ASSC 4, Brussels, Belgium

Change Detection, Attention, and the Contents of Awareness

Ron Rensink, Cambridge Basic Research
Dan Simons, Harvard University

- Part 1: **Intentional** detection of change (Rensink)
- Part 2: **Incidental** detection / attn capture (Simons)
- Part 3: New findings and new issues (everyone)

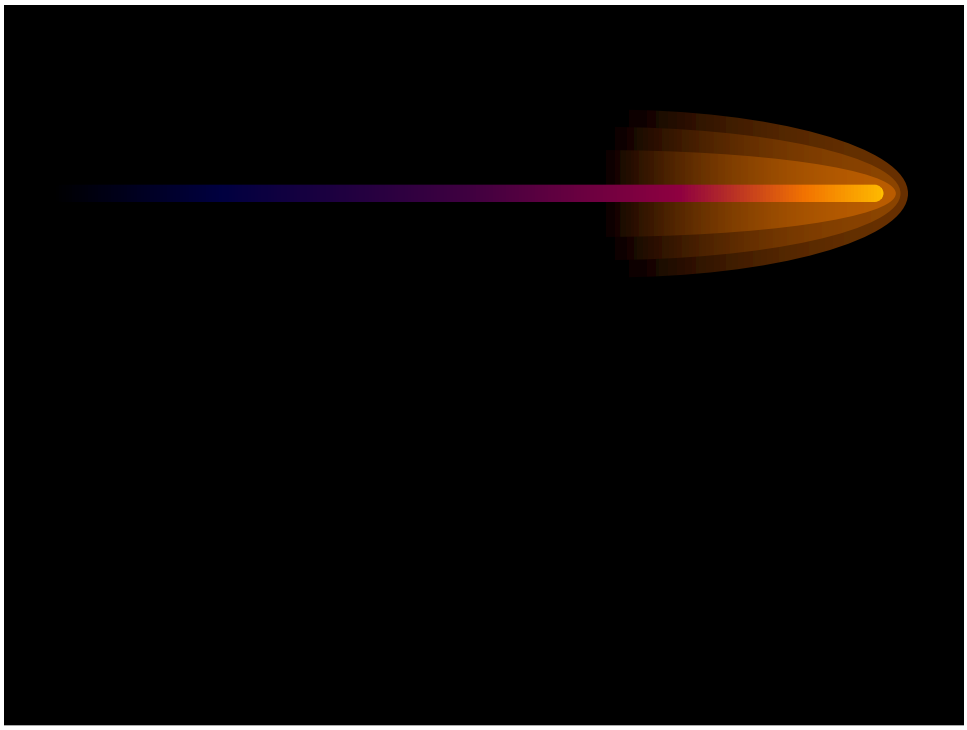
1

2

Part 1: Intentional Search for Change

- **Change blindness**
  - basic phenomenon
- **Representation of scenes**
  - virtual representation
- **Attentional system**
  - attentional capacity; binding problem
- **Nonattentional systems**
  - implicit perception of change

3



4

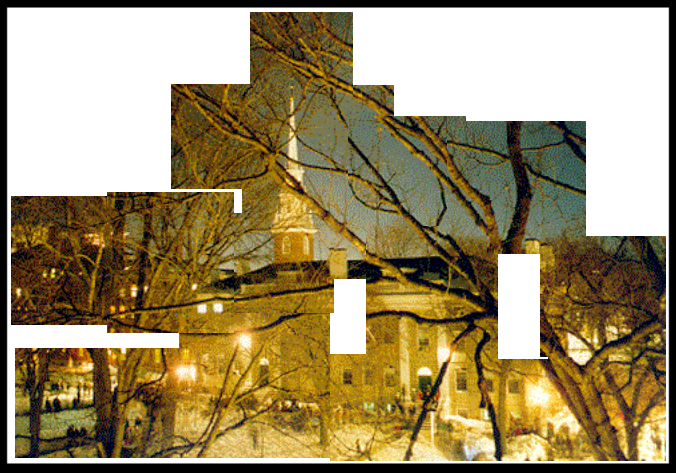
1. Change Blindness

1.0 How Do People See Scenes?

5

6

Visual buffer: accumulates information



7

8

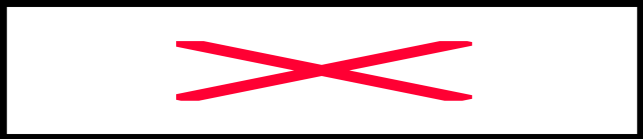
9

Question:

How is this accumulation implemented?

Suggestions:

- shift retinal contents (e.g., Trehub, 1991)
- data fusion (e.g., Clark & Yuille, 1990)



10

11

1.1. Looking Again at the Basic Assumptions

Questi

Hov **cumulated?**

12

13

Why believe we accumulate detailed information?
- always seems easy to notice changes...

14

15

But is this always true?

Make change during brief blank interval between original and changed images

(flicker paradigm - Rensink et al., 1997)

Demo

16

17

This is known as change blindness

- induced for changes simultaneous with:

- image flicker (e.g., Rensink et al., 1995, 1997)
- saccades (e.g., Grimes, 1996; Henderson, 1997)
- eyeblinks (O'Regan, Deubel, et al., 1999)
- "splats" not on change (Rensink et al., 2000)
- movie cuts (Levin & Simons, 1997)
- real-world interruptions (Simons & Levin, 1998)

Proposal: All these have the same cause

18

19

Proposal: Attention is needed to perceive change in an object.

- Under normal circumstances, a change creates a motion transient, which draws attention.
- When change is made same time as other event, transients interfere with drawing of attention, causing change to become “invisible”.

20

21

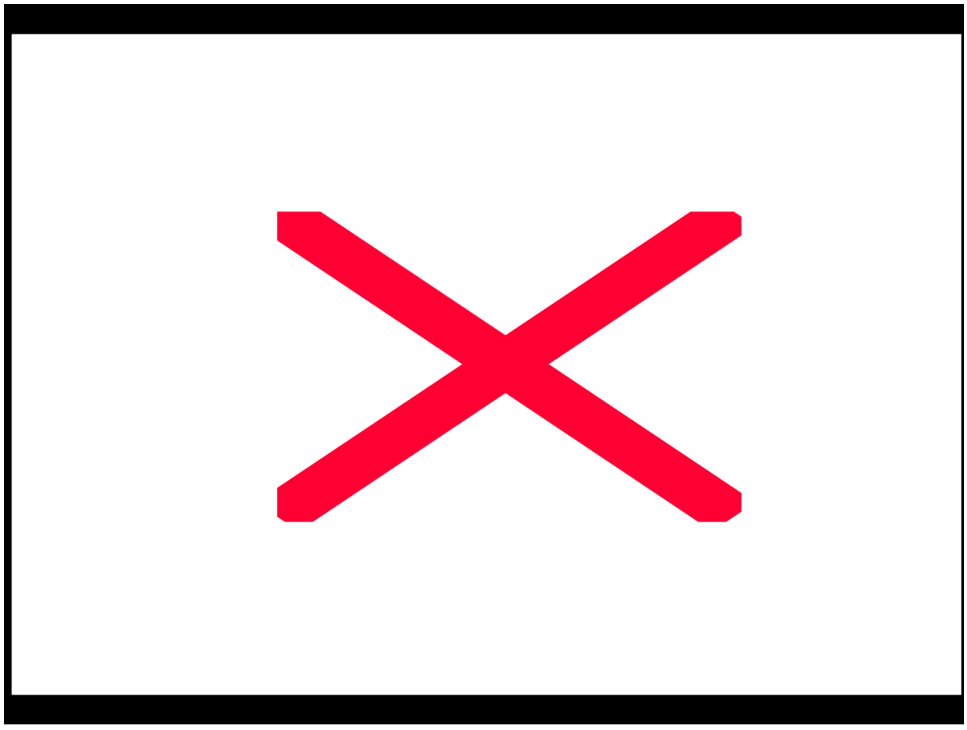
1.2 Coherence Theory

Background assumptions

- Complex pre-attentive structures (**proto-objects**) are formed rapidly across the visual field.
  - 3D slant (Enns & Rensink, 1991)
  - object completion (Rensink & Enns, 1998)
- These are what attention has rapid access to, **not image properties**.
  - lower-level structures (pixels, etc.) are **pre-empted** by higher-level structures. (Rensink & Enns, 1995, 1998)

22

23

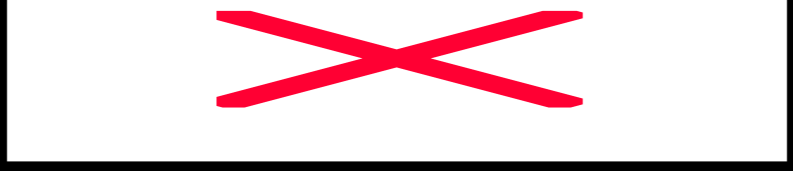


24

25

Proposal – Coherence theory

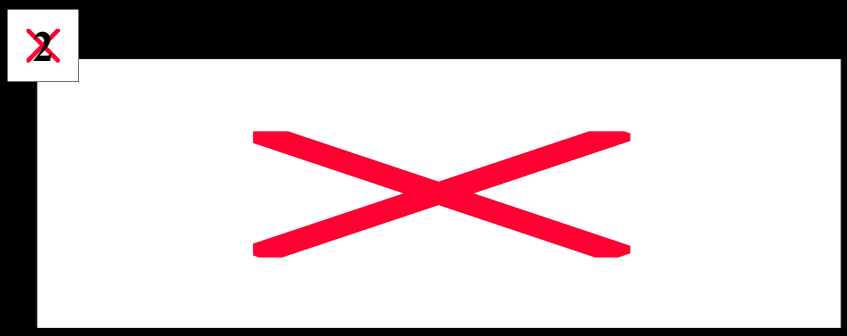
1



26

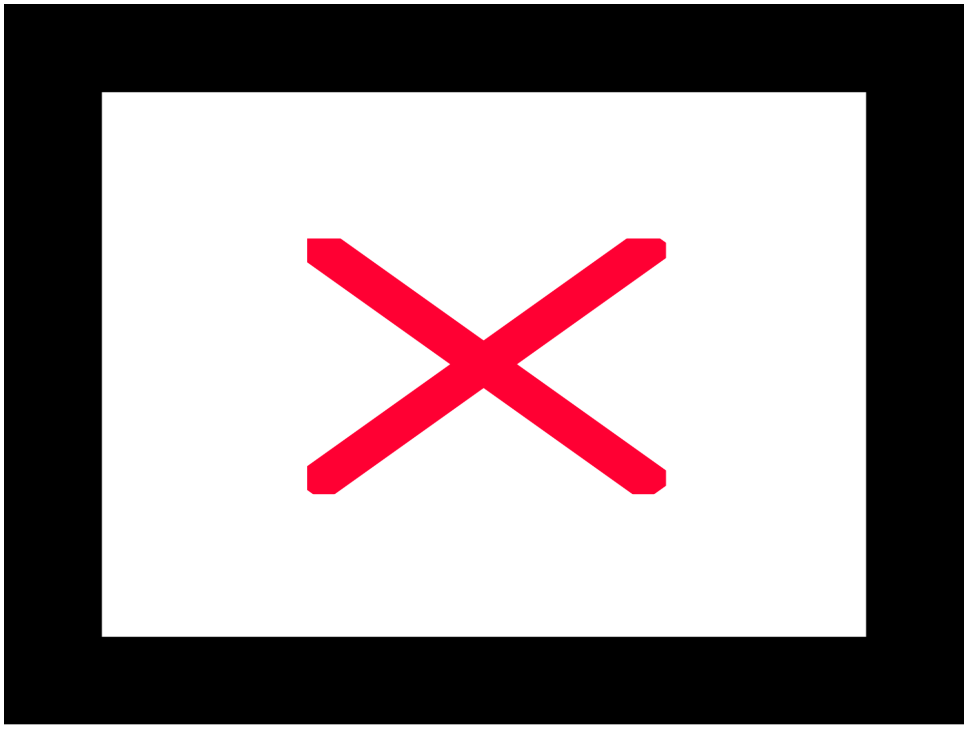
27

Coherence theory— (cont'd):

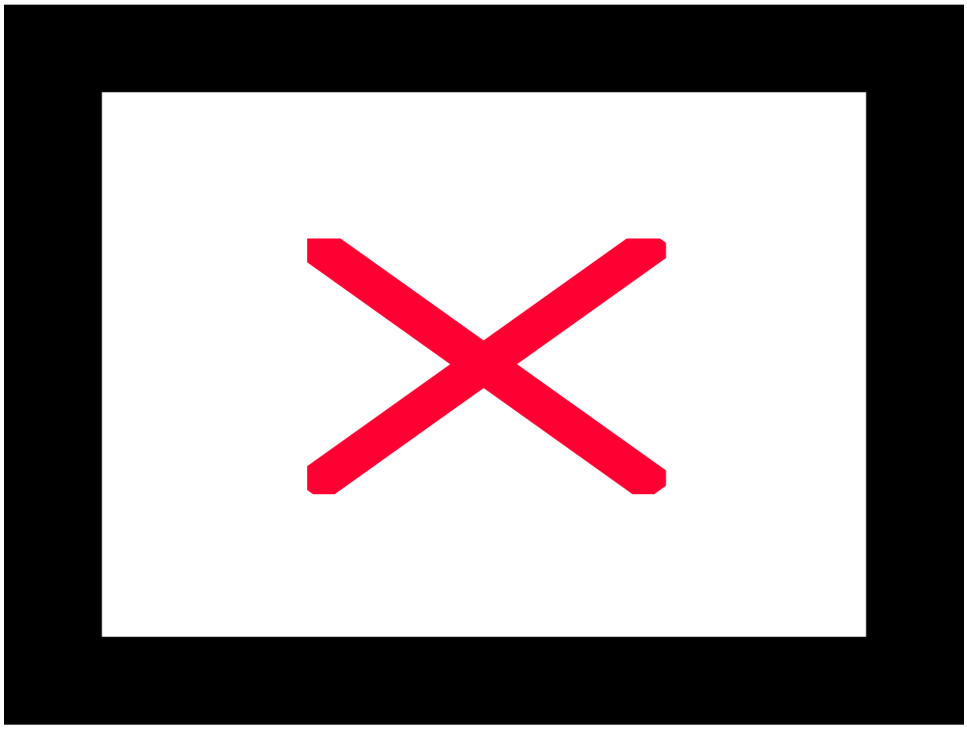


28

29



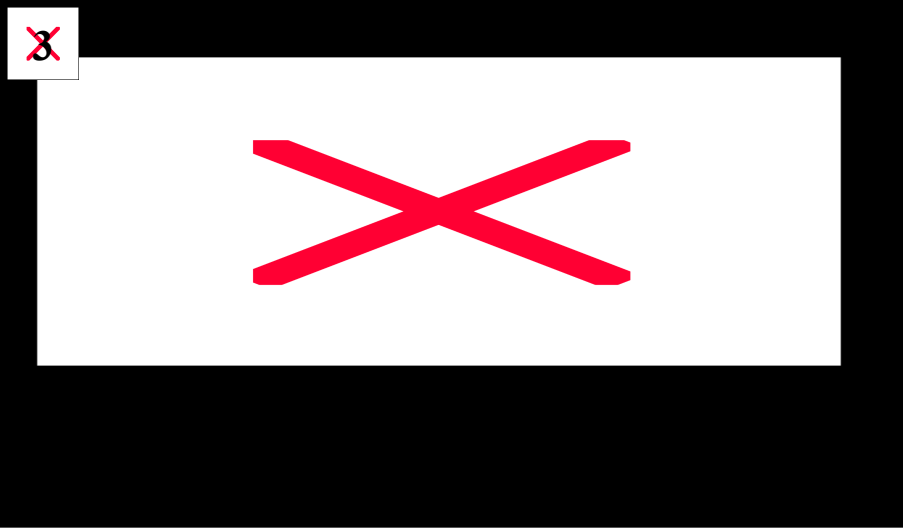
30



31

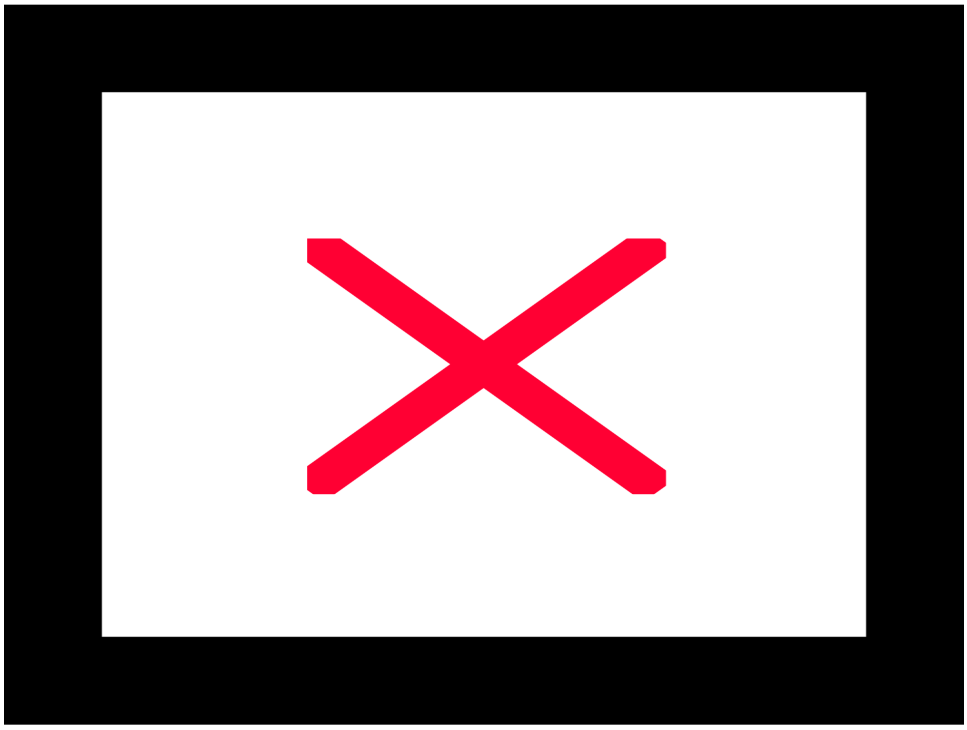
32

Coherence theory— (cont'd):

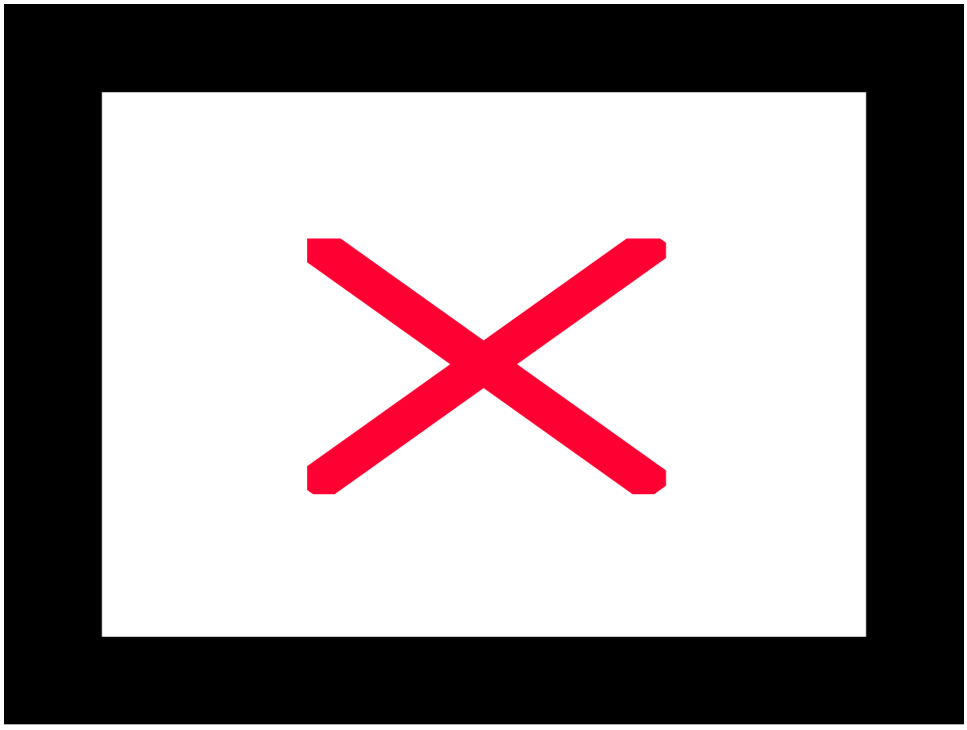


33

34



35



36

37

Important points (Section 1)

- Attention is needed to see change.
  - otherwise, **change blindness** results
- Unattended proto-objects are volatile, but still provide context for interpreting attended objects.
  - **low-level** representation isn't sparse
  - only **stable** representations are sparse.

38

39

2. Seeing

2.0 Implications for scene perception

Older view: scene perception carried out by a sequence of operations:

pixels > edges > regions > objects > scenes > …

-> Build up a complete description of scene

40

41

However...

Change blindness shows that we can only integrate what we attend to

And we can't attend to much

Thus, can't form representations that are both detailed **and** complete.

-> **have coherent representation of only a few objects at any moment**

42

43

If only a few object representations exist
at any given instant...

Why do we have the impression of perceiving lots of coherent objects simultaneously?

44

45

2.1 Virtual Representation (Rensink, 2000)

Observation:

- Although objects are simultaneously present, do not need to be simultaneously **represented**
- All that is needed is that the properties of each object can be **accessed when requested.**

46

47

This is **virtual representation**
- coherent representations
are formed "just in time"

If co-ordination is successful, it will appear to higher levels as if representation is "real", i.e., as if all items present simultaneously.

48

49

Can this work for the visual system? **Yes!!**

-can always **obtain information from the world**
- use the world itself as a memory
(Stroud, 1955; Brooks, 1988)

-to build a coherent representation of an object, focus eyes and attention on appropriate location whenever that object is needed

-only 1-2 independent objects needed for a task at any given time

50

51

Note 1:

-although world is an **external memory**, it is **not an external representation** (as proposed by Brooks, 1988; Noë et al, 2000)

- representations needed at early levels for various purposes

- e.g., compensating for object occlusion

- linking together elements in the image that are related in the scene

52

53

Note 2:

- using the world as an external memory means that **perception is not carried out in isolation** in the perceiver—rather, the perceiver and environment form a **partnership**.

- environment can act not only as **external memory** but as an **external processor**
  - embedded cognition (see Clark, 1997).

54

55

2.2 Triadic Architecture (Rensink, 2000)

Question:

How might a virtual representation be implemented in the human visual system?

Need to do this in a way that is compatible with what is known about the visual system.

56

57

- **Nonattentional** extraction of aspects of scene (I):
  - **Gist**: abstract meaning of scene (farm, harbor, etc.)
    - obtained within 200 ms (Biederman, 1974)
    - obtained without attention (Oliva & Schyns, 1997)

58

59

60

61

- **Nonattentional** extraction of aspects of scene (I):
  - **Gist**: abstract meaning of scene (farm, harbor, etc.)
    - obtained within 200 ms (Biederman, 1974)
    - obtained without attention (Oliva & Schyns, 1997)
  - Possibly derived via statistics of low-level structures (e.g. Swain & Ballard, 1991; Oliva & Torralba, 2000)

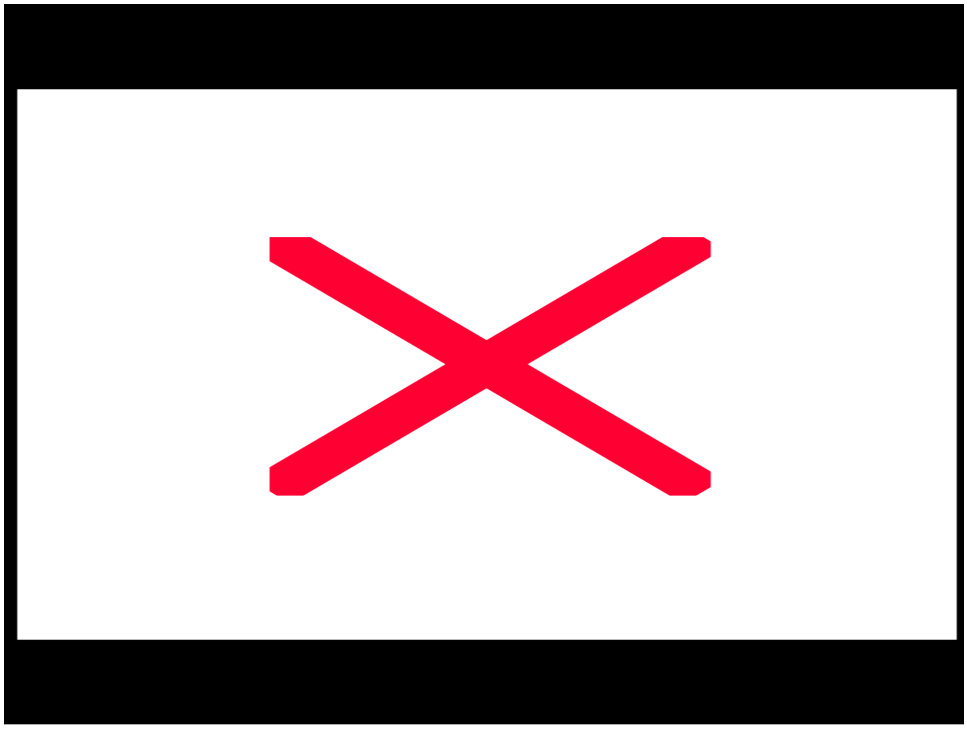
62

63

- **Nonattentional** extraction of aspects of scene (II):
  - **layout**: arrangement of items in the scene.
    - nonvolatile (Simons, 1996; Sanocki & Epstein, 1997)
    - can be learned without attention (Chun & Jiang, 1998)

64

65



66

67

Important Points (Part 2)

- Our impression that many coherent objects are represented simultaneously is only an impression.
  - scenes represented via **virtual representation**
- **Attention not a central processing "gateway"**
  - stream specialized for **coherent objects**

68

69

3. Visual Attention

Change blindness is not just an interesting effect in its own right.

Can use the experimental techniques and theoretical structure to find out more about **attention itself**.

- **capacity** — how many items are "held" at a time?
- **speed** — how fast are attended objects formed?
- **selectivity** — how well can properties be selected?

70

71

Approach: Visual Search for Change

Visual Search:

- display is an array of identical items (distractors)
- on half the trials, one of the items differs (target)
- observer must report if target present or absent

target absent

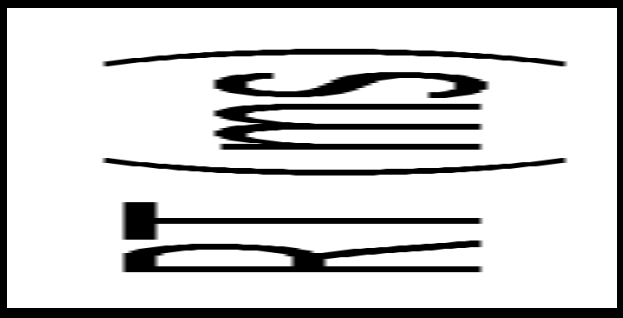
target present

72

73

Measure: Reaction time (RT) vs. set size

RT often a linear function of number of items



search slope = Δ (reaction time) / Δ (# of items)

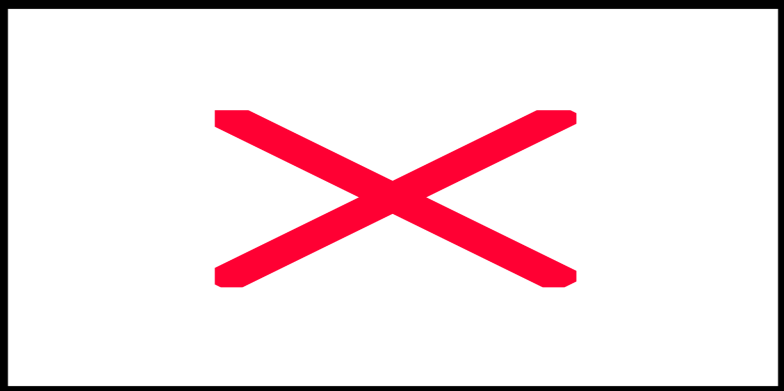
target-absent slope usually 2x target-present slope
-> **serial, self-terminating search**

74

75

Visual search for change (Rensink, 1998, 2000)

- display is an array of flickering non-changing items
- on half the trials, one of the items changes (target)
- observer must report if target present or absent



76

77

Example: To determine attentional capacity

- if only 1 item held by attention,
  search slope = alternation time

- if 10 items held by attention,
  search slope = alternation time / 10

- more generally, if n items held by attention,
  search slope = alternation time / n

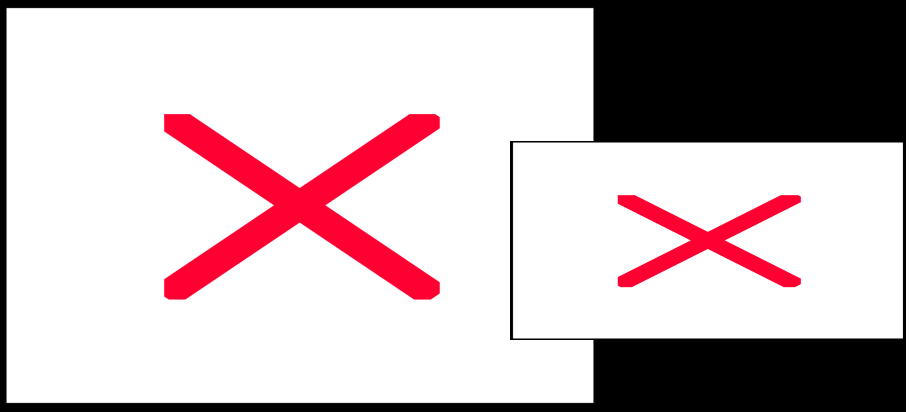
items held = alternation time / search slope

Capacity = maximum number of items held

78

79

*Results: Search for orientation **change***
(changing target, nonchanging distractors)

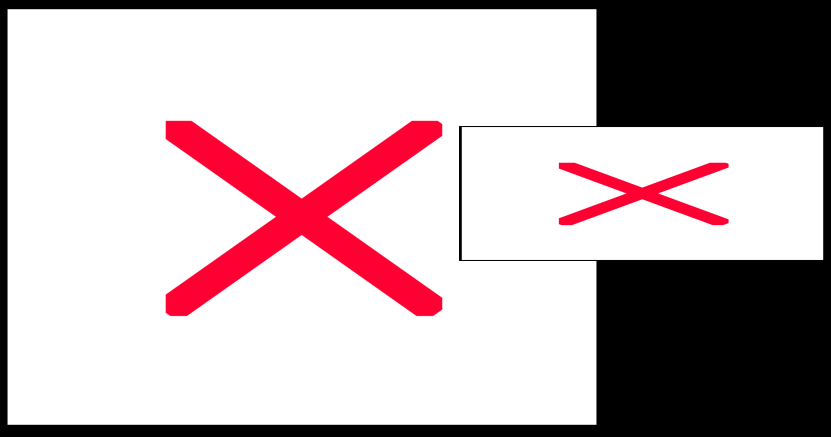


Focused attention can hold on to ≤ 5.5 items

80

81

Dependence on ISI (Rensink et al., 2000)



Focused attention can hold on to 4 items

82

83

Results:

- Attention loads up over time
  - loading rate = 8 items/sec
- Attention has a capacity of about 4 items
  - similar to other estimates of attentional capacity
- Demonstrates that visual detail is not built up
  - otherwise, capacity estimate would be unlimited

84

85

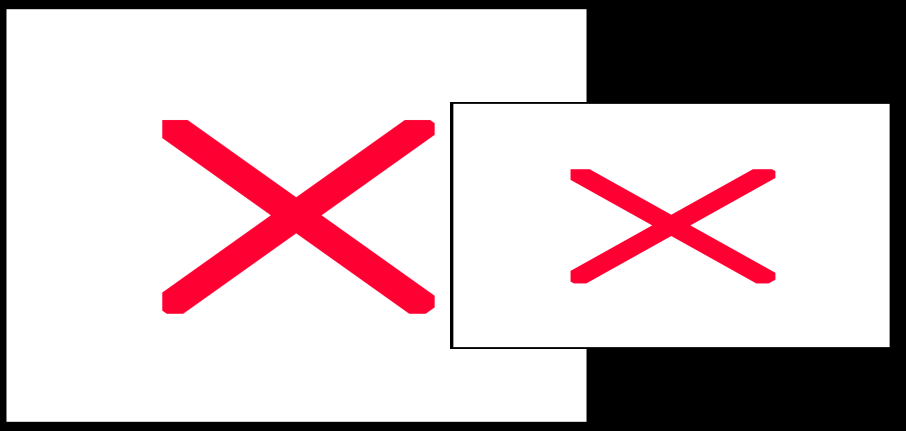
What about search for <u>absence</u> of change?

- targets remain constant
- nontargets continually changing orientation

86

87

*Results: Search for **absence of change***
(non-changing target, changing distractors)



Focused attention can hold on to only 1 item

88

89

Why is there a search asymmetry?

Reminiscent of asymmetry in "classic" visual search
- detecting presence of feature (change)
is much easier that detecting absence

Asymmetry explained via **pooling** of signal
over image (Treisman & Gormican, 1985)

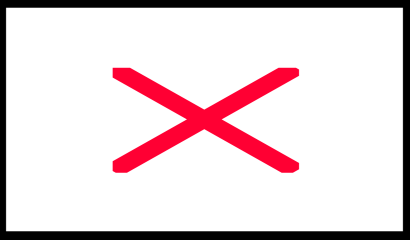
- detecting presence: 1 vs 0
- detecting absence: n vs n-1

90

91

Hypothesis: *Information from 4-5 links pooled into a* ***single nexus***

(a) Searching for presence of change



If change present, nexus value = 1

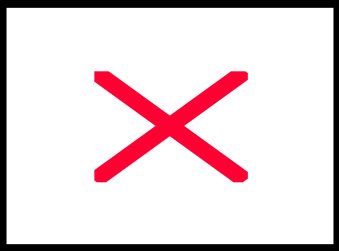
If change absent, nexus value = 0

Thus, present vs absent is 1 vs 0 — strong signal

92

93

(b) Searching for absence of change



If nonchange present, nexus value = 4

If nonchange absent, nexus value = 5

Thus, present vs absent is 4 vs 5 — weak signal
-check links one at a time to get strong signal

94

95

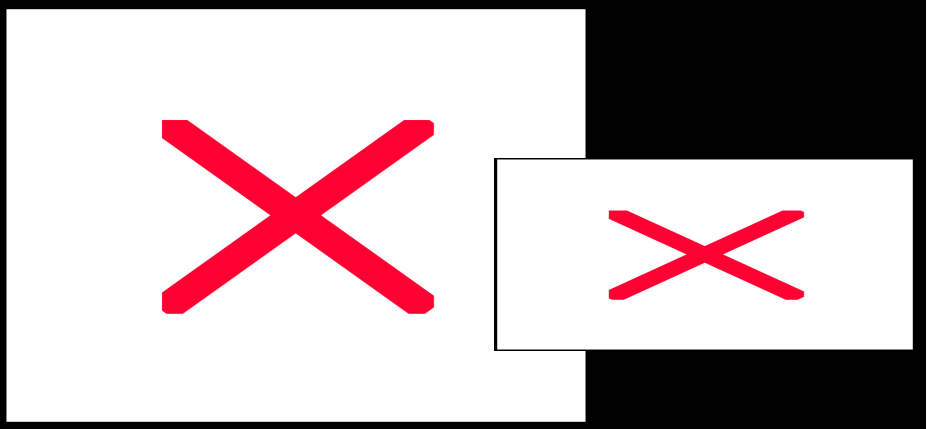
What about search for <u>conjunction</u> of change?

- nontargets continually change on **<u>one</u>** dimension (orientation **<u>or</u>** contrast sign)
- target continually changes on both dimensions (orientation **<u>and</u>** contrast sign)

96

97

*Results: Search for **conjunction of change***
*(target changes in orientation **and** contrast)*

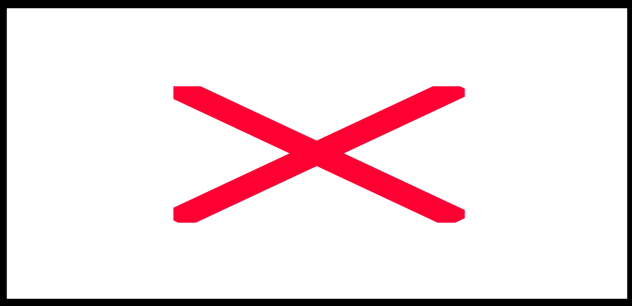


Focused attention can hold on to only 1 item

98

99

Binding of attended items is incomplete:



Information from each item is not held separately—features are pooled into nexus

100

101

Binding problem: how to correctly assign visual properties to different objects

- if only one object represented at a time, no problem with assignment of properties

-> **If only one collection point (nexus), all attended properties assigned to that object**
- further distinctions (e.g. binding of parts) inaccessible to higher levels

102

103

Binding problem becomes irrelevant
(cf. problem of integrating saccades)

- only appears that than one object is represented simultaneously

Problem becomes one of gating:
- when to allow which properties into the coherence field?

104

105

Important points (Section 3)

- Only one nexus (object) can be attended at a time; inputs funnelled in from 4 links (parts).
- Binding problem may be largely irrelevant.
  In its place is the **gating problem**:
  - which items/properties to put into coherence field?

106

107

4. Nonattentional Streams

Virtual representation implies an important role for **nonattentional streams** in vision

These streams not primarily concerned with explicit perception of change (coherence theory)

-> Mapped out via **implicit** detection of change?

108

109

4.1 Implicit Detection of change — visuomotor

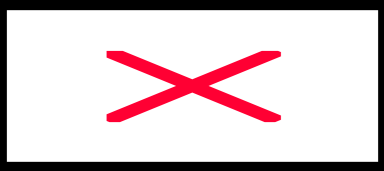
- Bridgeman et al. (1975) — **oculomotor response**
  - target moves while observer saccades to it
  - eye makes corrective saccade, even though observers have no explicit perception of change
- Goodale et al. (1986) — **manual pointing**
  - target moves while observer saccades to it
  - hand reaching towards target corrects its trajectory, even though observers have no explicit perception of change

110

111

4.2 Implicit Detection of change — perceptual

- Fernandez-Duque & Thornton (2000)
  - observers view 2-display sequence; each display is a simple array of rectangles
  - observers tested on two items: the item changed, and the item diagonally across from it



  - If observer did not notice change, asked to **guess which item changed**.

112

113

Results

- Observers could guess **better than chance** (55-63%) even though change was not consciously noticed
  - (a form of blindsight in normal observers)
  - ➞ **involvement of limited-capacity system**
- No attentional priming at location of unnoticed change
  - ➞ **involvement of purely nonattentional system**

114

115

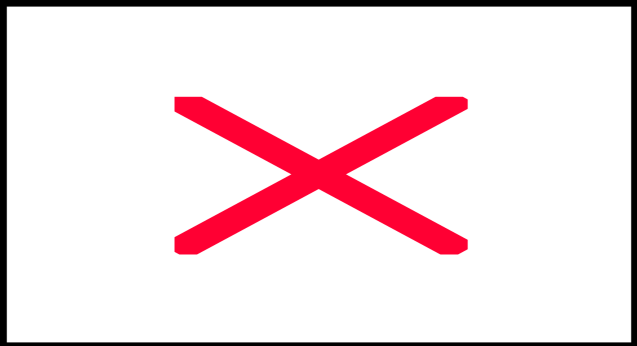
4.3. Visual awareness without visual experience

Origin - spontaneous reports by observers in original flicker experiments that they were **aware** of the change long before they **visually experienced** it.

116

117

- Rensink (1998, 2000)
  - observers view continuous flicker sequence (natural images)
  - asked to hit button (t1) when change was **felt**
  - then hit button (t2) when change was **seen**



118

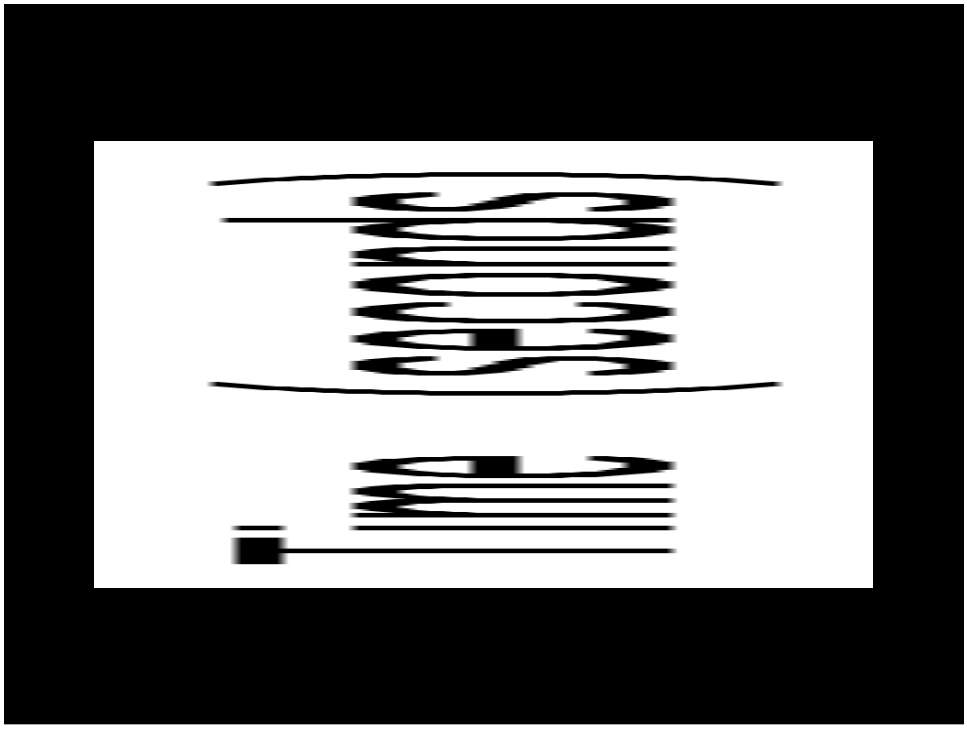
119

Results

- 1/2 of observers had no feeling of change without visual experience of it
- 1/4 of observers could feel a change before seeing it
  - (t2-t1) > 1 second on 20% of trials
  - average duration = 3.7 seconds
    - increases to 8.1 seconds if color field used
- not a result of guessing:
  - accuracy on catch trials is good (82%)

120

121



122

123

Not a "weakened" form of seeing:

- response times for sensing are **later** than for seeing
- behavior different for different kinds of change:
  - RTs for color changes exactly the same
  - RTs for location changes are 3.5 seconds later for sensing ($p < .005$)

Mindsight: Conscious (mental) awareness without an accompanying visual experience

124

125

Possible Explanation

- Mindsight observers have no attentional problems
  - responses generally as fast as for other observers

Mindsight due to nonattentional system (alert?)
-observers experiencing mindsight take longer to disengage from it?

126

127

Important points (Section 4)

- There exists a limited ability to detect change by **nonattentional perceptual** systems
- Detection of change is **nonexperiential**:
  - no conscious awareness at all (cf. blindsight)
  - aware, but no visual experience (mindsight)

128

129

Summing Up…

- **Change blindness**
  - attention needed to see change
- **Representation of scenes**
  - virtual representation (based on nonattentional systems)
- **Attentional system**
  - 4 links pooling into 1 nexus; binding problem avoided
- **Nonattentional systems**
  - implicit perception of change; mindsight

130