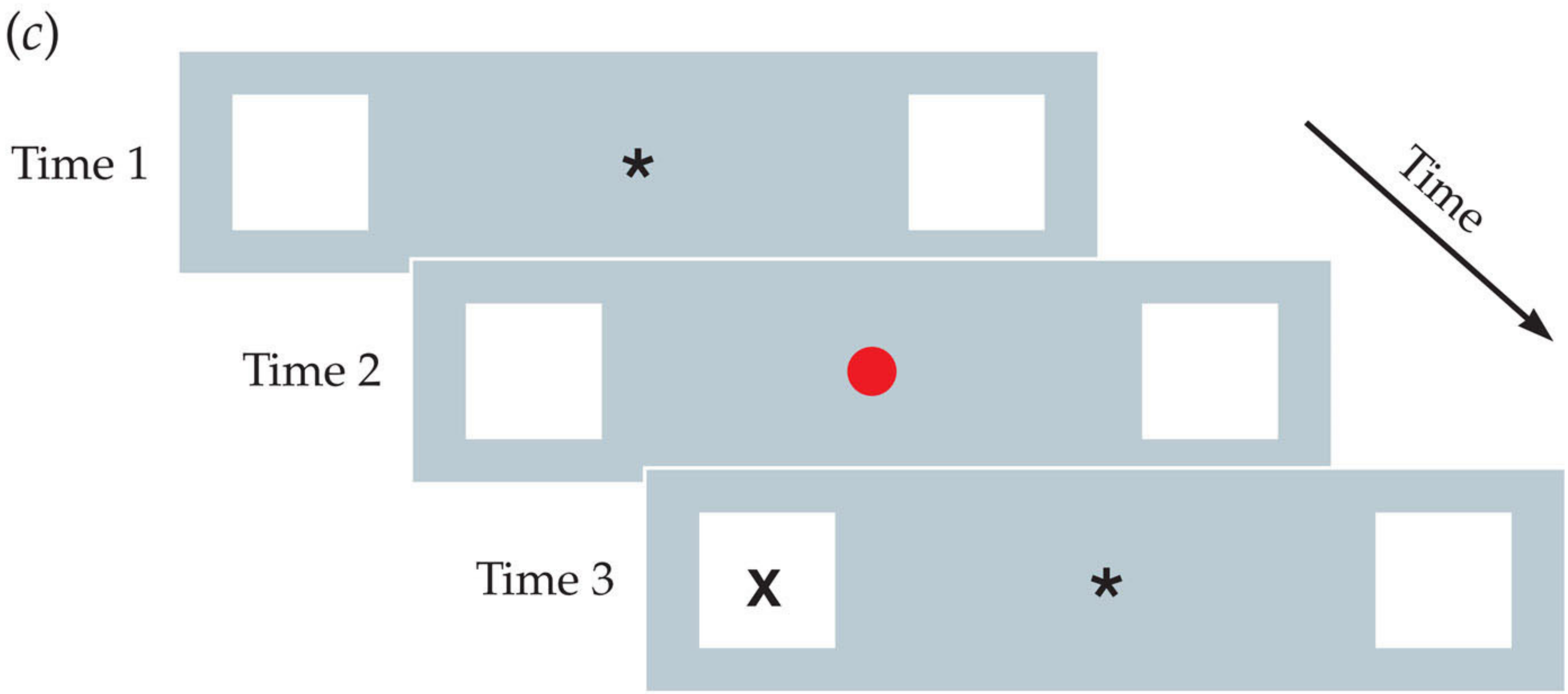
RTs are longer on invalid cue trials

Posner cueing paradigm



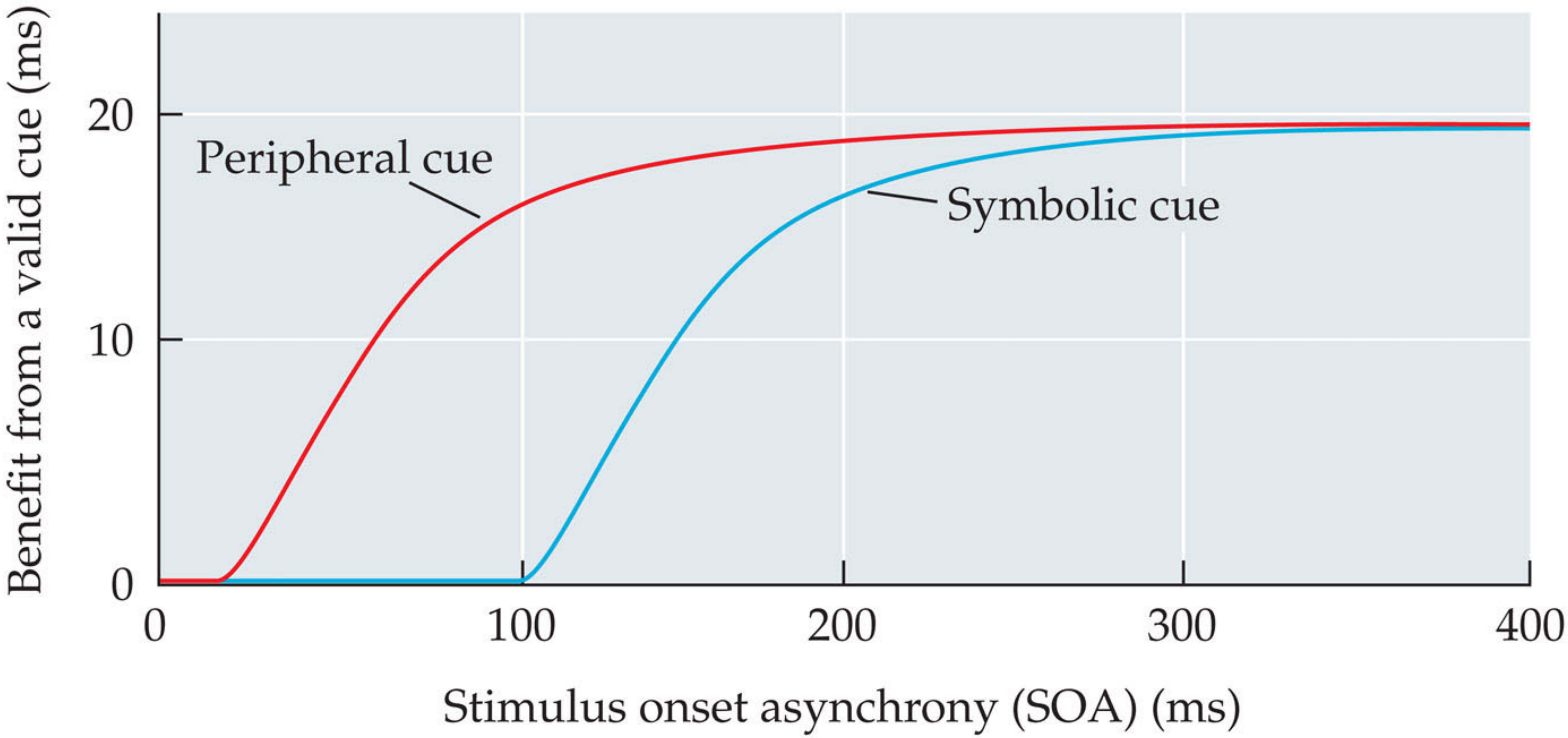
SENSATION & PERCEPTION 4e, **Figure 7.3 (Part 2)**

Posner cueing paradigm



SENSATION & PERCEPTION 4e, Figure 7.3 (Part 3)

The effect of a cue develops over time



SENSATION & PERCEPTION 4e, Figure 7.4

Selection in space

Theories of Attention

"Spotlight" model: Attention is restricted in space and moves from one point to the next. Areas within the spotlight receive extra processing.

"Zoom lens" model: The attended region can grow or shrink depending on the size of the area to be processed.

Remember the search for the unicorn or Waldo at the beginning of the lecture? As if we could only recognise an object if the "spotlight" of attention was right on it—this leads us directly to ...

Visual search

Visual search: Looking for a target in a display containing distracting elements.

Examples: Finding weeds in your lawn or the remote control on the coffee table

Target: The goal of a visual search.

Distractor: In visual search, any stimulus other than the target.

Set size: The number of items in a visual search display.

The efficiency of visual search is the average increase in RT for each item added to the display.

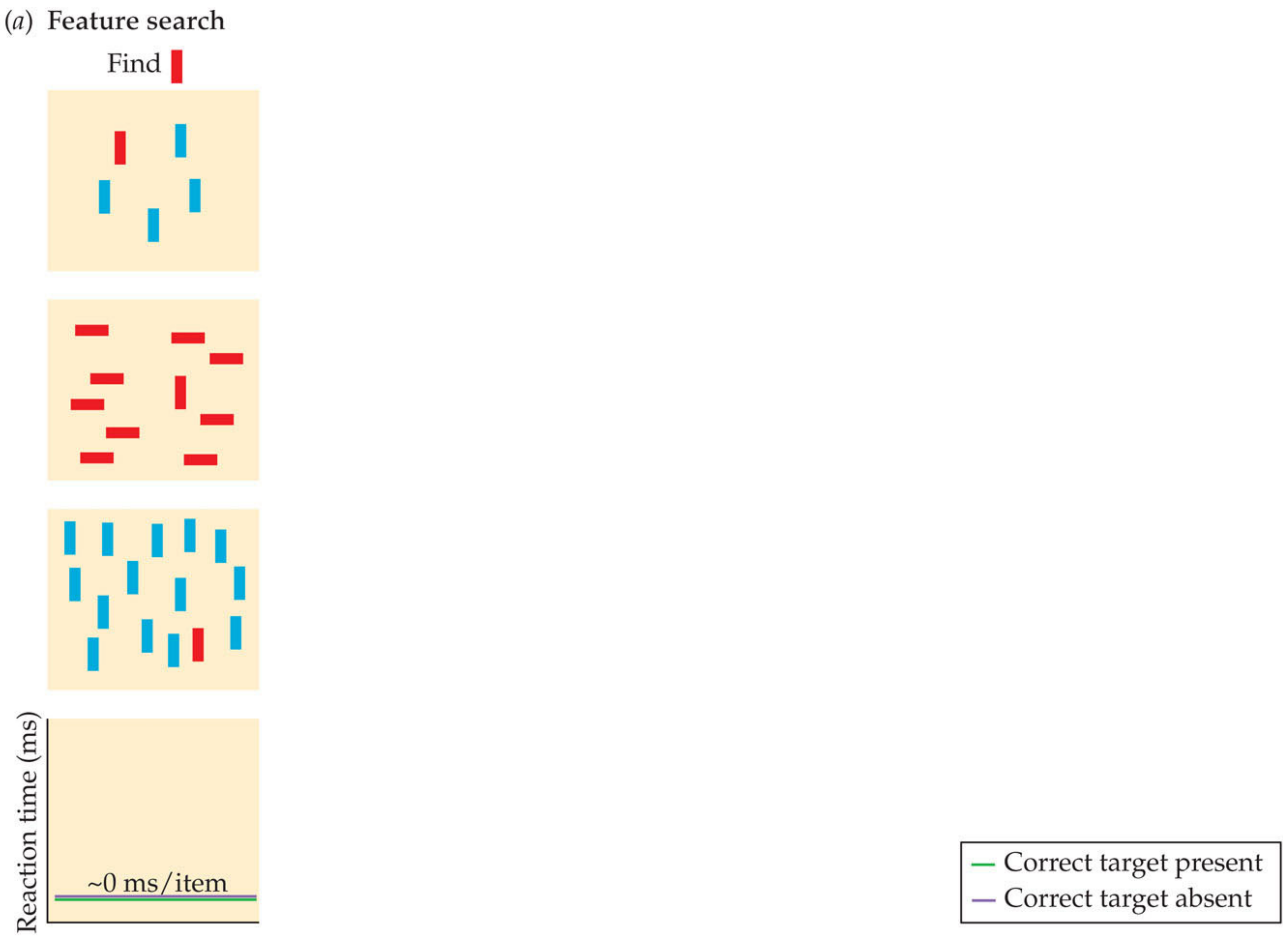
Measured in terms of search slope, or ms/item

The larger the search slope (more ms/item), the less efficient the search

Some searches are efficient and have small slopes

Some searches are inefficient and have large slopes (e.g. looking for Waldo …)

Laboratory visual search task



SENSATION & PERCEPTION 4e, **Figure 7.6**

Visual search

Feature searches are efficient.

Feature search: Search for a target defined by a single attribute, such as a salient color or orientation.

Salience: The vividness of a stimulus relative to its neighbours **(see our next lecture!)**.

Parallel: In visual attention, referring to the processing of multiple stimuli at the same time.

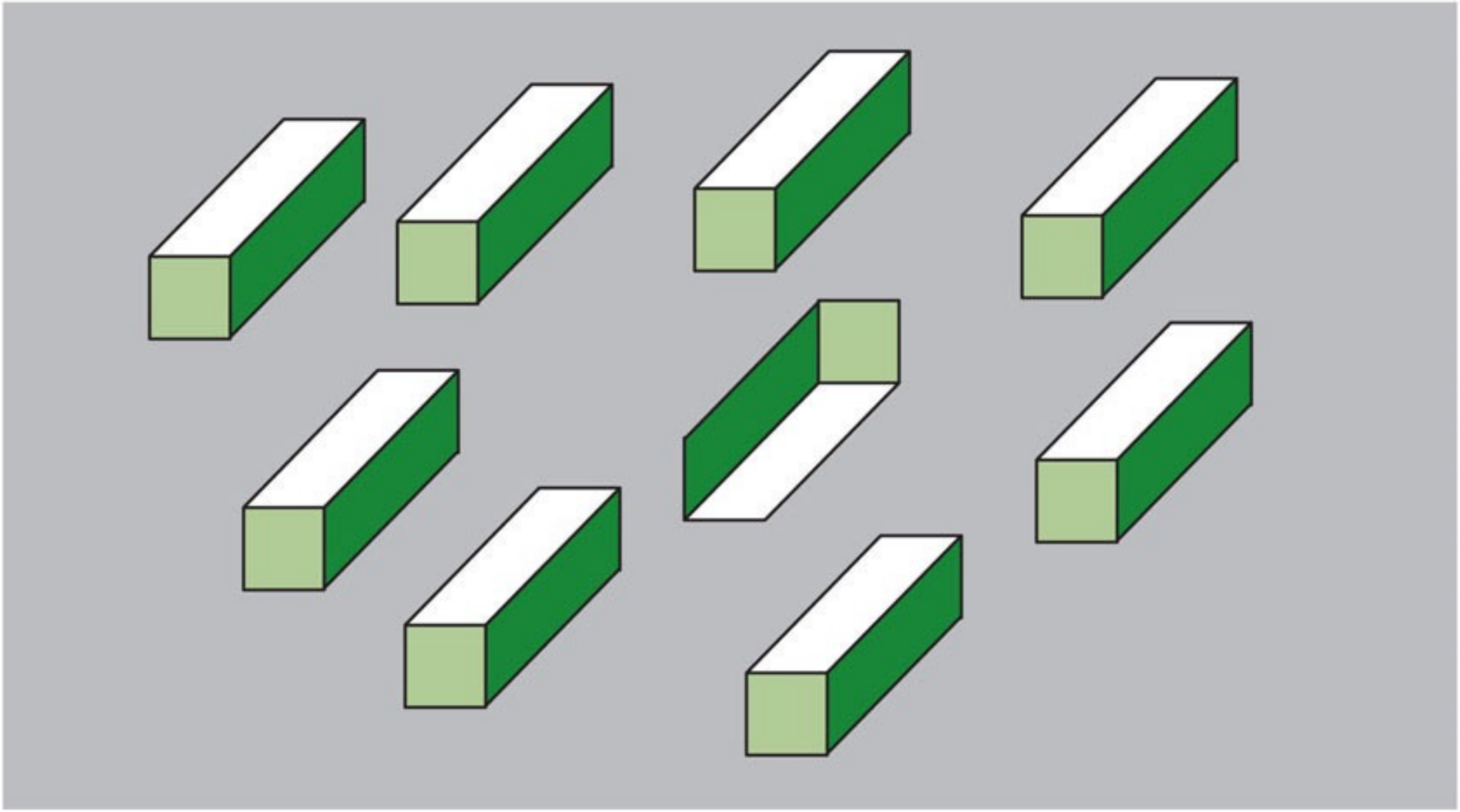
Many searches are inefficient, however.

Serial self-terminating search: A search from item to item, ending when a target is found.

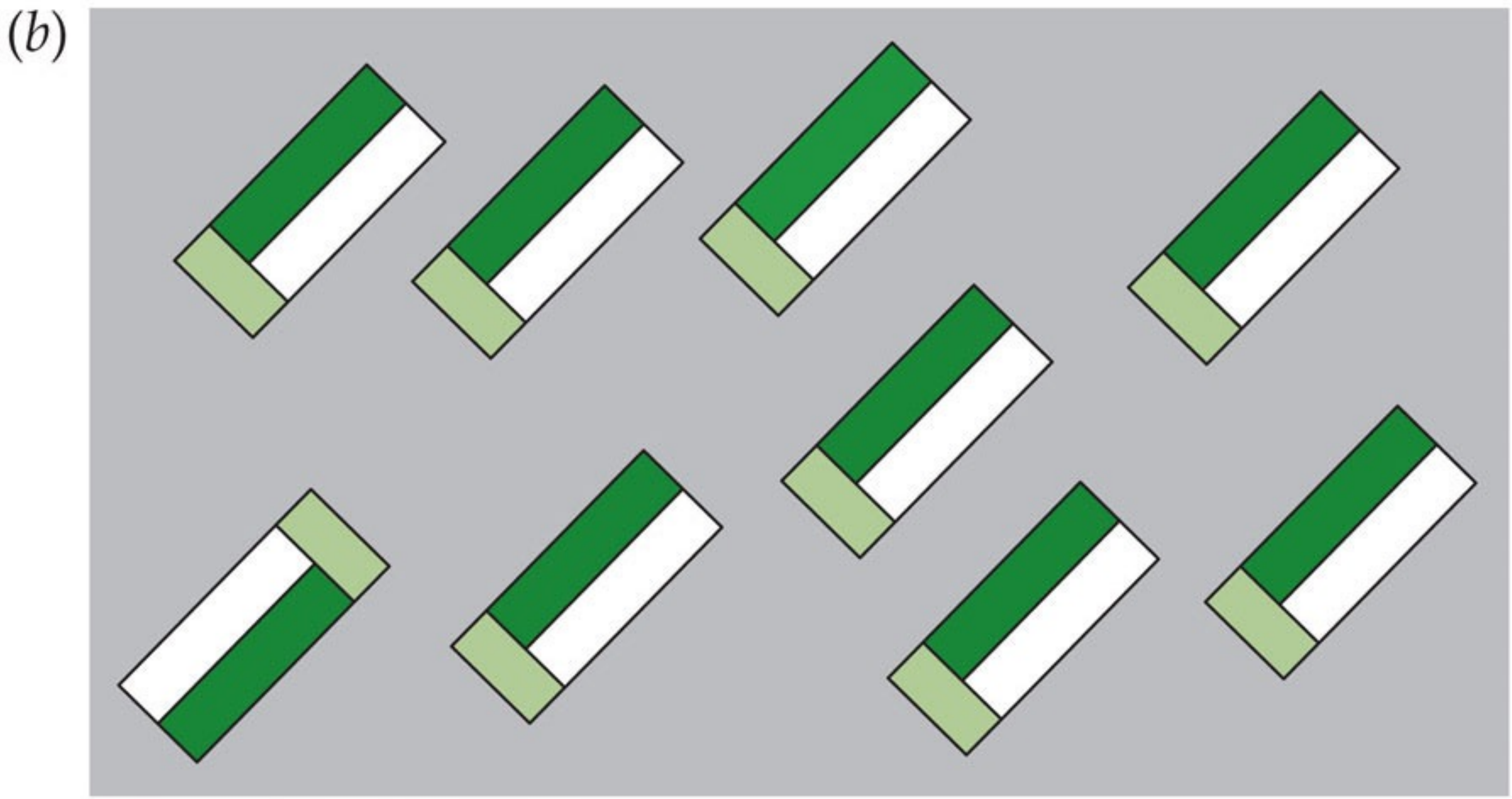
Search can be much more laborious if you're not familiar with what you're searching for

Search can sometimes be surprisingly fast ...

(a)



… and search can sometimes be surprisingly slow!



Visual search

In real-world searches, basic features can guide visual search.

Guided search: Attention is restricted to a subset of possible items based on information about the item’s basic features (e.g., color or shape).

Conjunction search: Search for a target defined by the presence of two or more attributes.

No single feature defines the target

Defined by the co-occurrence of two or more features

A real-world conjunction search: Big, round, red tomatoes

SENSATION & PERCEPTION 4e, **Figure 7.9**

Visual search

In real-world searches, the real world guides visual search (remember: Scene perception)

Scene-based guidance: Information in our understanding of scenes that helps us find specific objects in scenes.

For instance, a mug will typically be found on a horizontal surface and a picture will typically be found on a vertical surface.

Search for arbitrary objects is not very efficient

SENSATION & PERCEPTION 4e, **Figure 7.10**

Scene-based guidance would help you find the water cooker in this scene quickly

***SENSATION & PERCEPTION 4e,* Figure 7.11**

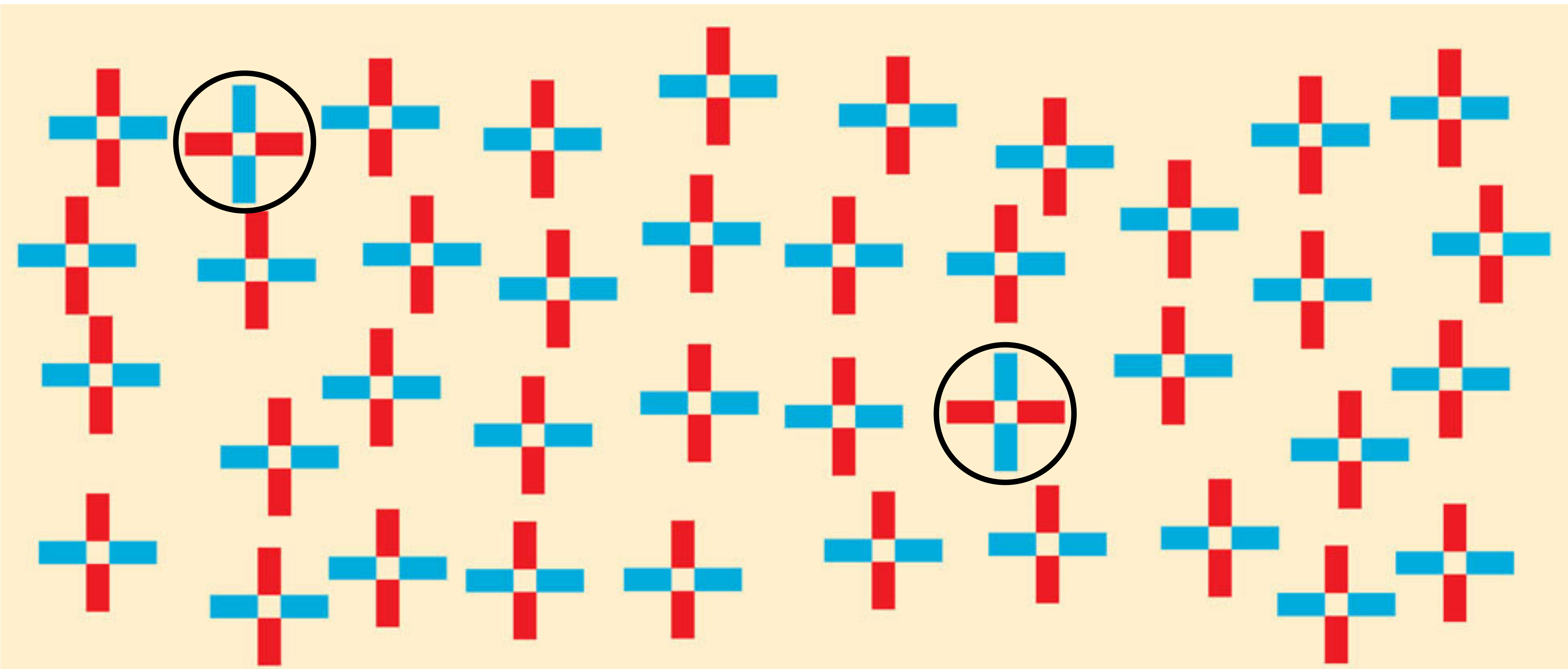
Visual search

The binding problem: The challenge of tying different attributes of visual stimuli, which are handled by different brain circuits, to the appropriate object so we perceive a unified object.

Example: A vertical red bar moving to the right

How do we combine these features when perceiving the bar?

A conjunction search with a binding problem



SENSATION & PERCEPTION 4e, Figure 7.12

Top-down visual attention and feature binding

"Common wisdom" (Rosenholtz et al., 2012) in visual perception that—one of the many roles of attention—is to bind individual features into objects.

[Remember: The early visual system is often thought of as analysing the world along individual dimensions, c.f. local low-dimensional measurements of the plenoptic function.]

Feature integration theory (Treisman & Gelade, 1980) and its many modifications and variants is one of the most prominent and most often cited theories in visual perception.

Feature integration theory: Anne Treisman's theory of visual attention, which holds that a limited set of basic features can be processed in parallel preattentively, but that other properties, including the correct binding of features to objects, require attention.

Preattentive stage: The processing of a stimulus that occurs before selective attention is deployed to that stimulus.

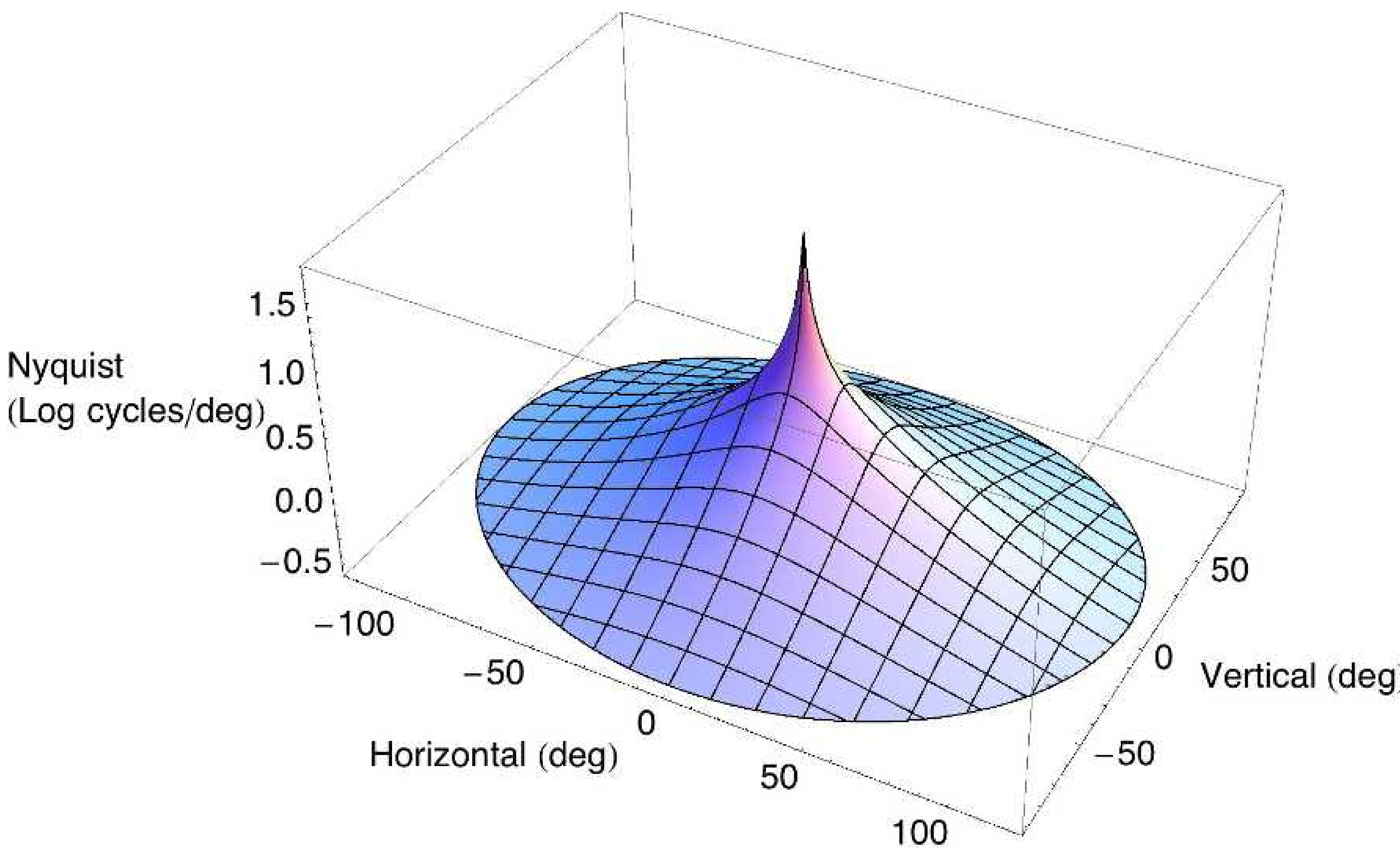
Simple targets that do not need "attentional binding" (sometimes referred to as *pre-attentive*) can be processed by the visual system in parallel, whereas looking for a more complex target requiring attentional feature binding (leading to serial search).

What is wrong with feature integration theory?

… it ignores the profound inhomogeneity of the visual system:

Strikingly different resolution across the visual field!

Human visual sensitivity is not uniform

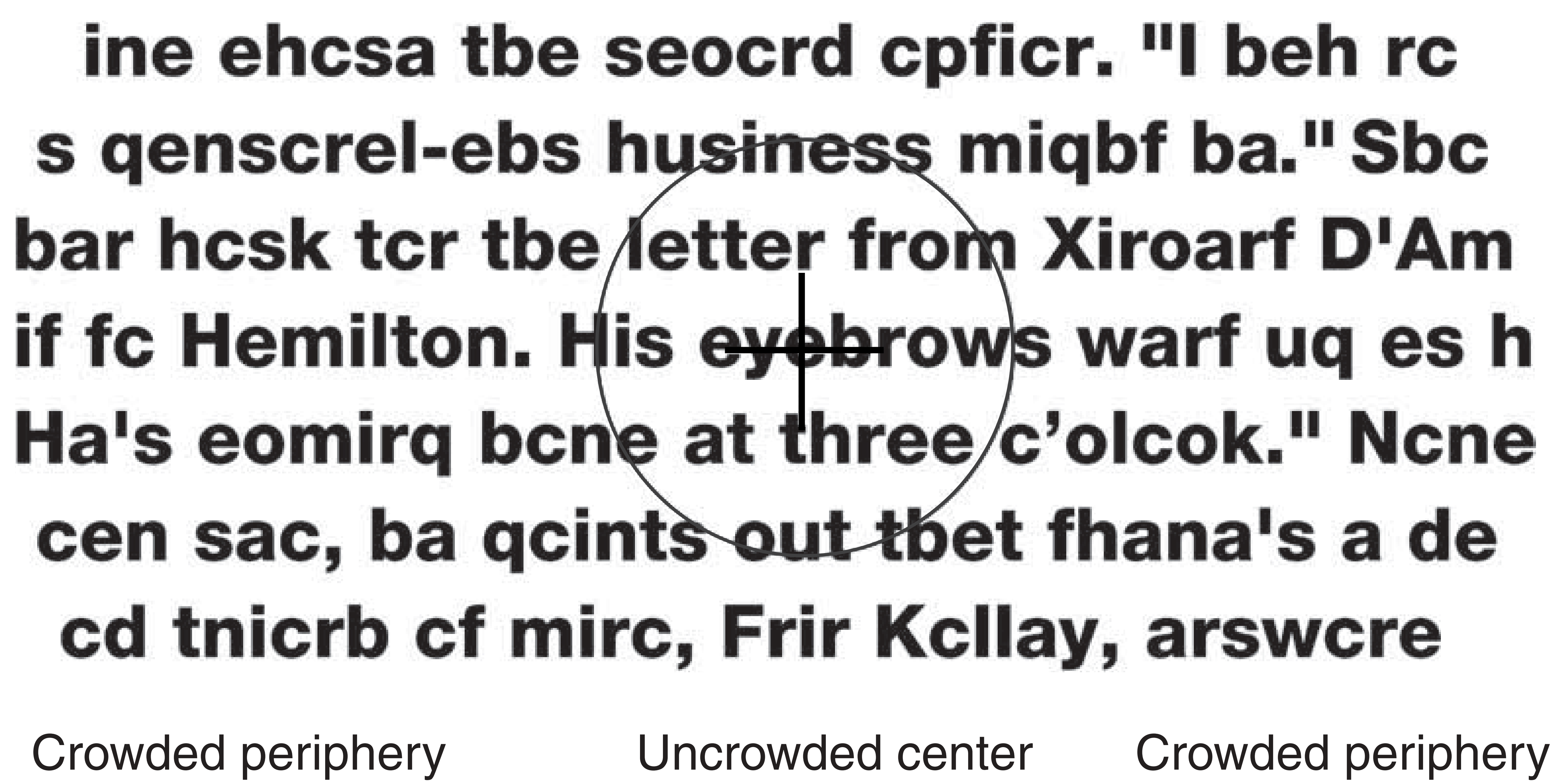


from Watson, A. B. (2014). A formula for human retinal ganglion cell receptive field density as a function of visual field location. Journal of Vision, 14(7), 15:1–17.

Crowding

DCO + C

Crowding



from Pelli, D. G., & Tillman, K.A. (2008). The uncrowded window of object recognition. Nature Neuroscience, 11(10), 1129–1135.

Crowding

Crowding is a phenomenon in which it is more difficult to identify an object when surrounded by flanking objects.

Not masking: detecting the presence of the object is not affected – a bottleneck of recognition

Primarily mediated by cortical neurons: flankers in a different eye to the target are almost as effective at causing crowding as flankers in the same eye.

Primarily peripheral but can be found in the fovea too (just need really small objects!)

Crowding zones scale with eccentricity according to ''Bouma's law'': flankers begin to cause crowding at approximately half the retinal eccentricity of the target

Peripheral vision as “compulsory texture perception”

Rosenholtz et al, 2012


Bundesliga
M'GLADBACH 0-1 MAINZ
Eine Bank fürs Leben.
Postbank
BORUSSIA-PARK
Bitburger
Kappa
Football
Mainz goalkeeper takes his eye off the ball
theguardian

What is wrong with feature integration theory?

… it ignores the profound inhomogeneity of the visual system:

Strikingly different resolution across the visual field!

Rosenholtz et al. (2012; see also Geisler & Chou, 1995) argue that serial vs. parallel search is simply a function of how easily the target & the distractors can be discriminated *in the periphery:*

Serial search if you need to foveate on a target (or distractor) to tell which is which …

Parallel search if the difference is robust enough to be visible with our low resolution periphery.

The decodability of a search task from summary statistics is predictive of search times in humans

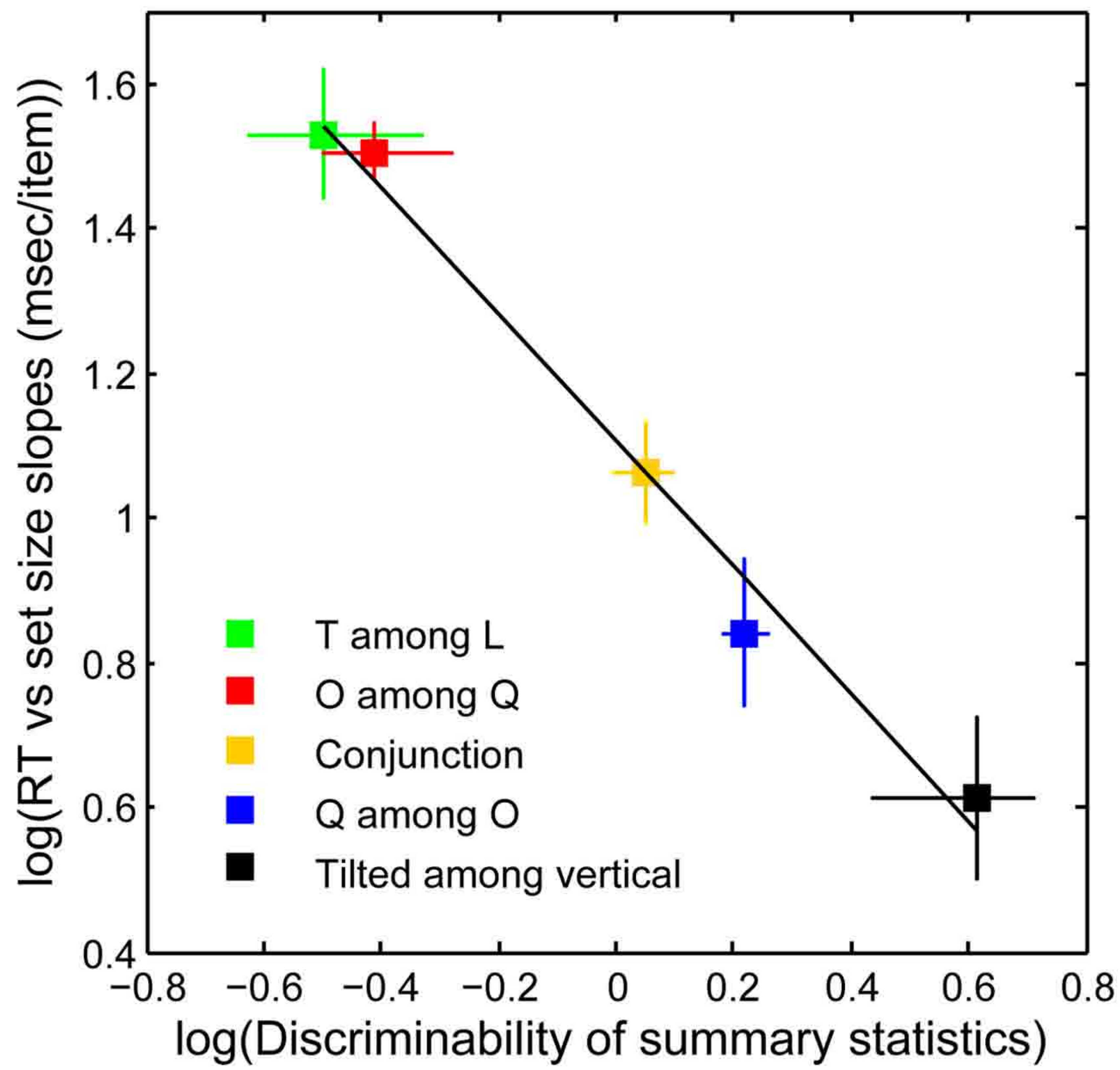


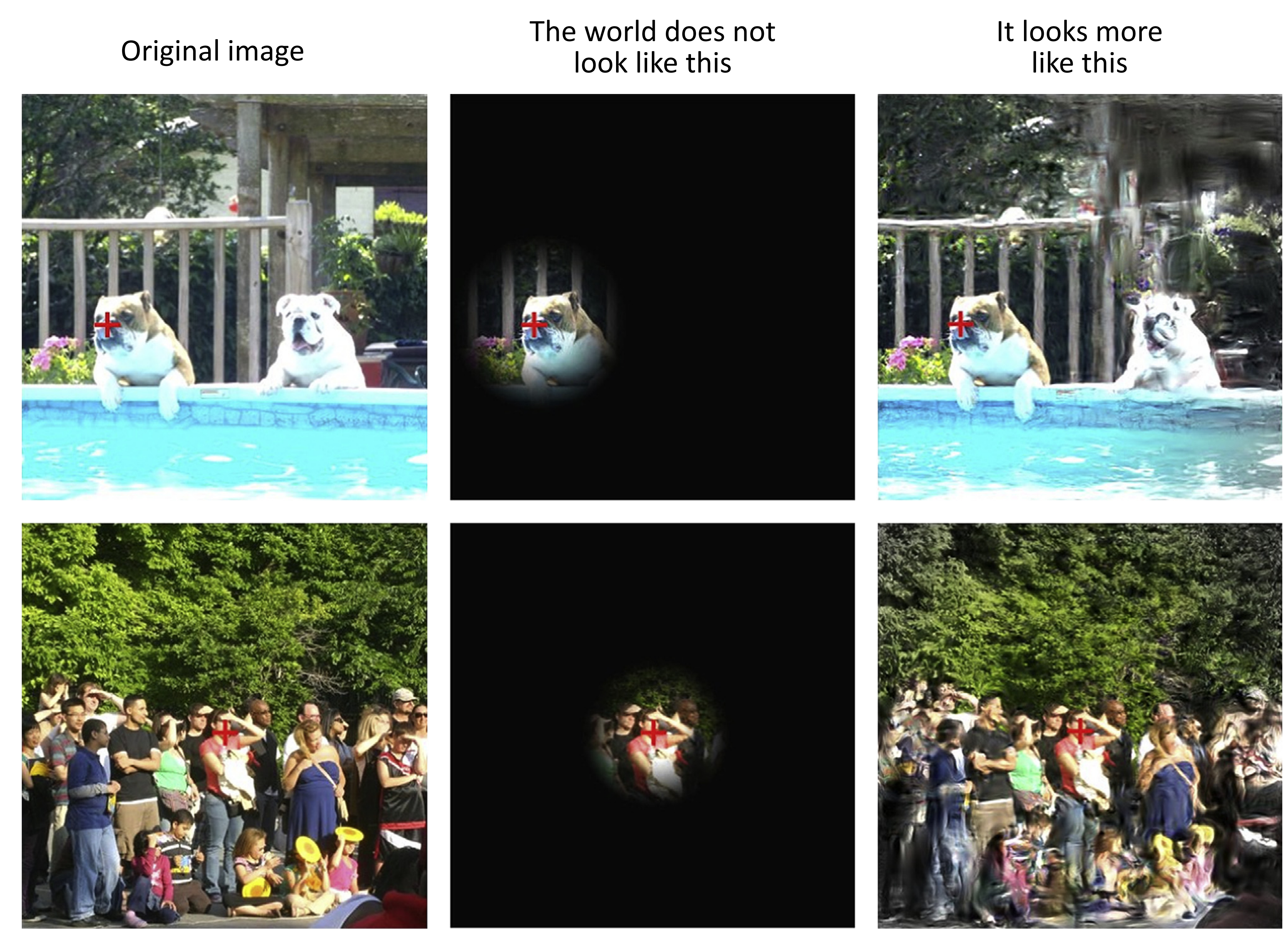
FIGURE 6 | Search performance vs. statistical discriminability. *y*-Axis: search performance for correct target-present trials, as measured by log 10 (search efficiency), i.e., the mean number of milliseconds (ms) of search time divided by the number of display items. *x*-Axis: "statistical discriminability" of target-present from target-absent patches based on the empirical discriminability, *d′*, of the corresponding mongrels. There is a strong relationship between search difficulty and mongrel discriminability, in agreement with our predictions. [*y*-axis error bars = SE of the mean; *x*-axis error bars = 95% confidence intervals for log 10 (*d′*)].

Opinion

What is the Bandwidth of Perceptual Experience?

Michael A. Cohen,[1,*] Daniel C. Dennett,[2] and
Nancy Kanwisher[1]

Although our subjective impression is of a richly detailed visual world, numerous empirical results suggest that the amount of visual information observers can perceive and remember at any given moment is limited. How can our subjective impressions be reconciled with these objective observations? Here, we answer this question by arguing that, although we see more than the handful of objects, claimed by prominent models of visual attention and working memory, we still see far less than we think we do. Taken together, we argue that these considerations resolve the apparent conflict between our subjective impressions and empirical data on visual capacity, while also illuminating the nature of the representations underlying perceptual experience.



The End

Felix Wichmann

Neural Information Processing Group
Eberhard Karls Universität Tübingen