Lecture 2: Quantifiers and Approximation

Case study: Most vs More than half

Jakub Szymanik

university of groningen

Outline

Outline

Departure point

Observation

Quantifiers are associated with truth-conditions.

Departure point

Observation

Quantifiers are associated with truth-conditions.

Observation

Quantifiers are associated with verification strategies.

Number knowledge

Observation

Quantifiers embed number knowledge in language.

Number knowledge

Observation

Quantifiers embed number knowledge in language.

Question

What is the characteristics of that ability in humans?

Number Sense

Definition

An intuitive understanding of numbers, their magnitude, relationships, and how they are affected by arithmetical operations.

Dehaene, The Number Sense: How the Mind Creates Mathematics, 1999

Two Core Systems of Numbers

1. Approximate representations
2. Precise representations

Feigenson et al., Core Systems of Numbers, Trends in Cognitive Science, 2004

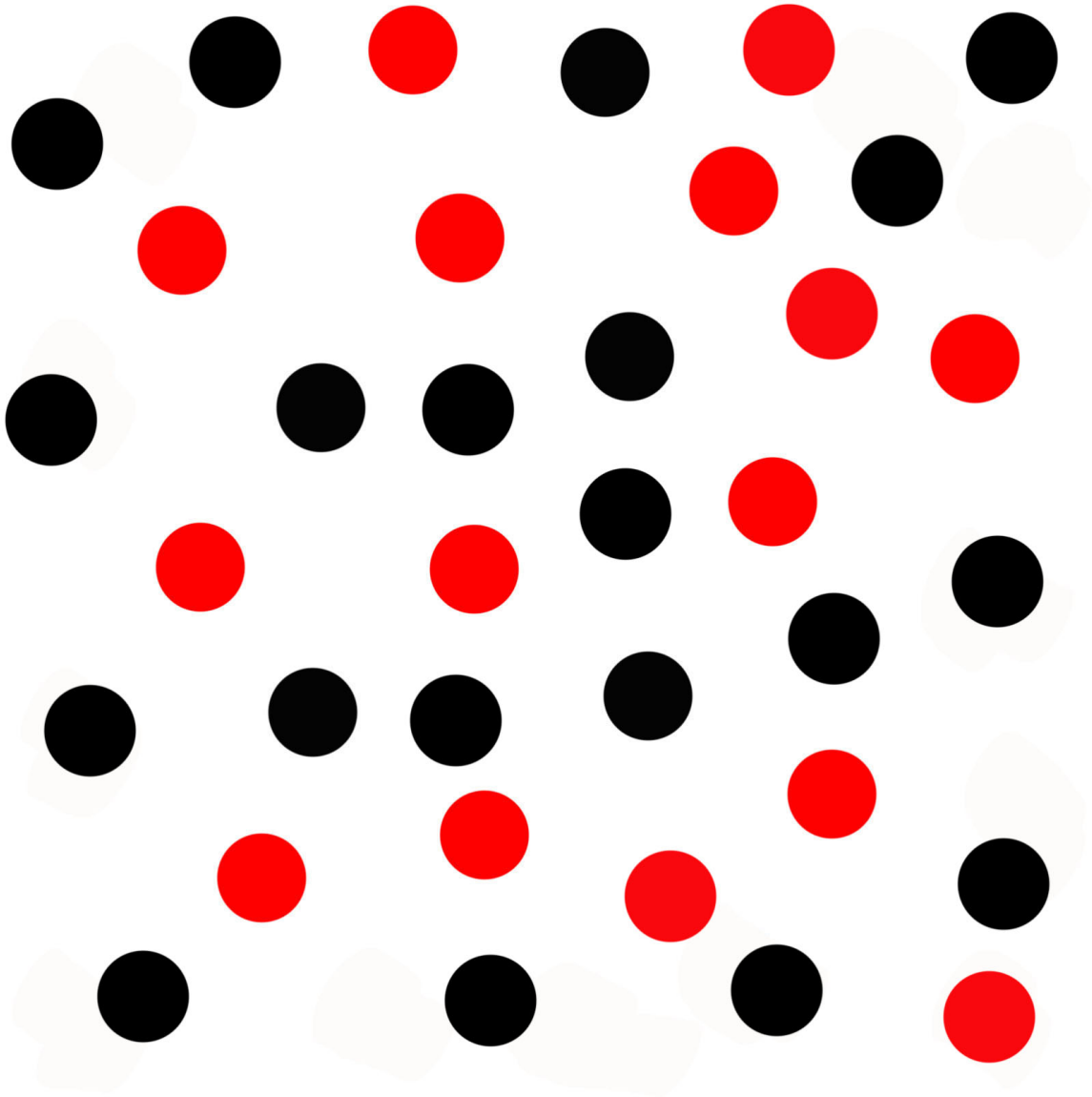
Outline

Test Your Approximate Number Sense

Test

Approximate Number Sense (ANS)

Approximate Number Sense (ANS)



ANS Cont.

- common to many nonverbal animals;

ANS Cont.

- common to many nonverbal animals;
- an evolutionary ancient cognitive resource;

ANS Cont.

- common to many nonverbal animals;
- an evolutionary ancient cognitive resource;
- generates representations of numerosity;

ANS Cont.

- common to many nonverbal animals;
- an evolutionary ancient cognitive resource;
- generates representations of numerosity;
- across multiple modalities;

ANS Cont.

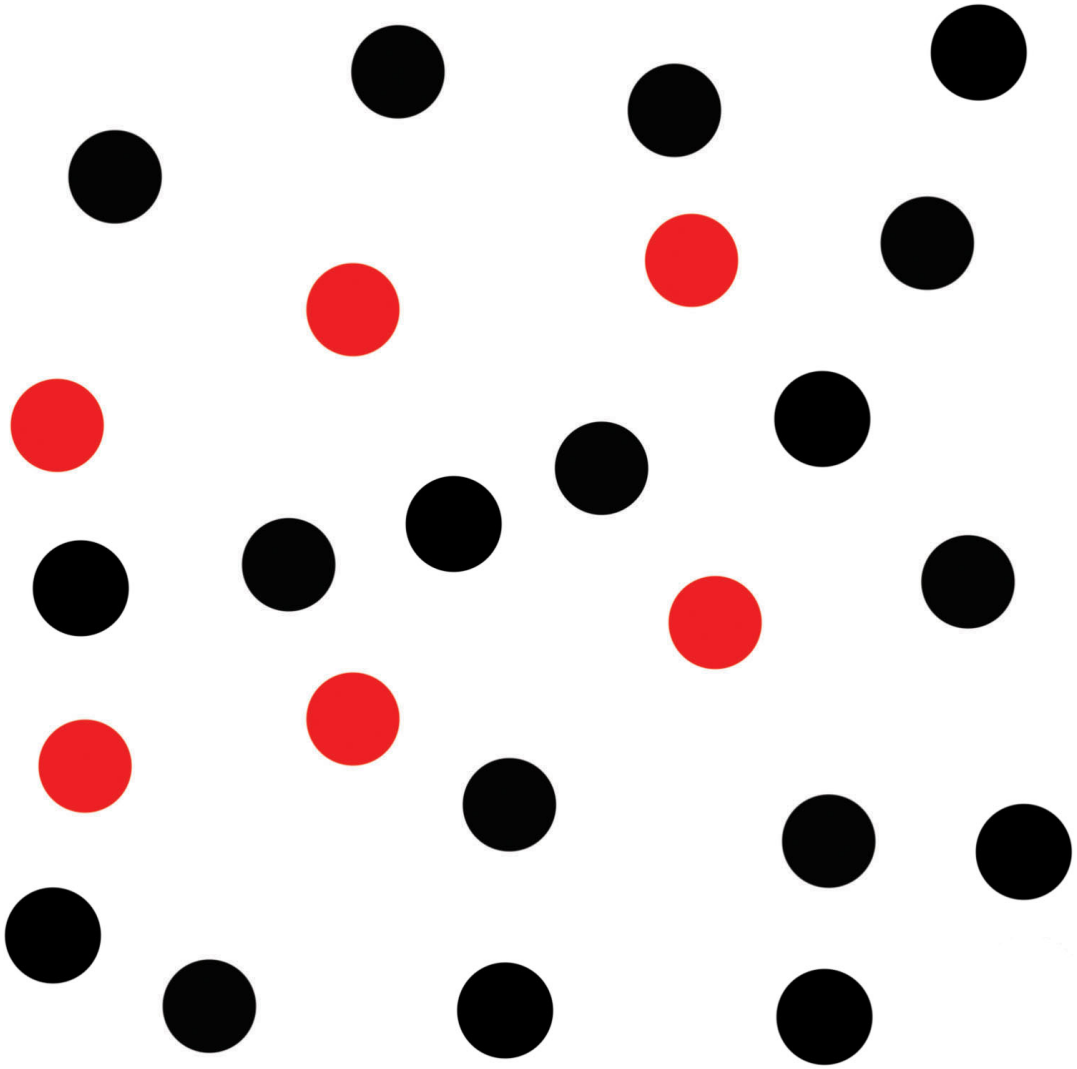
- common to many nonverbal animals;
- an evolutionary ancient cognitive resource;
- generates representations of numerosity;
- across multiple modalities;
- doesn't require explicit training;

ANS Cont.

- common to many nonverbal animals;
- an evolutionary ancient cognitive resource;
- generates representations of numerosity;
- across multiple modalities;
- doesn't require explicit training;
- but works within certain limits.

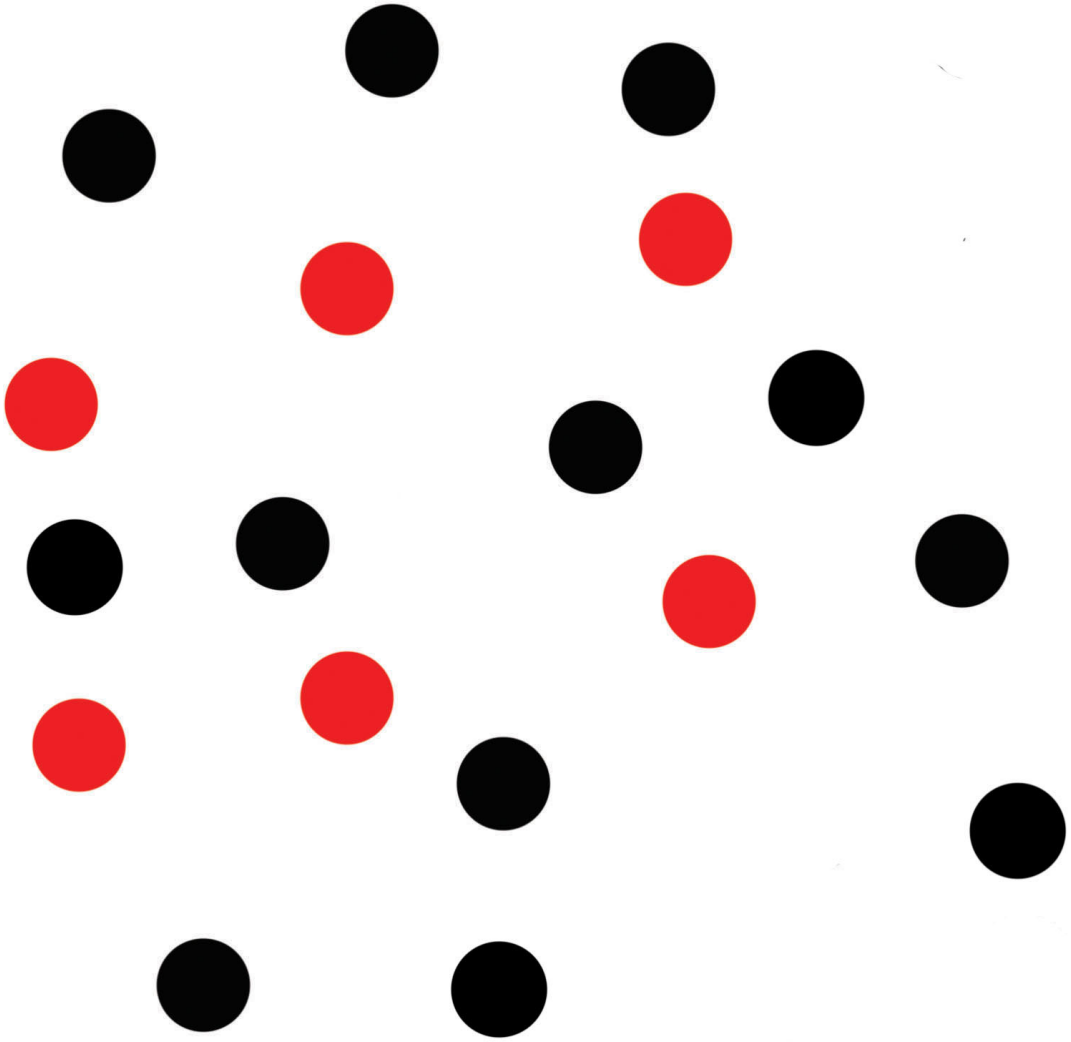
Distance-effect

Distance-effect



Distance-effect

Distance-effect



Distance-effect Cont.

Observation

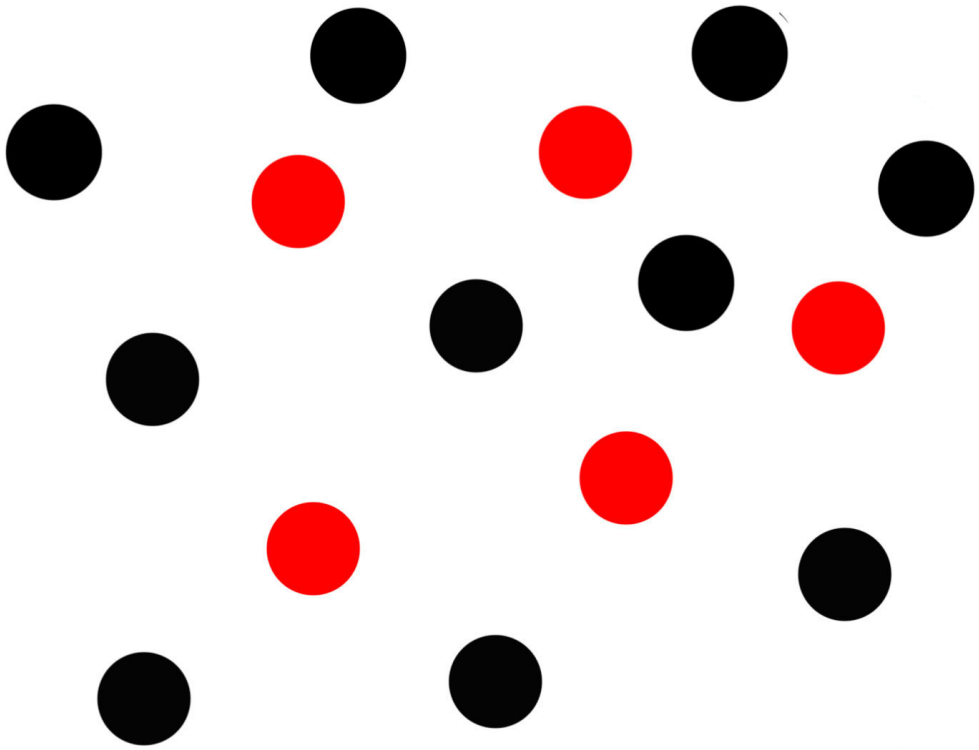
There is a systematic, monotonous decrease in numerosity discrimination performance as the numerical distance between the numbers decreases.

Example

From 6 vs 18 to 6 vs 12.

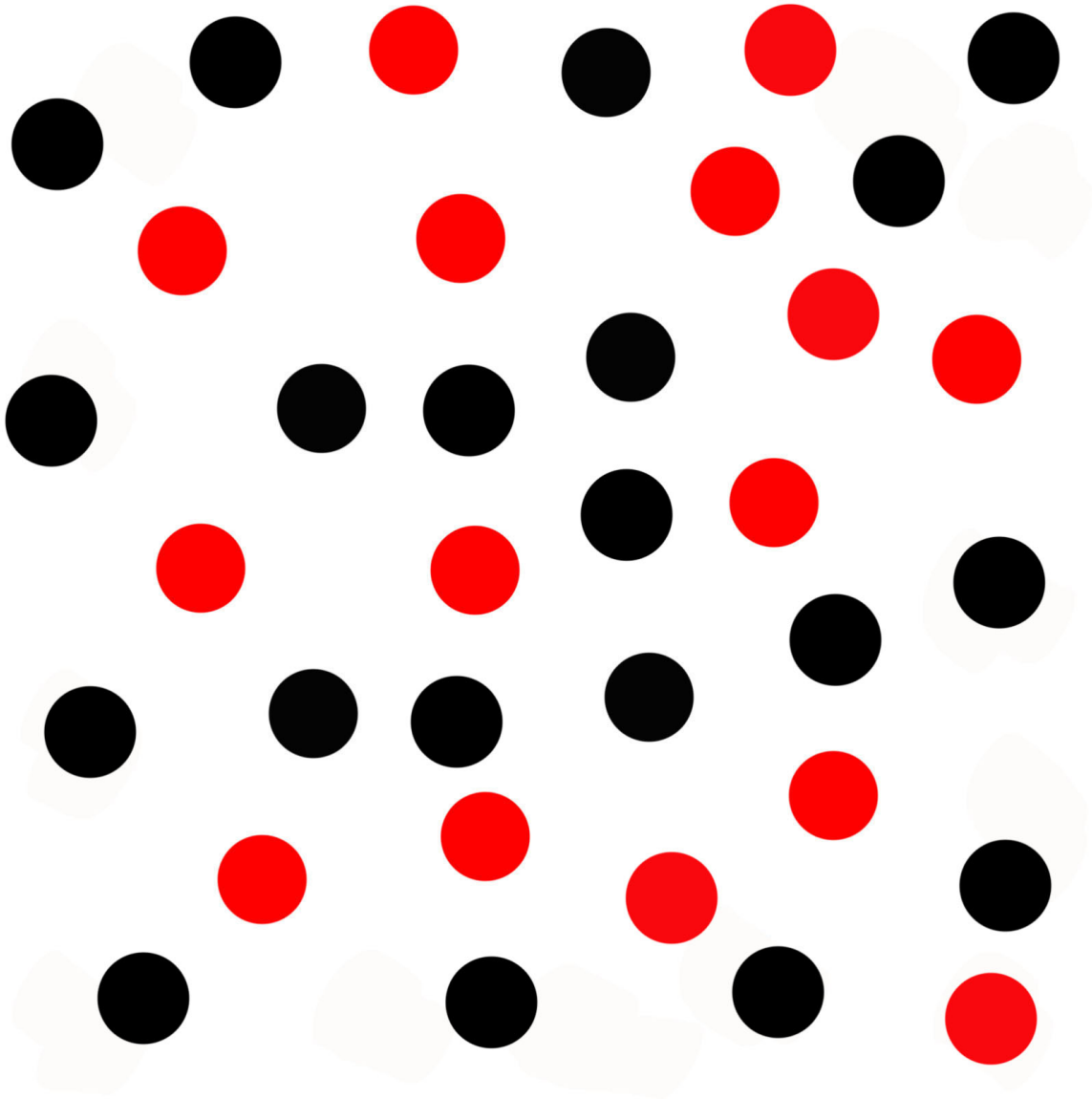
Size-effect

Size-effect



Size-effect

Size-effect



Size-effect Cont.

Observation

For equal numerical distance, performance also decreases with increasing number size.

Size-effect Cont.

Observation

For equal numerical distance, performance also decreases with increasing number size.

Example

5 things are detectable different from 10 then 15 from 20.

Weber's Law

Weber's Law

Discriminability depends on the ratio of relevant representational values.

Weber's Law

Weber's Law

Discriminability depends on the ratio of relevant representational values.

$$WR = \frac{\text{larger set}}{\text{smaller set}}$$

Weber's Law

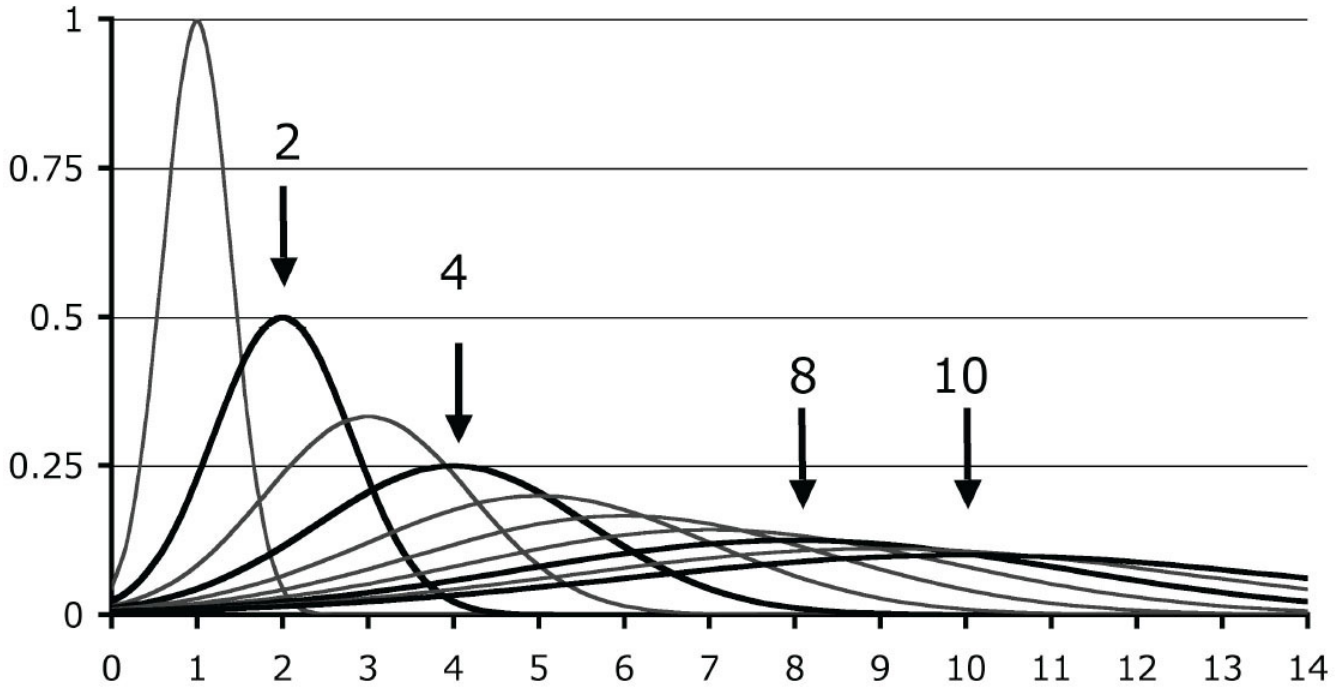
Weber's Law

Discriminability depends on the ratio of relevant representational values.

$$WR = \frac{\text{larger set}}{\text{smaller set}}$$

It explains ratio dependent performance.

Mental Number Line



The more overlap, the poorer the discriminability

Development of ANS

- The acuity of ANS improves during childhood.

Development of ANS

- The acuity of ANS improves during childhood.
- Adults can discriminate 7:8 ratios.

Development of ANS

- The acuity of ANS improves during childhood.
- Adults can discriminate 7:8 ratios.
- But 6-months infants only 1:2.

ANS and Counting

- By 5yrs ANS is mapped onto discrete number words.

ANS and Counting

- By 5yrs ANS is mapped onto discrete number words.
- It is activated anytime we deal with numbers.

ANS and Counting

- By 5yrs ANS is mapped onto discrete number words.
- It is activated anytime we deal with numbers.
- Distance and size effect for Arabic numerals!

Outline

Meanings of ‘most’

$$\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}$$

Meanings of ‘most’

$$\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}$$

$$\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \text{card}(A - B)$$

Yet Another Meaning of ‘most’

$$\text{most}[A, B] = 1 \text{ iff } \mathit{OneToOnePlus}(A \cap B, A - B)$$

Definition

$\mathit{OneToOnePlus}(X, Y) \iff \exists X' \subseteq X$ s.t. there is a one-to-one function between X' and Y but not X and Y.

Corresponding Procedures

They suggest different strategies:

Corresponding Procedures

They suggest different strategies:

- $\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}$
- Comparing number of target dots to the half of all dots.

Corresponding Procedures

They suggest different strategies:

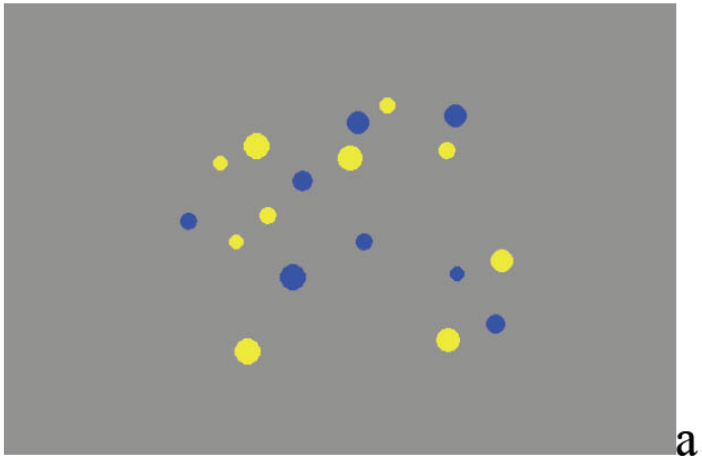
- $\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}$
- Comparing number of target dots to the half of all dots.
- $\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \text{card}(A - B)$
- Comparing blue and yellow dots directly.

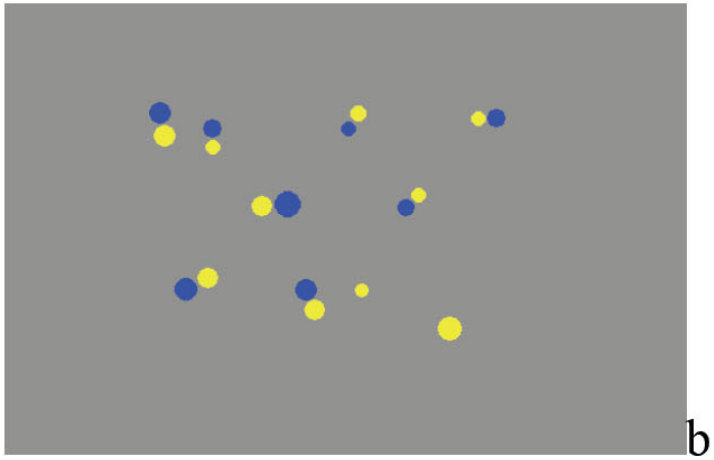
Corresponding Procedures

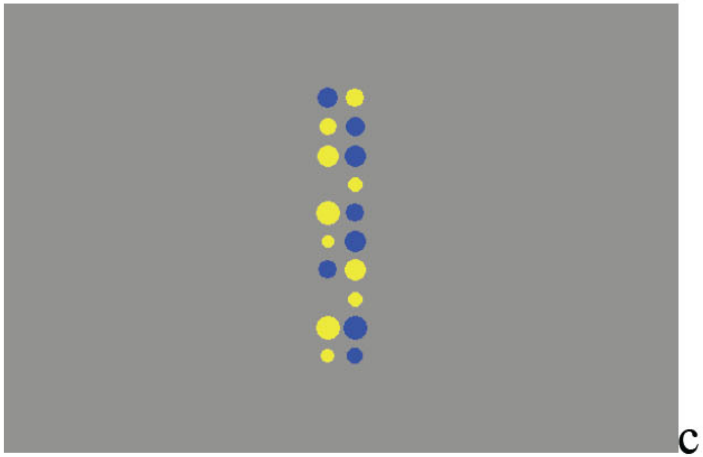
They suggest different strategies:

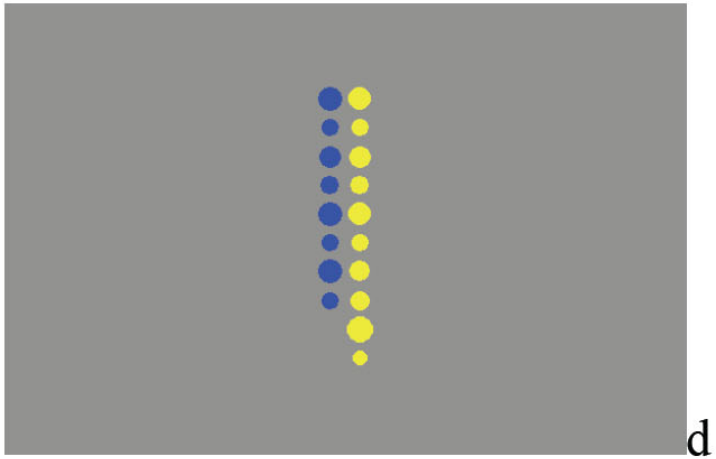
- $\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}$
- Comparing number of target dots to the half of all dots.
- $\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \text{card}(A - B)$
- Comparing blue and yellow dots directly.
- $\text{most}[A, B] = 1 \text{ iff } \mathit{OneToOnePlus}(A \cap B, A - B)$
- Searching for 1-1 map.

Triggers


a


b


c


d

Experimental Questions

Question

Do people use OneToOnePlus strategy?

Experimental Questions

Question

Do people use OneToOnePlus strategy?

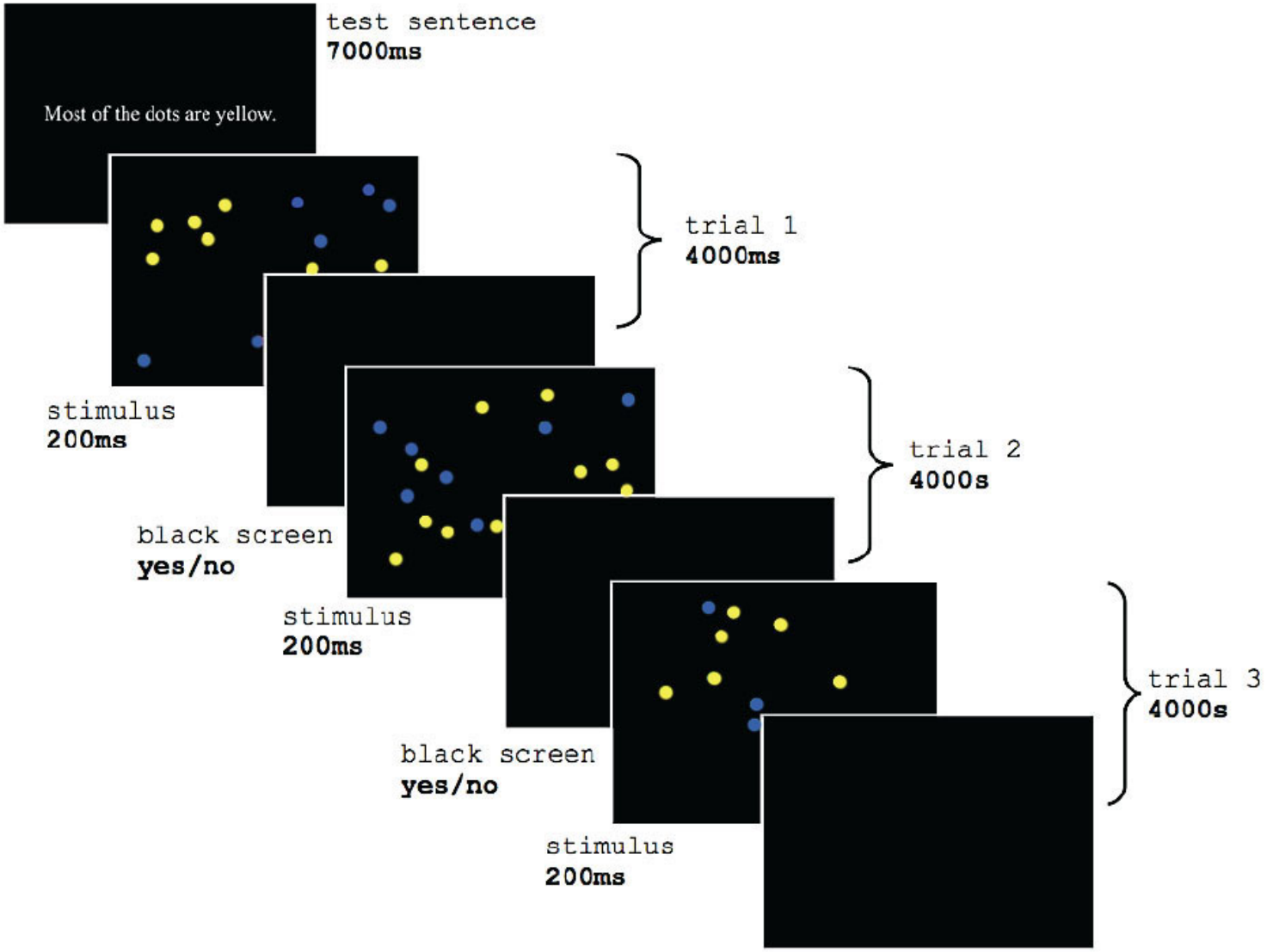
Question

Do people use ANS to judge truth-value?

Dehaene & Cohen, Cultural Recycling of Cortical Maps, Neuron, 2007

Pietroski et al., The Meaning of ‘Most’: semantics, Numerosity, and Psychology, Mind and Language, 1999

Pietroski's et al. Experimental Design



Experimental Design cont.

- 12 subjects

Experimental Design cont.

- 12 subjects
- 360 trials each

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10
- Area-controlled vs. size-controlled

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10
- Area-controlled vs. size-controlled
- 10 trials in each bin for 4 conditions:
  - Scattered Random

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10
- Area-controlled vs. size-controlled
- 10 trials in each bin for 4 conditions:
  - Scattered Random
  - Scattered Pairs

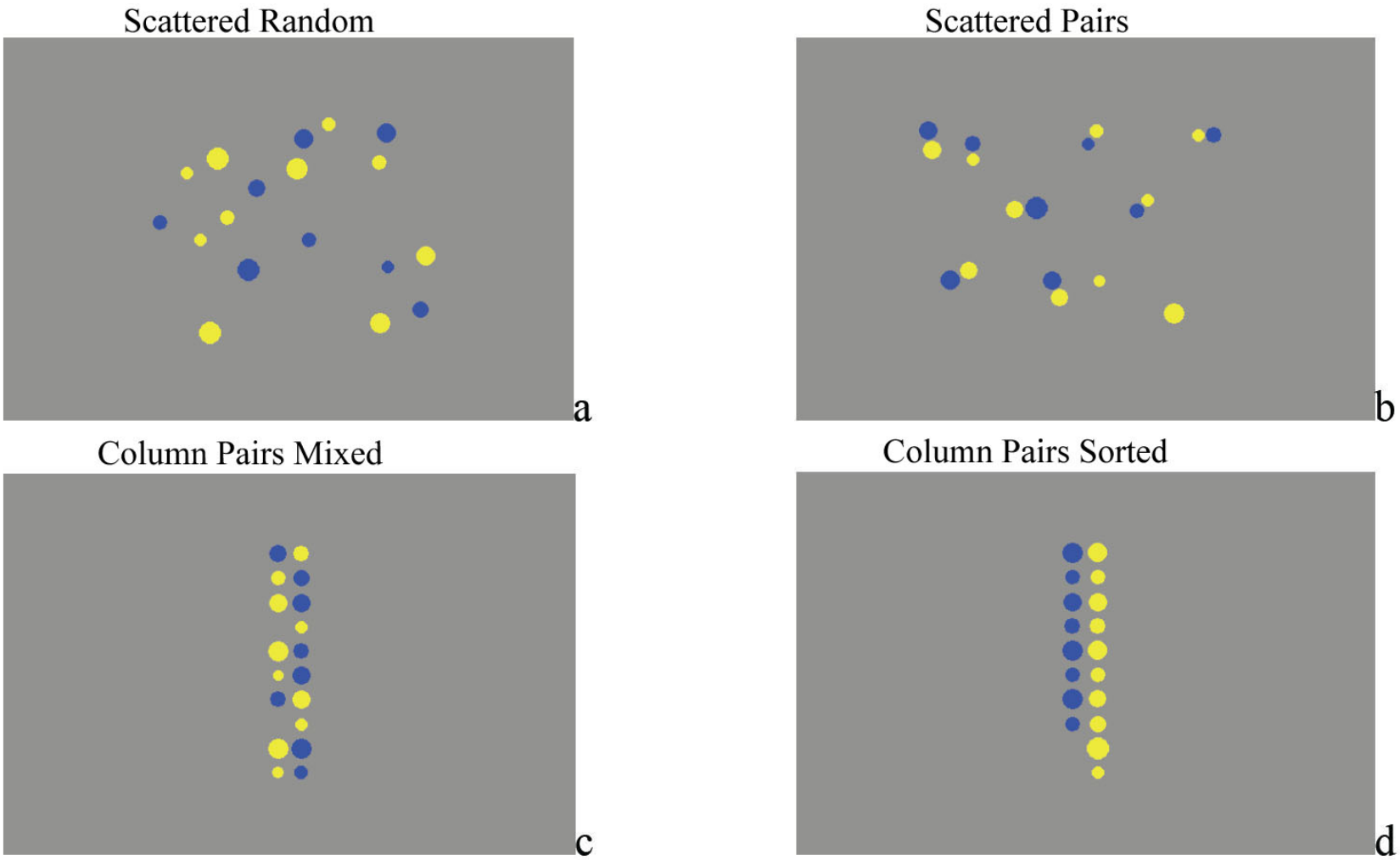
Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10
- Area-controlled vs. size-controlled
- 10 trials in each bin for 4 conditions:
  - Scattered Random
  - Scattered Pairs
  - Column Pairs Mixed

Experimental Design cont.

- 12 subjects
- 360 trials each
- 2 color dots: blue and yellow
- The number of dots in each color: 5–17
- 9 bins: 1:2, 2:3, 3:4, 4:5, . . . , 9:10
- Area-controlled vs. size-controlled
- 10 trials in each bin for 4 conditions:
  - Scattered Random
  - Scattered Pairs
  - Column Pairs Mixed
  - Column Pairs Sorted

Trial Types


Scattered Random
a
Scattered Pairs
b
Column Pairs Mixed
c
Column Pairs Sorted
d

Results

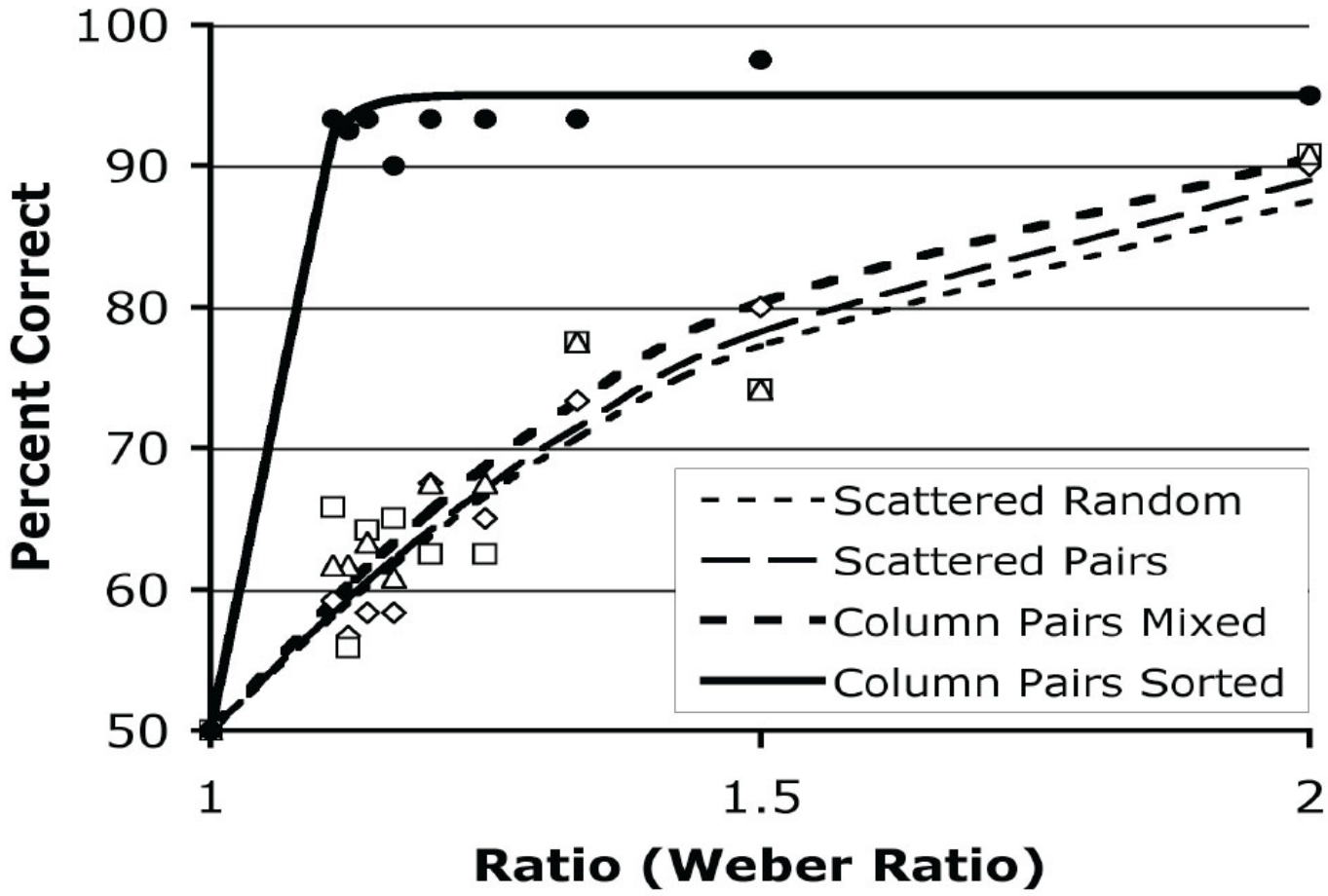
- Participants did better with larger ratios.
- They did best on Column Pairs Sorted,
- but no significant differences among other trial types.
- No influence of size- or area-control.

Results

- Participants did better with larger ratios.
- They did best on Column Pairs Sorted,
- but no significant differences among other trial types.
- No influence of size- or area-control.

Conclusion

OneToOnePlus is out.


100
90
80
70
60
50
Percent Correct
1
1.5
2
Ratio (Weber Ratio)
Scattered Random
Scattered Pairs
Column Pairs Mixed
Column Pairs Sorted

Agreement with ANS

Observation

Agreement between the ANS and performance.

Agreement with ANS

Observation

Agreement between the ANS and performance.

Conclusion

Participants relied on ANS to evaluate 'most'.

Follow-up Question

Question

How is the cardinality of the non-blue set estimated?

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \overline{\mathit{Yellow}})$$

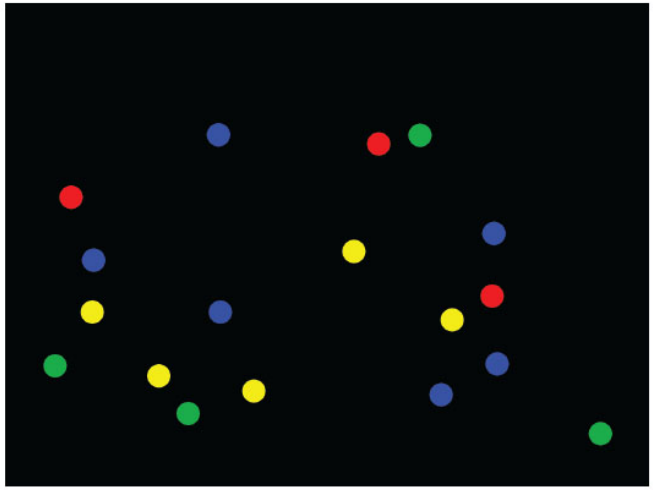
Lidz et al., Interface Transparency and the Psychosemantics of ‘most’, Natural Language Semantics, in press

2 Strategies: Selection

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \mathit{Blue}) \; + \; \text{card}(\mathit{Dot} \cap \mathit{Red}) + \\ \ldots \; + \; \text{card}(\mathit{Dot} \cap \mathit{Green})$$

2 Strategies: Selection

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \mathit{Blue}) + \text{card}(\mathit{Dot} \cap \mathit{Red}) + \ldots + \text{card}(\mathit{Dot} \cap \mathit{Green})$$

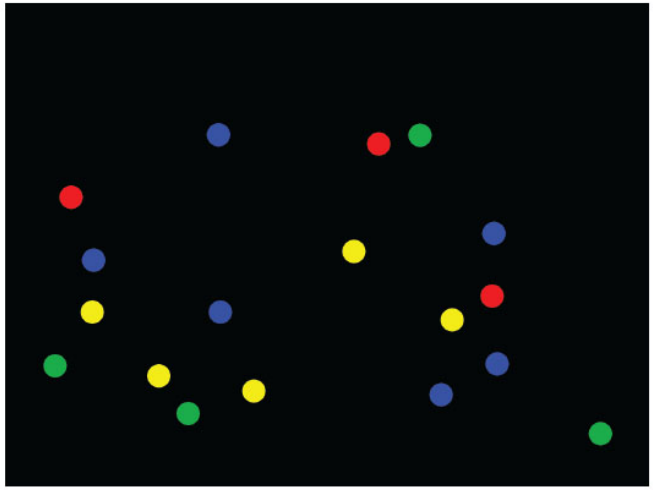


2 Strategies: Subtraction

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot}) - \text{card}(\mathit{Dot} \cap \mathit{Yellow})$$

2 Strategies: Subtraction

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot}) - \text{card}(\mathit{Dot} \cap \mathit{Yellow})$$



Comparing 2 Strategies

- For 2 color screens selection is easier.
- For multi-color screens subtraction might be more optimal.

Question

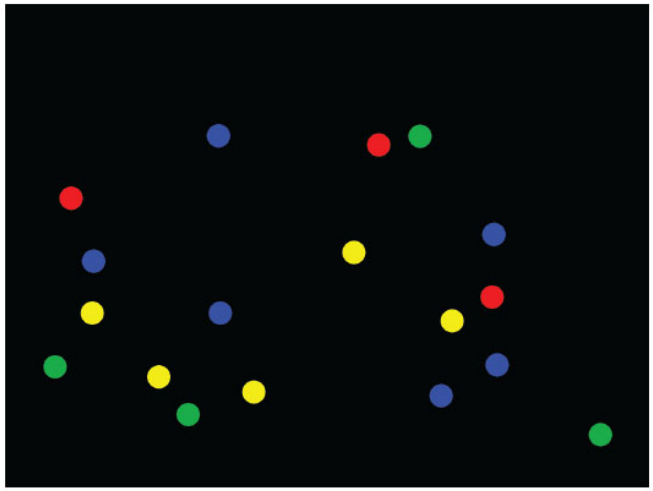
Is one of them the default strategy?

Restrictions of Our Visual System

- We can generate estimates only for up to 3 sets: the total set of dots and 2 color subsets.
- So, selection might be out as a non-universal strategy.

Restrictions of Our Visual System

- We can generate estimates only for up to 3 sets: the total set of dots and 2 color subsets.
- So, selection might be out as a non-universal strategy.

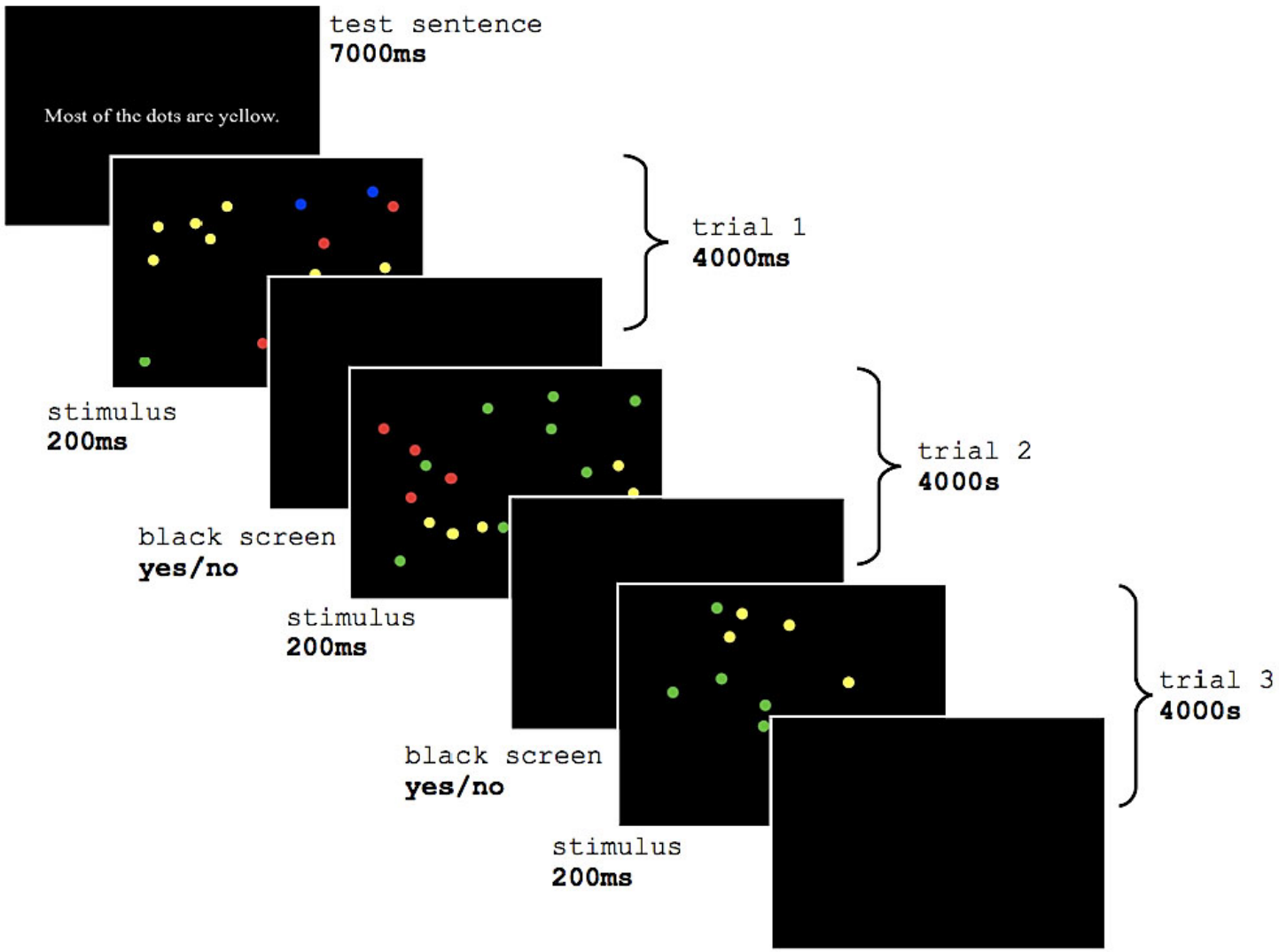


Halberda et al., Multiple Spatially Overlapping Sets Can be Enumerated in Parallel, Psychological Science, 2006. DEMO

Halberda's et al. 2006 Experiment

- 10 subjects;
- 450 trials each;
- 500 ms displays with 1-35 dots in 1-6 colors;
- probe before or after diverged from 3 colors on.

Lidz's et al. Experimental Design



Findings

Observation

There was no difference in accuracy as the function of number of colors on the display, but only as the function of the ratio.

Observation

Accuracy was not higher on 2 color screens.

Conclusion

Subtraction was always used.

Canonical Meaning of ‘most’

$$\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \text{card}(A) - \text{card}(A \cap B)$$

Discussion

Observation

For 2 colors selection is more efficient than subtraction but the later was nevertheless used.

> *. . . our data support Interface Transparency Thesis (ITT), according to which speakers exhibit a bias towards the verification procedures provided by canonical specification of truth-conditions. (Lidz et al., 2010)*

Meaning as a Collection of Procedures

Subtractions might be the default strategy only when:

- Brief flash;
- The same sentence over and over again;

Question

What are the other alternatives?

Outline

Tomaszewicz's Study

Most1: większość, 'majority'

więk-	*-sz-*	*-ość*
'many/great'	adjectival comparative '-er/more'	nominalizer

Most2: najwięcej, 'largest subset / the most'

naj-	*-więc-*	*-ej*
adverbial superlative '-est/most'	'many/great'	adverbial comparat. '-er/more'

$$\text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \mathit{Blue})\ \&$$
$$\&\ \text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \mathit{Red})\ \&$$
$$\ldots \&\ \text{card}(\mathit{Dot} \cap \mathit{Yellow}) > \text{card}(\mathit{Dot} \cap \mathit{Green})$$

(Stepwise Selection)

Experiments

- Design similar to Pietroski et al. 2009 and Lidz et al. 2010.
- 20 subjects on-line;
- 200ms;
- 2x180 displays (for *Most 1* and *Most 2*);
- 2-5 colors (1-3 distractors).

Results

- For *Most 1* the result of Lidz et al. and of Pietroski et al. were replicated.
- *Most 1* = most.

Results

- For *Most 1* the result of Lidz et al. and of Pietroski et al. were replicated.
- *Most 1* = most.
- For *Most 2* : effect of ratio and number of colors.
- Stepwise selection.
- Superlative meaning.

Results

- For *Most 1* the result of Lidz et al. and of Pietroski et al. were replicated.
- *Most 1* = most.
- For *Most 2* : effect of ratio and number of colors.
- Stepwise selection.
- Superlative meaning.
- Different accuracy patterns for *Most 1* and *Most 2* in the same 2-color displays.

Discrepancy with Halberda et al. 2006?

Question

So, can we attend more than 3 colors?

Hypothesis

2 different tasks:

1. *enumeration of subsets and*
2. *recognizing the largest one.*

Question

Does most have a superlative reading in English?

Kotek et al, *Most Meanings are Superlative, to appear in Syntax Semantics*

Main Claim

Hypothesis

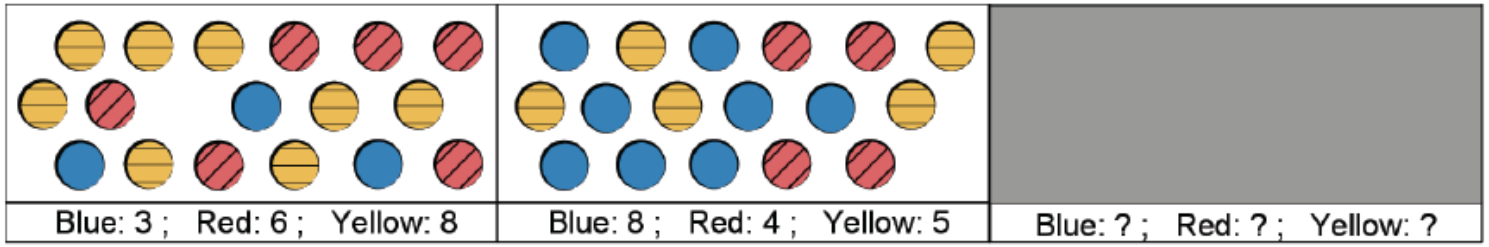
Most is ambiguous between:

1. *dominant proportional reading;*
2. *latent superlative reading.*

Can the speakers really access a superlative meaning?

Experiment 1: Covered Box

- A sentence is shown;
- with ‘more than half’ or ‘most’;
- and a picture:

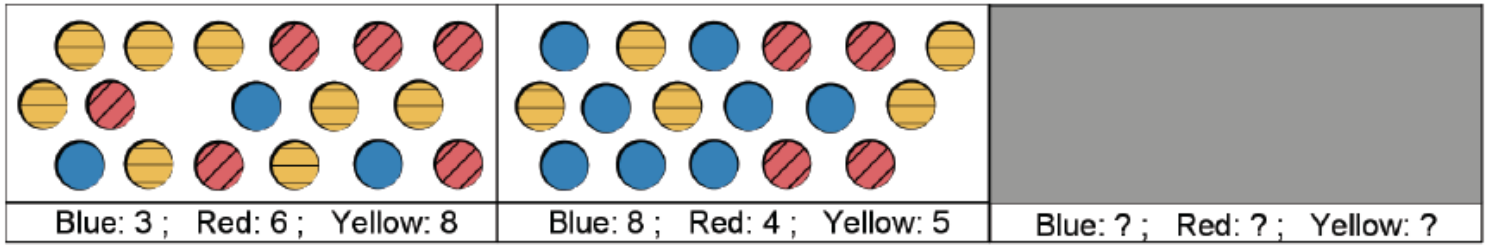


(a) False picture (b) Superlative picture (c) Covered box

- ‘Exactly one picture matches the sentence. Choose.’

Experiment 1: Covered Box

- A sentence is shown;
- with ‘more than half’ or ‘most’;
- and a picture:



(a) False picture (b) Superlative picture (c) Covered box

- ‘Exactly one picture matches the sentence. Choose.’

Results:

- Most: (c) 67.4%; (b) 32.6%.
- More than half: (c) 100%.

Meaning of ‘most’

$$\text{most} = \{\text{approximation+subtraction}, \text{stepwise selection (superlative)}, \ldots\}$$

Outline

Unchallenged Alternative

Why not:

$$\text{most}[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \frac{\text{card}(A)}{2}?$$

Children and ‘most’

Observation

Successful verification of ‘most’ in cases with two salient subsets is achieved at 3 years 7 months of age.

Observation

The comprehension is independent of knowledge of exact number words.

Halberda et al., The Development of ‘Most’ Comprehension and Its Potential Dependence on Counting Ability in Preschoolers, Language Learning and Development, 2008. DEMO

Question

What happens under a different experimental paradigm?

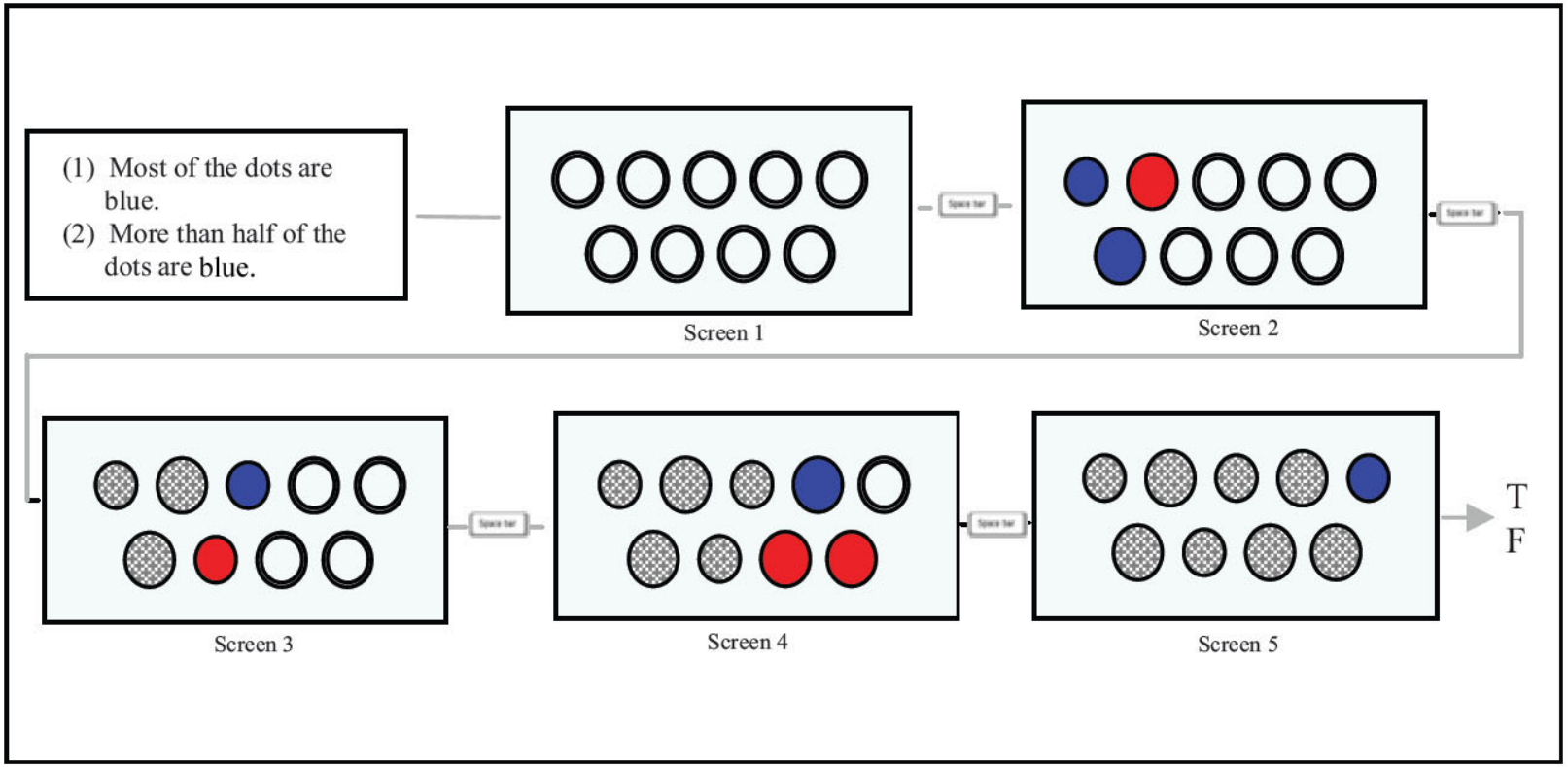
Hypothesis

$\mathrm{card}(A \cap B) > \frac{\mathrm{card}(A)}{2}$ *seems rather like 'more than half'?*

Question

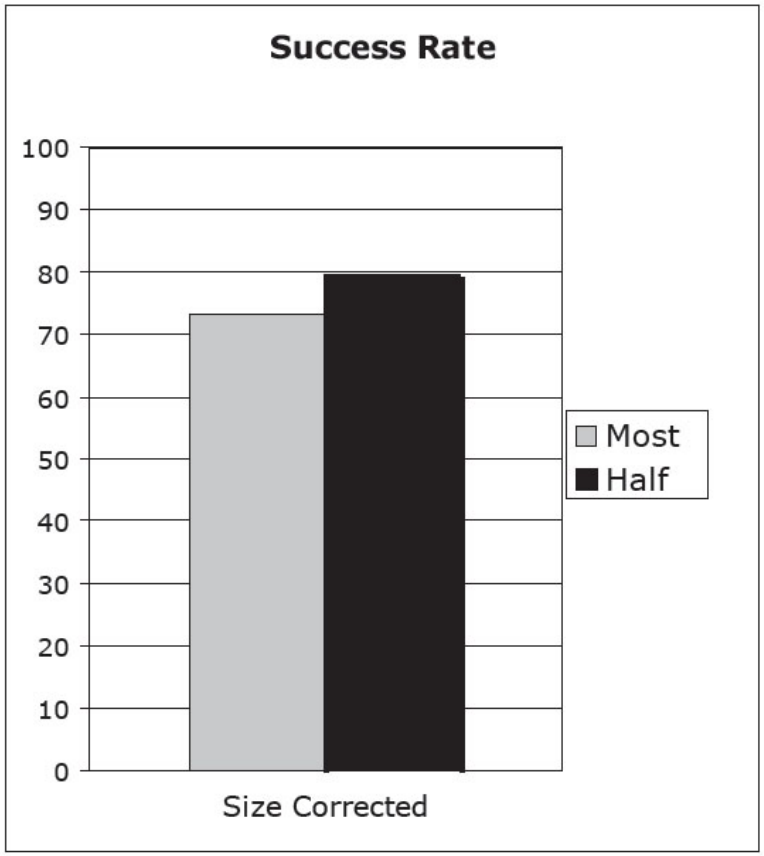
Are the strategies for 'most' and 'more than half' distinct?

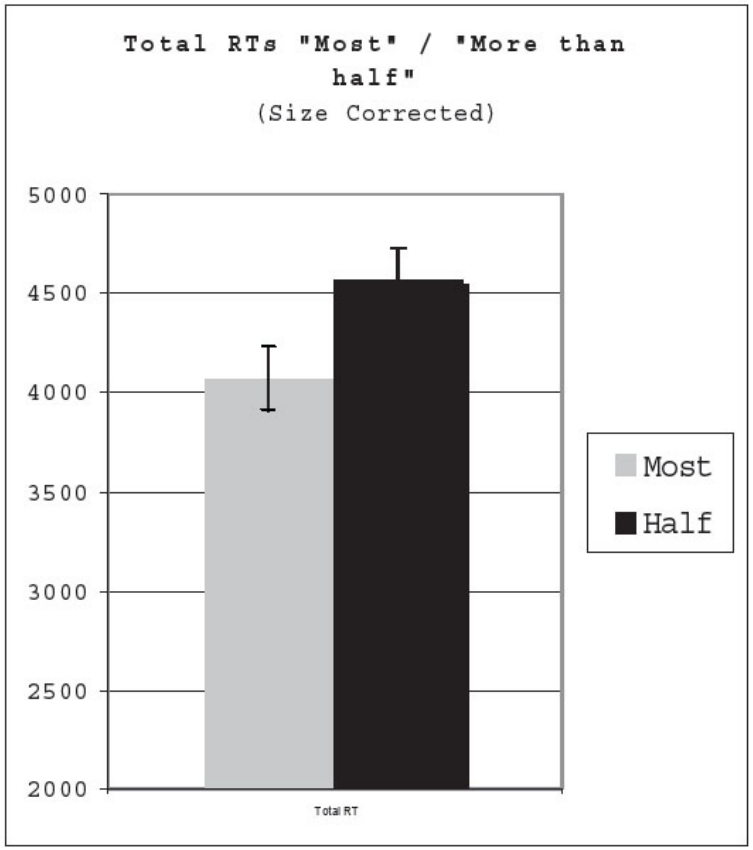
Hackl's Self-paced Counting



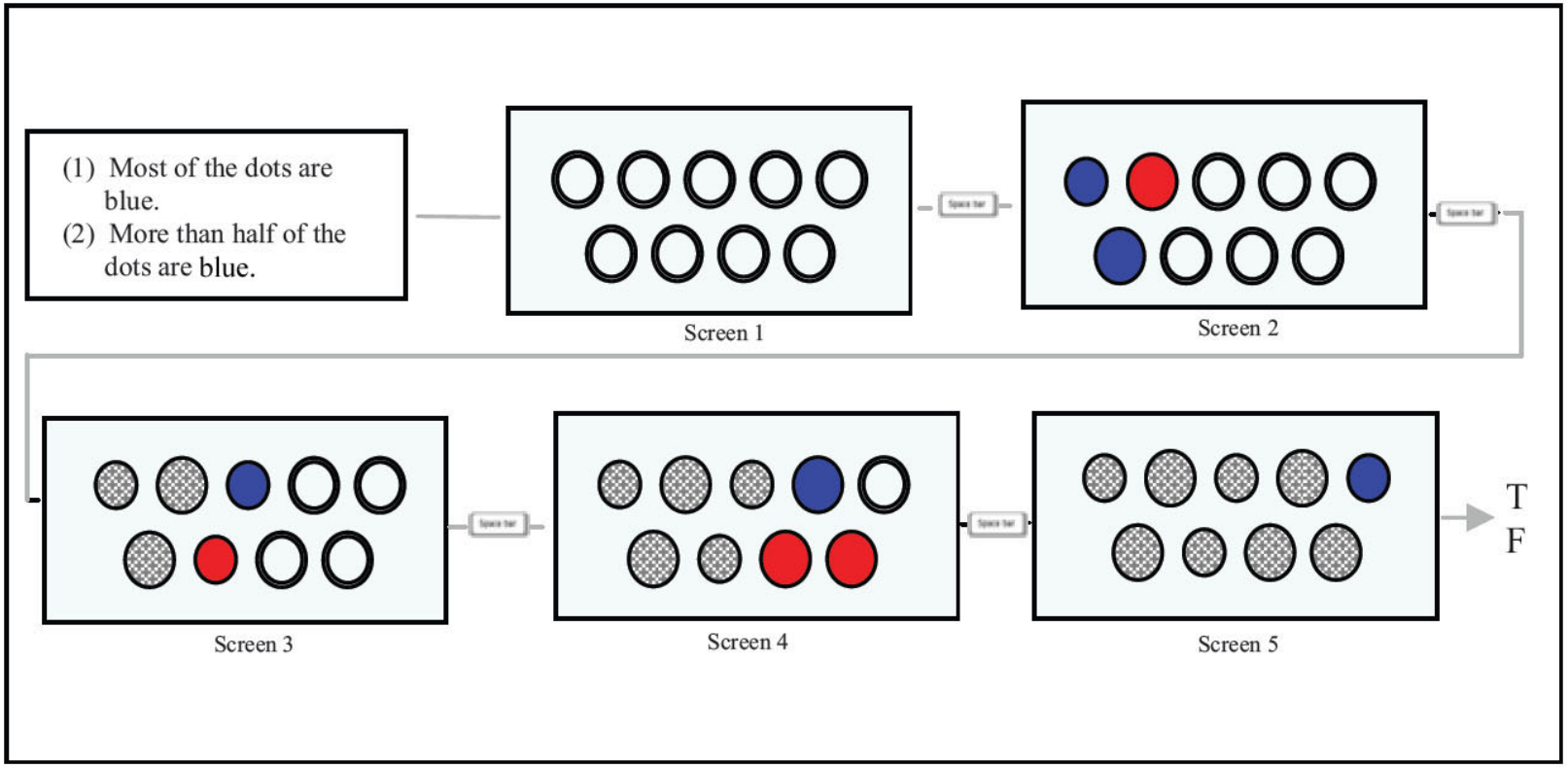
Hackl, On the Grammar and Processing of Proportional Quantifiers, Natural Language Semantics, 2009

Results Reported: Equivalence

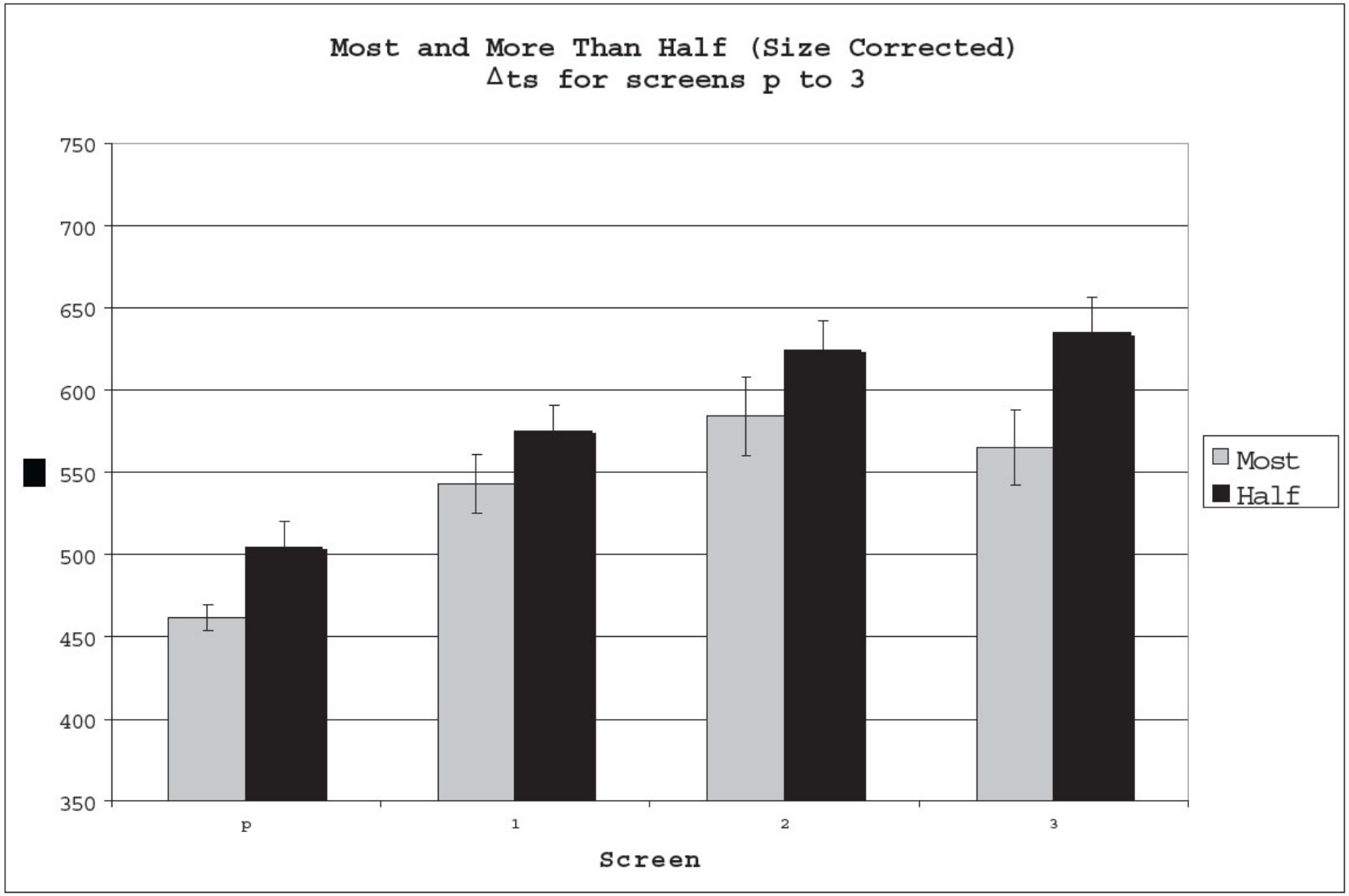




Let's Look Into The Verification Process



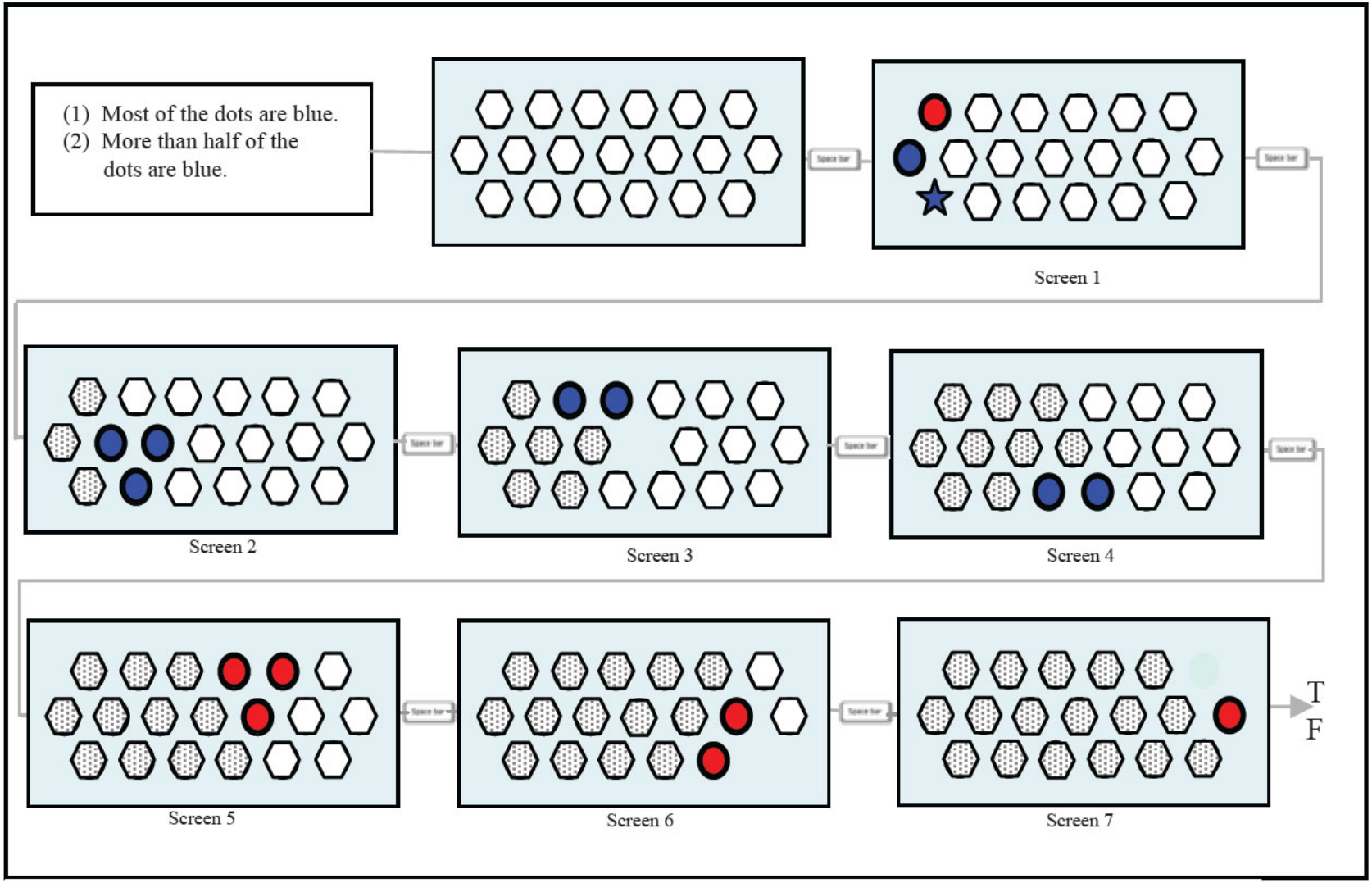
Results: Difference on 4 Screens



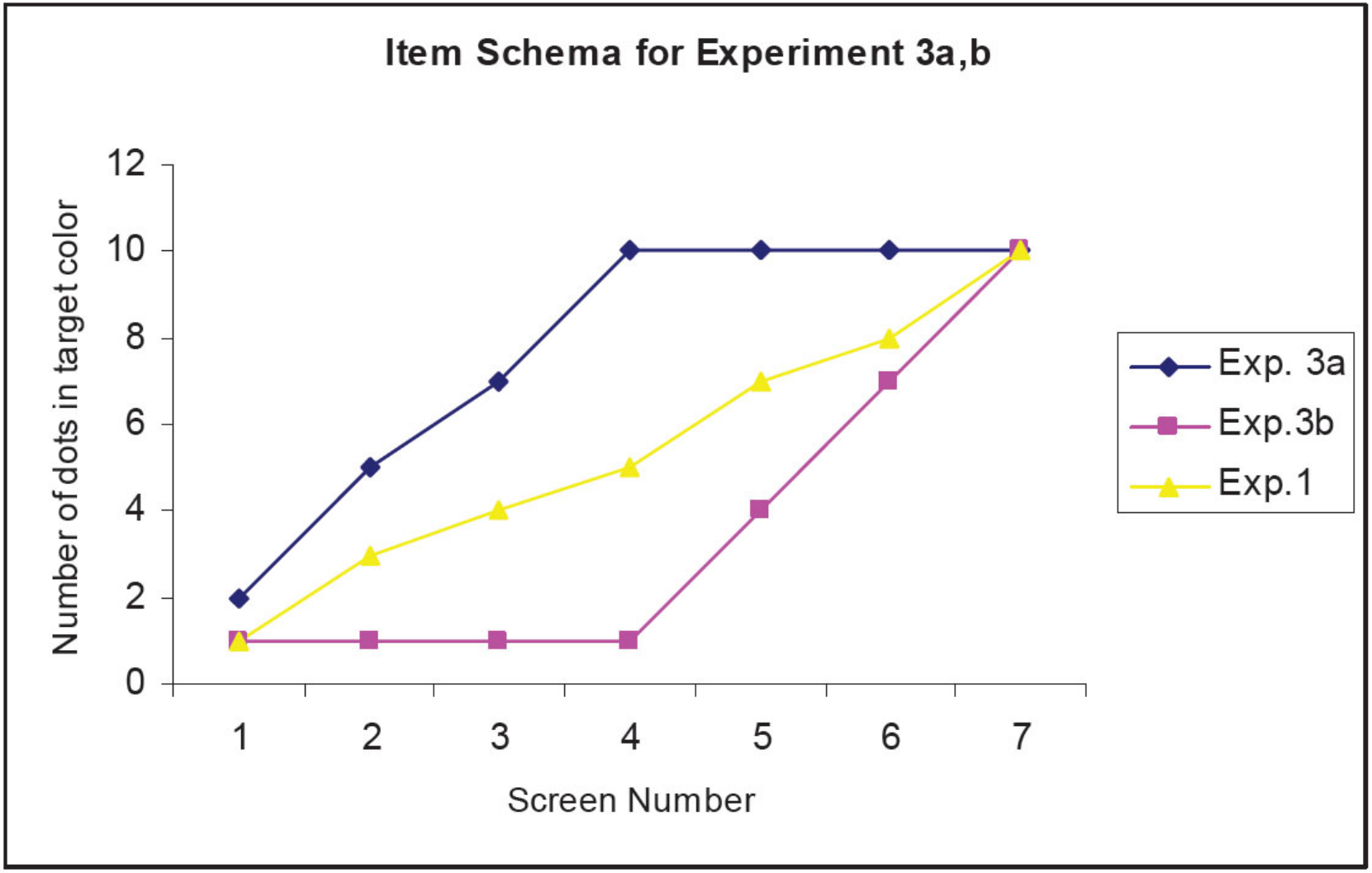
Discussion

- Speakers treat expressions as equivalent;
- but they use different verification strategies.

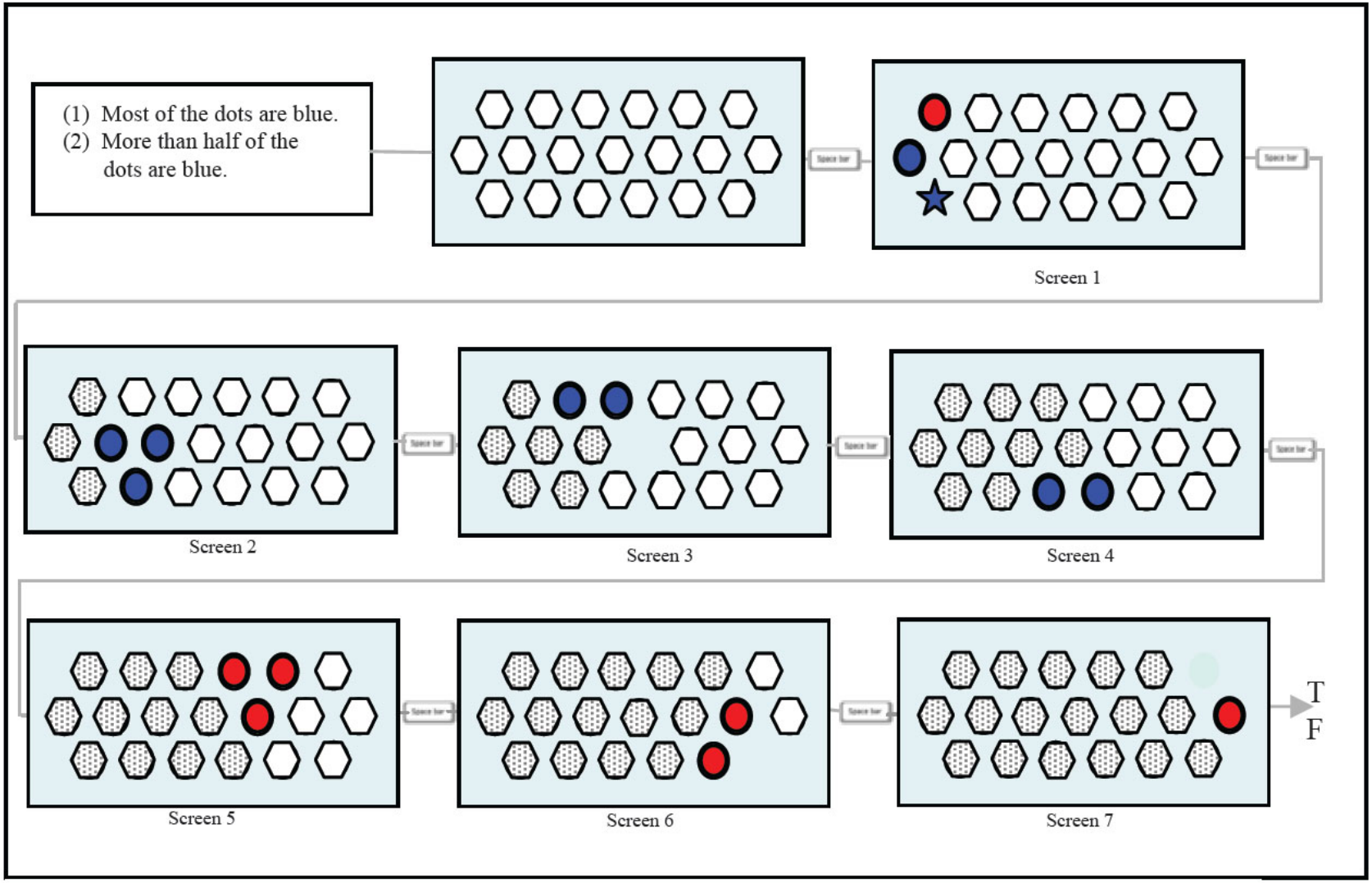
Modified SPC



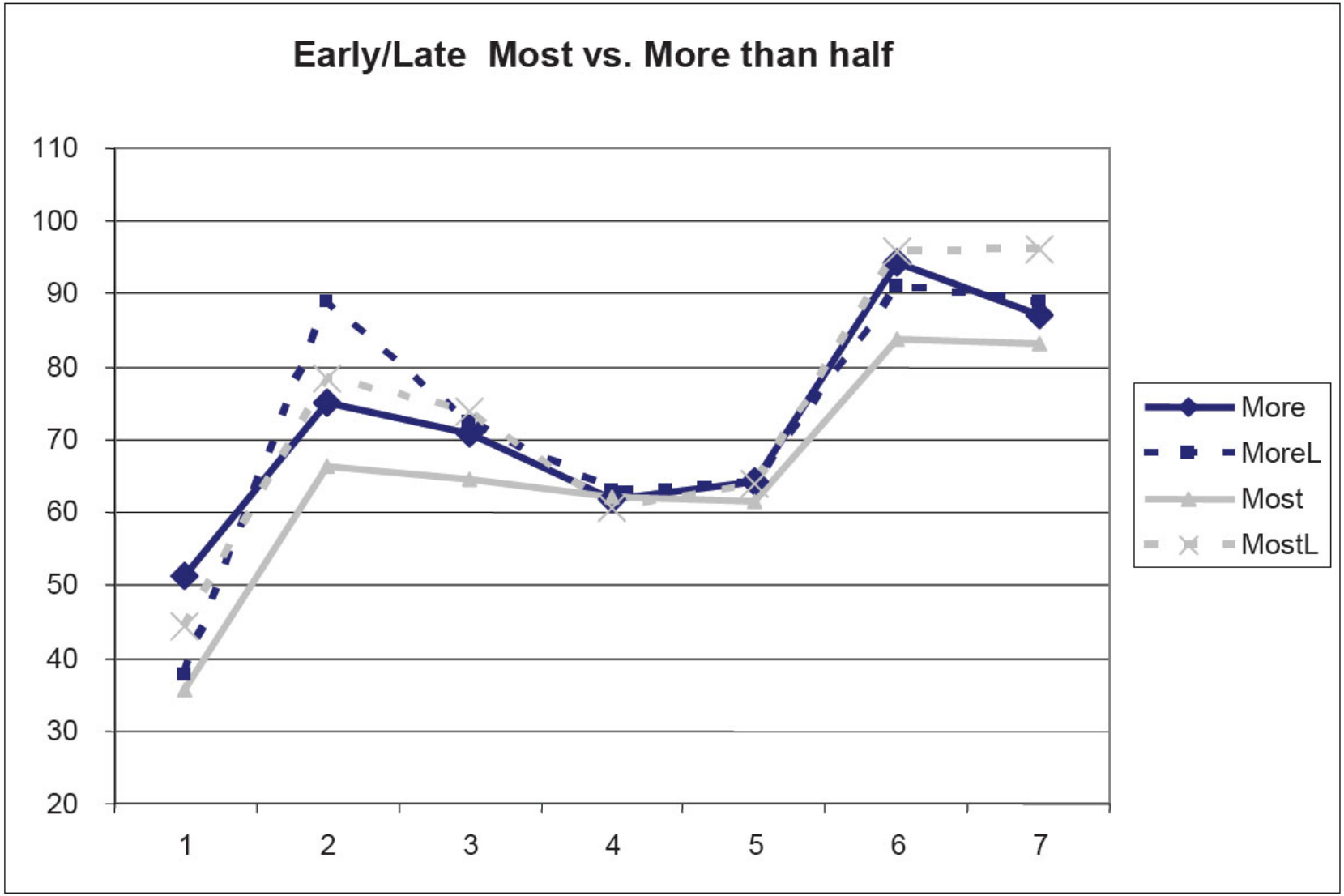
Manipulating Distribution in SPC



Modified SPC



Results


Early/Late Most vs. More than half
110
100
90
80
70
60
50
40
30
20
1
2
3
4
5
6
7
More
MoreL
Most
MostL

Discussion

- ‘most’ is more sensitive to distributional asymmetries.
- It needs more time in L-condition.

Discussion

- ‘most’ is more sensitive to distributional asymmetries.
- It needs more time in L-condition.
- ‘more than half’ is almost unaffected.

Discussion

- ‘most’ is more sensitive to distributional asymmetries.
- It needs more time in L-condition.
- ‘more than half’ is almost unaffected.

Hackl's Conclusions

- 'most' triggers 'lead counting':
  - How much target color leads at every screen.
  - Harder if it falls significantly behind.

Hackl's Conclusions

- 'most' triggers 'lead counting':
  - How much target color leads at every screen.
  - Harder if it falls significantly behind.
- but determining $\frac{A}{2}$ is insensitive to that.

Hackl's Conclusions

- 'most' triggers 'lead counting':
  - How much target color leads at every screen.
  - Harder if it falls significantly behind.
- but determining $\frac{A}{2}$ is insensitive to that.
- So:

  more than half$[A, B] = 1$ iff $\text{card}(A \cap B) > \frac{\text{card}(A)}{2}$

However,. . .

Question

1. *It cries out for comparison with non-SPC situation?*
2. *Aren't late and early in fact symmetric for 'lead counting'?*
3. *Is the estimation of $\frac{A}{2}$ really insensitive?*
4. *What about carrying the memorized values over?*

Let's us investigate the precise strategies closer!