Distinct Verification Strategies for *Most* and *More Than Half*:

Experimental Evidence for a Decompositional Analysis of Quantificational Determiners

Martin Hackl & Ben Acland
Pomona College

POMONA COLLEGE

Introduction

- Semantic analyses of a linguistic expression α typically address two questions:
  - What is the truth-conditional import of α ?
  - How much of the truth-conditional import is visible to the computational engine? What are the combinatorial properties of α?

Example:

(1) [[Most students are sick.]] = 1 iff $|ST \cap S| > |ST - S|$

(2) [[Most]] (A)(B) = 1 iff $|A \cap B| > |A - B|$

- Semanticists typically do not ask how truth-conditions are used by language external systems of the mind.

Introduction (cont.)

Example (cont.):

(3) [[More than half of the students are sick.]] = 1 iff $|ST \cap S| > \frac{1}{2} |ST|$

(4) [[More than half]] (A)(B) = 1 iff $|A \cap B| > \frac{1}{2} |A|$

Note: $|A \cap B| > |A - B| \Leftrightarrow |A \cap B| > \frac{1}{2} |A|$

- Does it matter how we state the truth-conditional import of *most*?
- How can we decide between competitors?

Introduction (cont.)

- Language internal reasons:

Better correspondence between LF of sentence that contains *most* and $|A \cap B| > |A - B|$ and *more than half and* $|A \cap B| > \frac{1}{2} |A|$
≈> Relies on a decomposition of *most* into MANY+est.

- Language external reasons:

Establish that $|A \cap B| > |A - B|$ and $|A \cap B| > \frac{1}{2} |A|$ are treated differently by some language external system.

Show that *most* triggers $|A \cap B| > |A - B|$ while *more than half* triggers $|A \cap B| > \frac{1}{2} |A|$

Introduction (cont.)

- Why does it matter?
  - If we have to distinguish between (1) and (2) we need a theory of quantification that is sensitive enough to the form of determiners to distinguish e.g. *most* and *more than half* in terms of their semantic components.

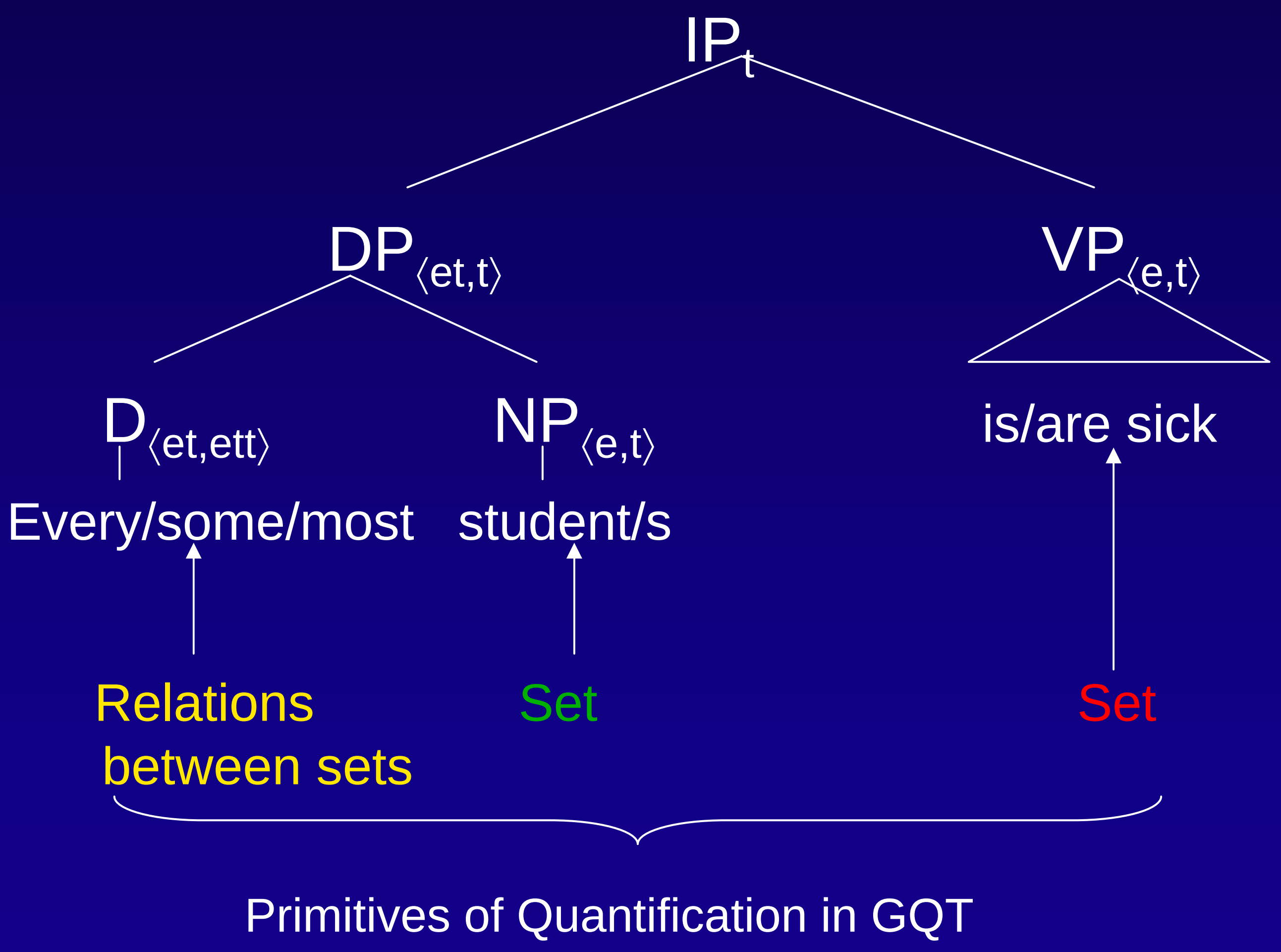
    (1) [[most]] (A)(B) = 1 iff $|A \cap B| > |A - B|$.

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

  - Generalized Quantifier Theory is too coarse to do that.
  - We need a theory that assumes a different set of semantic primitives for quantification in natural language.

Quantification in GQT

(1) Every/some/most student/s is/are sick.



Quantification in GQT (cont.)

- The internal make-up of determiners does not affect the external semantics

(2) [[zero]](A)(B) = 1 iff $A \cap B = \varnothing$

(3) [[fewer than one]](A)(B) = 1 iff $A \cap B = \varnothing$

(4) [[no]](A)(B) = 1 iff $A \cap B = \varnothing$

- Any two equivalent statements of the truth-conditions are equally good

(5) [[no]](A)(B) = 1 iff $A \cap B = \varnothing$, $|A \cap B| < 1$, $|A \cap B| = 0$, …

(6) [[more than 3]] (A)(B) = 1 iff $|A \cap B| > 3$, $|A \cap B| \geq \sqrt{16}$, …

(7) [[most]](A)(B) = 1 iff $|A \cap B| > ½|A|$, $|A \cap B| > |A - B|$, …

Does it matter to other Cognitive Systems?

$|A \cap B| > \frac{1}{2} |A|$

- Determine the total number of elements of A.
- Divide that by 2.
- Compare that to the number of As that are Bs

"More than half of the students are sick."

$|A \cap B| > |A - B|$

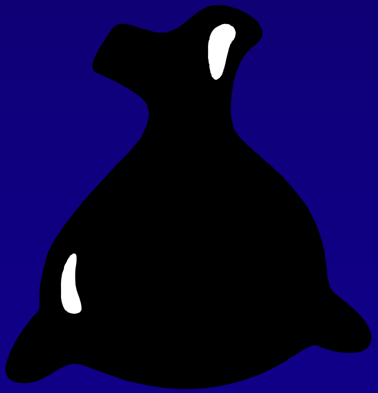
- Compare the number of As that are Bs with the number of As that are not Bs.

"There are more students that are sick than there are students that are not sick."

A new experimental Paradigm: “Self-Paced Counting”

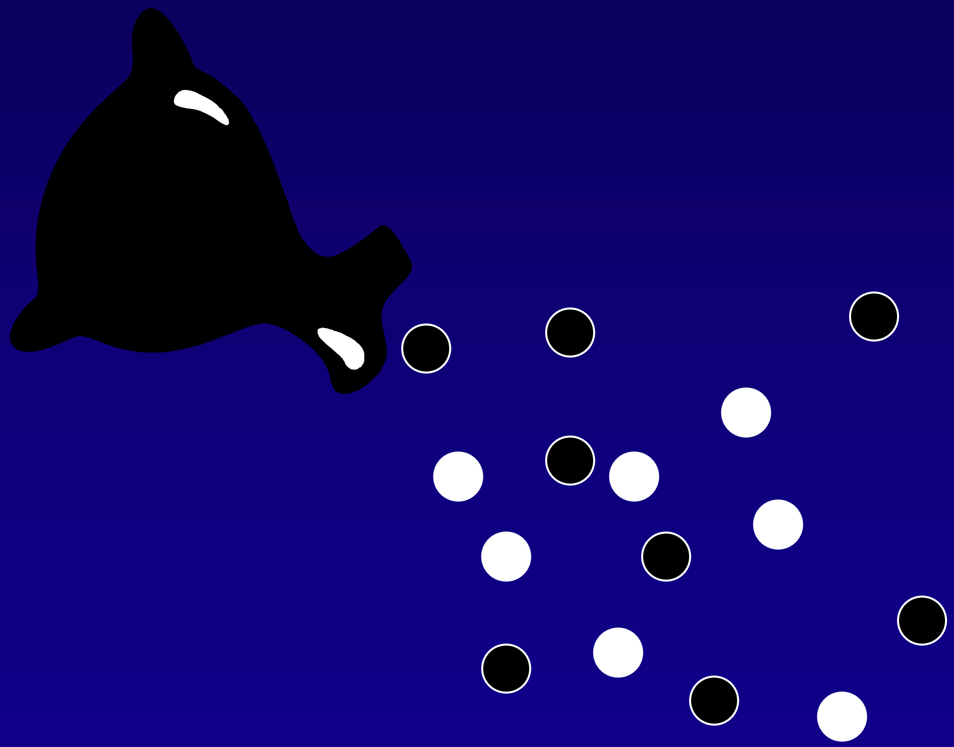
The basic idea behind SPC:

Imagine that you get a bag of marbles and your task is to find out whether most/more than half of the marbles in the bag are black.



“Self-Paced Counting”

Method 1: empty the bag all at once and count the number of black and the number of white marbles.

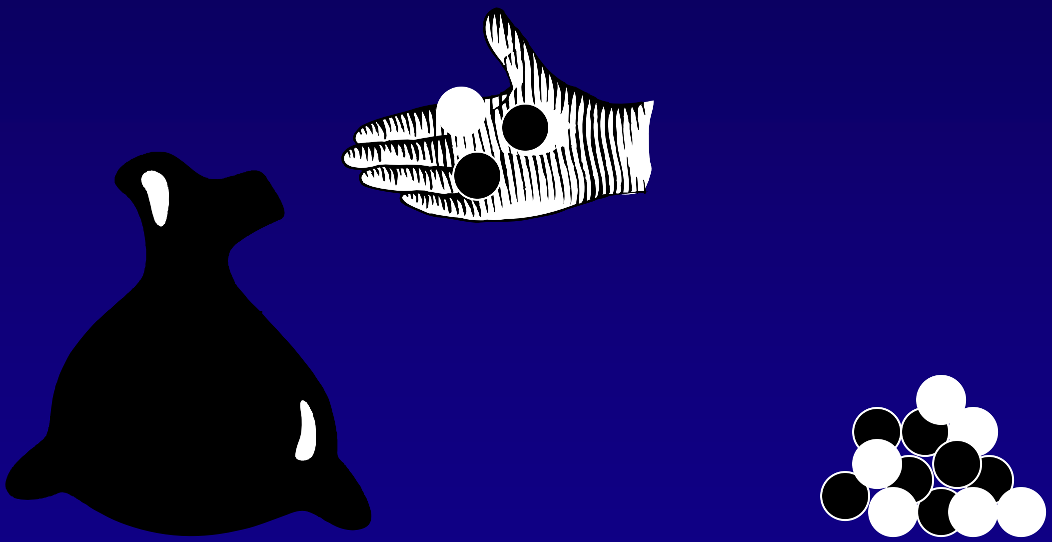


Problem:

There are too many degrees of freedom for solving the task. The geometric arrangement of marbles determines which strategy is being used.

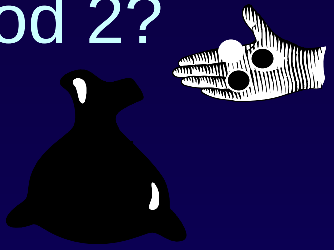
“Self-Paced Counting”

Method 2: Reach in with one hand and grab a handful of marbles to see how many black and white marbles there are. Repeat that as often as necessary.



Intuitively, one is faster using method 2 when the statement employs *most*.

Why is *most* intuitively easier for method 2?

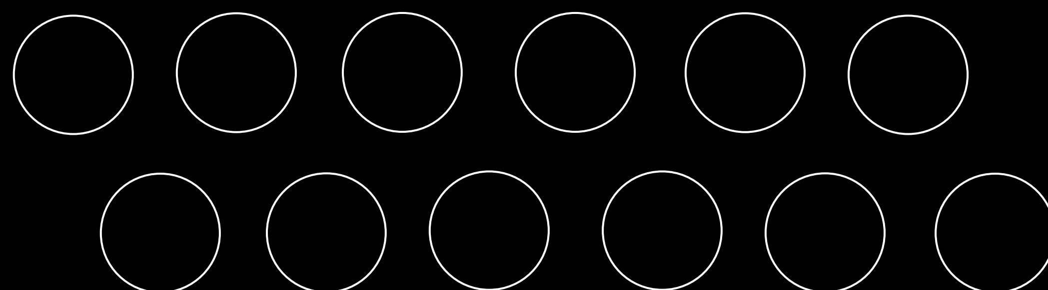


- “Vote Counting” (*Most*):
  - Every time you get a handful of marbles you check whether there are more black than white.
  - Keep track of whether black leads.
- “Counting to a criterion” (*More than half*):
  - Estimate how much half of the marbles is.
  - Check whether the number of black marbles is bigger than that.

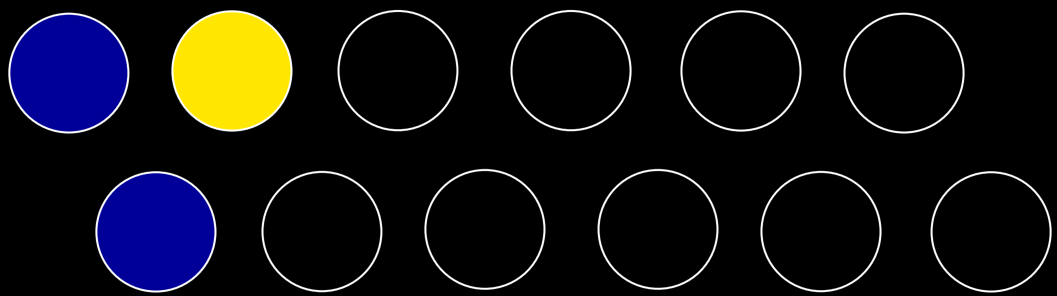
“Self-Paced Counting”

- Self-paced Counting is basically a computerized version of the “bag” modeled after Self-Paced Reading:
  - Subjects hear a sentence whose truth/falsity relative to an array of dots they have to determine as fast and as reliable as possible. “Most of the dots are blue.”
  - Subjects see an array of initially empty dots.
  - The dots are incrementally filled in as subjects press the space bar.
  - Previously seen dots are masked.
  - Subjects can answer as soon as they have enough information.

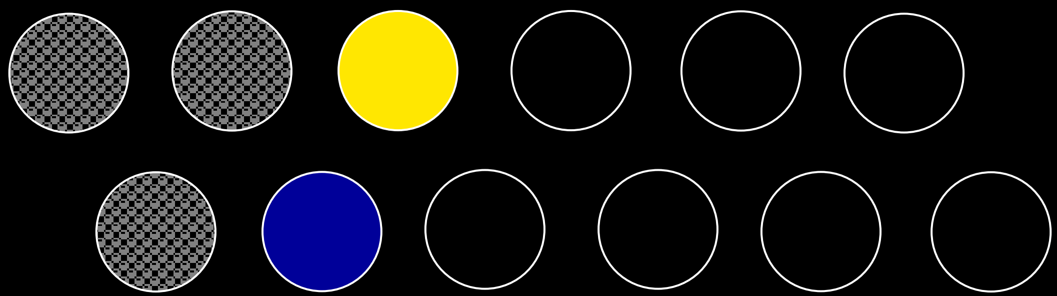
“Most of the dots are blue.”



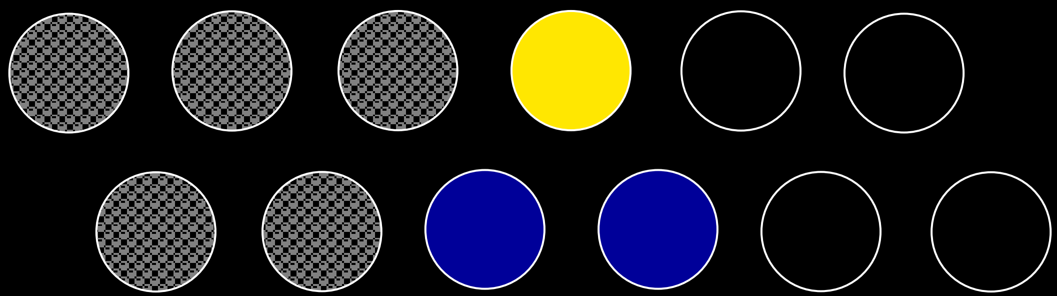
“Most of the dots are blue.”



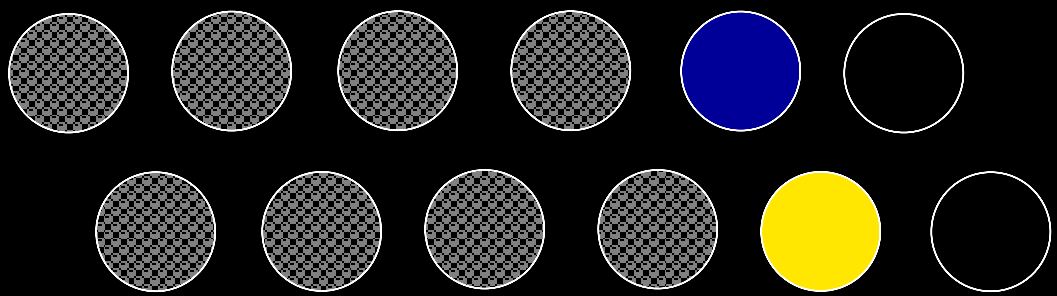
“Most of the dots are blue.”



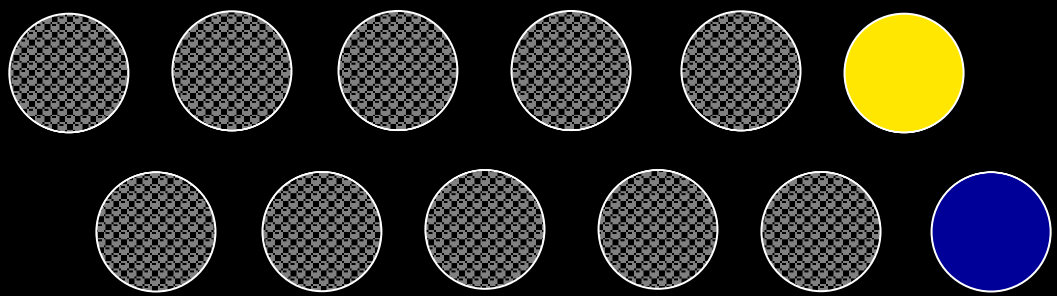
“Most of the dots are blue.”



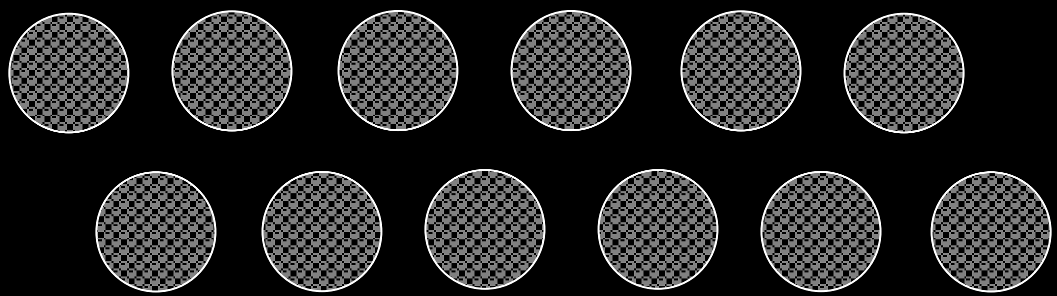
“Most of the dots are blue.”



“Most of the dots are blue.”

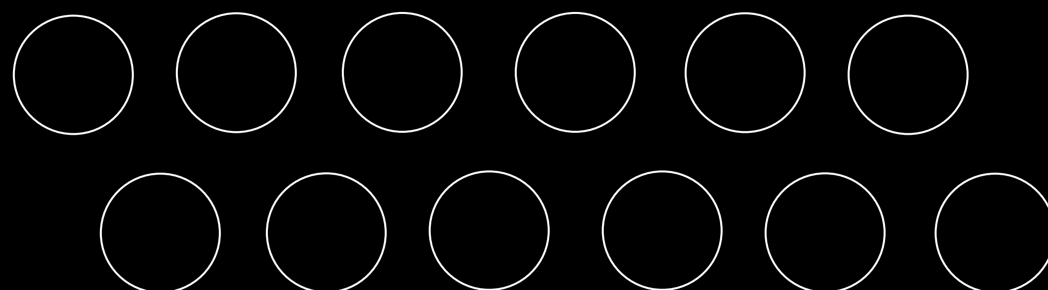


“Most of the dots are blue.”

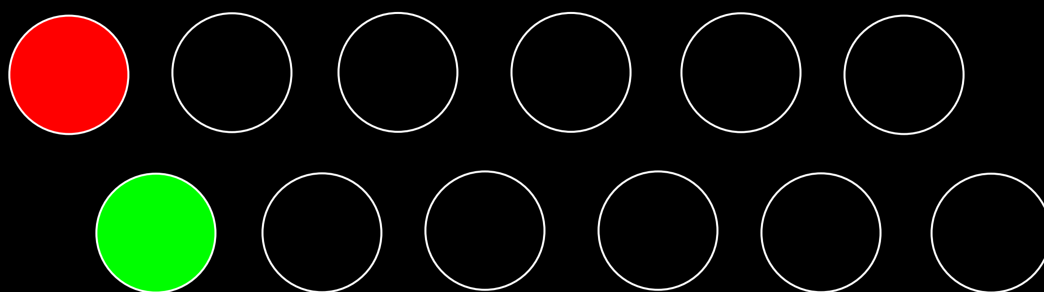


“More than half of the dots are red.”

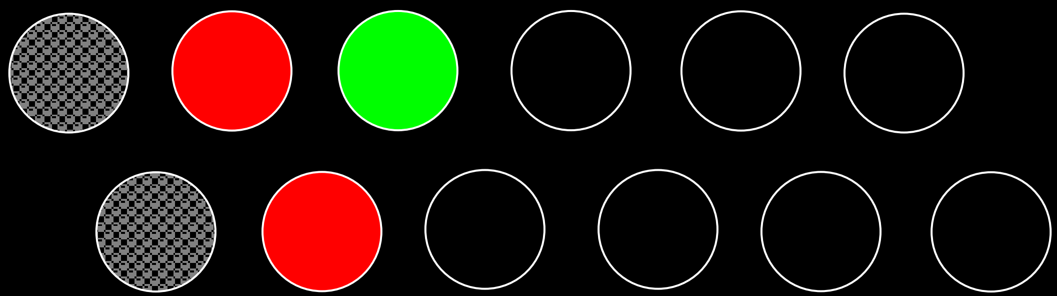
“More than half of the dots are red.”



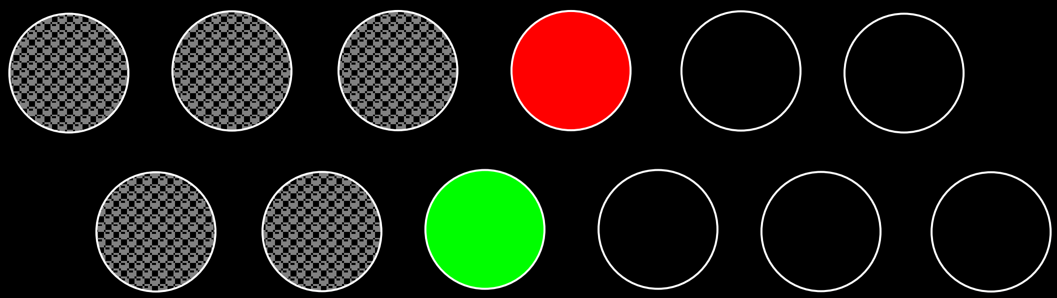
“More than half of the dots are red.”



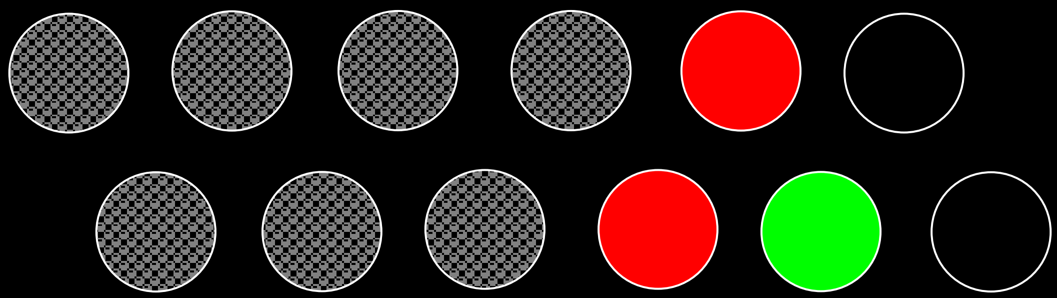
“More than half of the dots are red.”



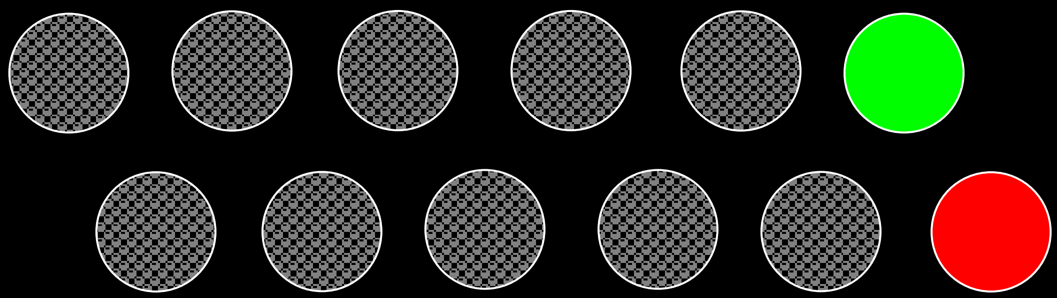
“More than half of the dots are red.”



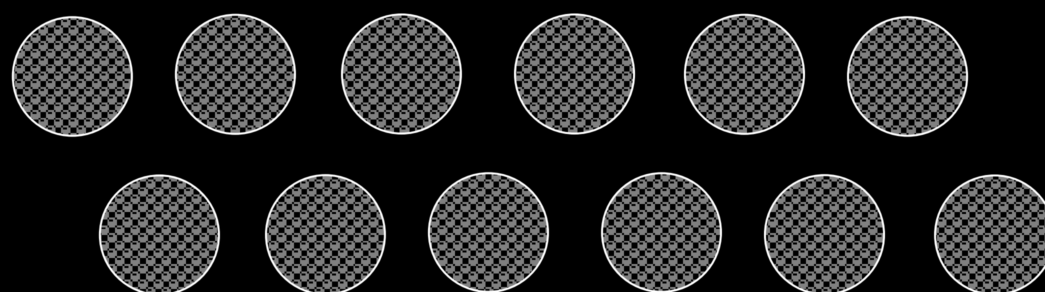
“More than half of the dots are red.”



“More than half of the dots are red.”



“More than half of the dots are red.”



SPC Methods: Most vs. More than half

Target Items :

- 24 target items: 12 *most* and 12 *more than half*
- There are as many true as false target items.
- Target items differed only wrt. what sound was played before (most vs. more than half)
- Dot arrays varied in length between 10 and 12
- Within the first 3 frames one cannot decide whether the sentence is true or false.
- Within the first three frames there are no frames with dots in only one color and there are no frames with only one dot.

Filler Items:

- 36 Filler items: *more than 5, only n, n, many*, *few, some.*
- 18 true, 18 false.
- dot arrays ranged from 7 to 12.

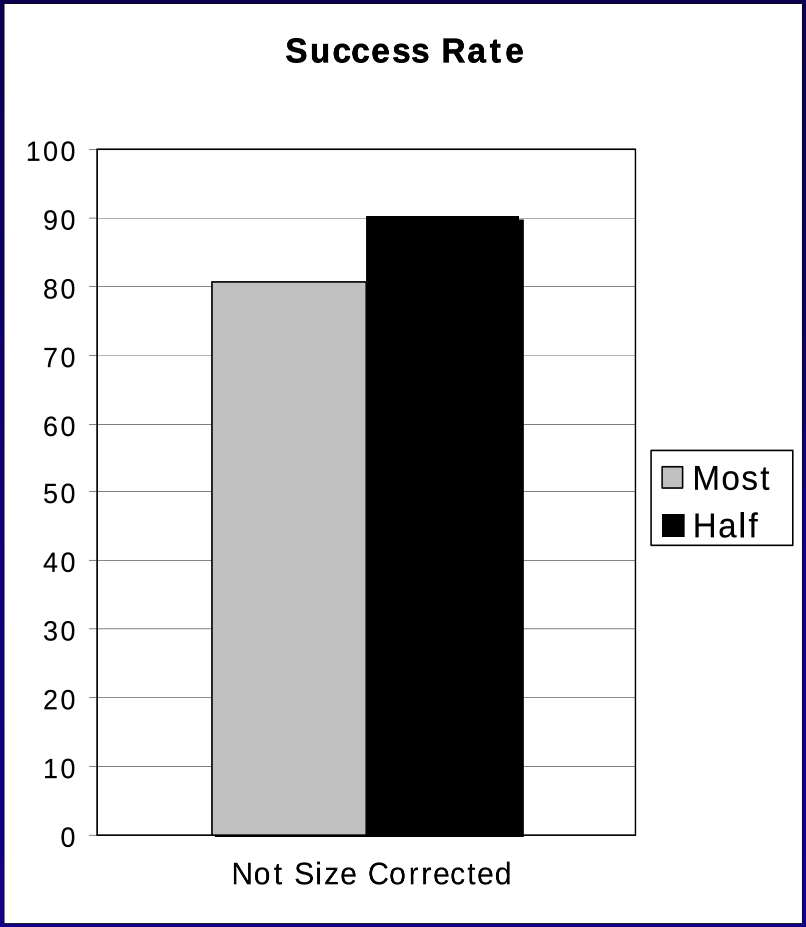
Practice Items: 10

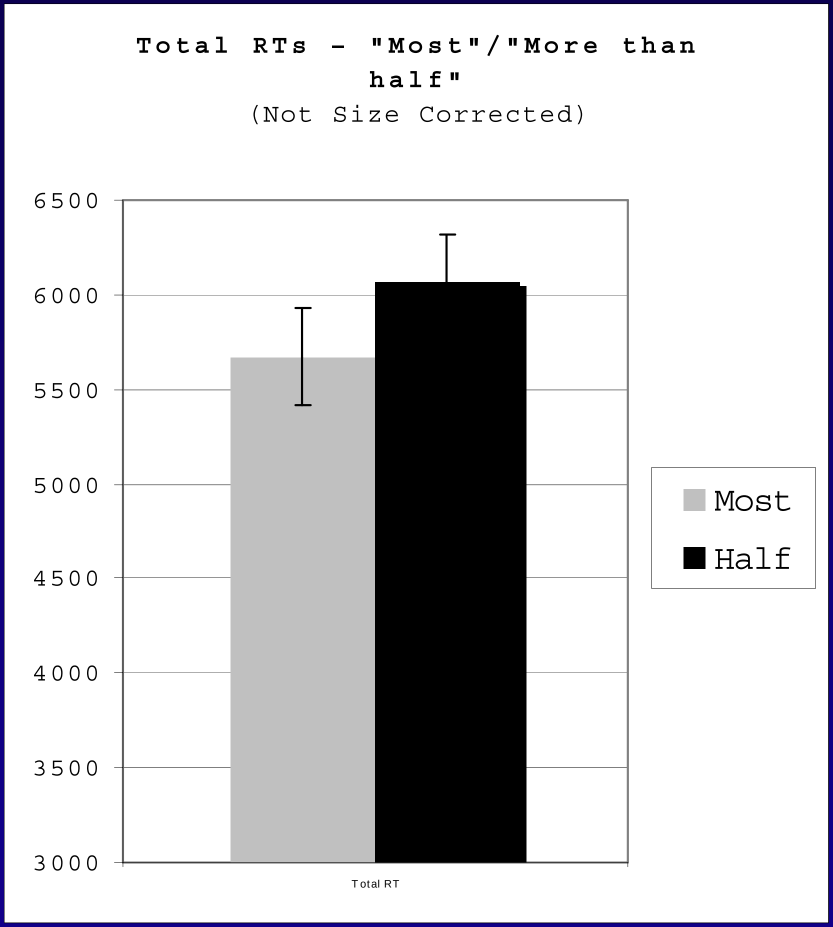
SPC Methodology (cont.)

Results:

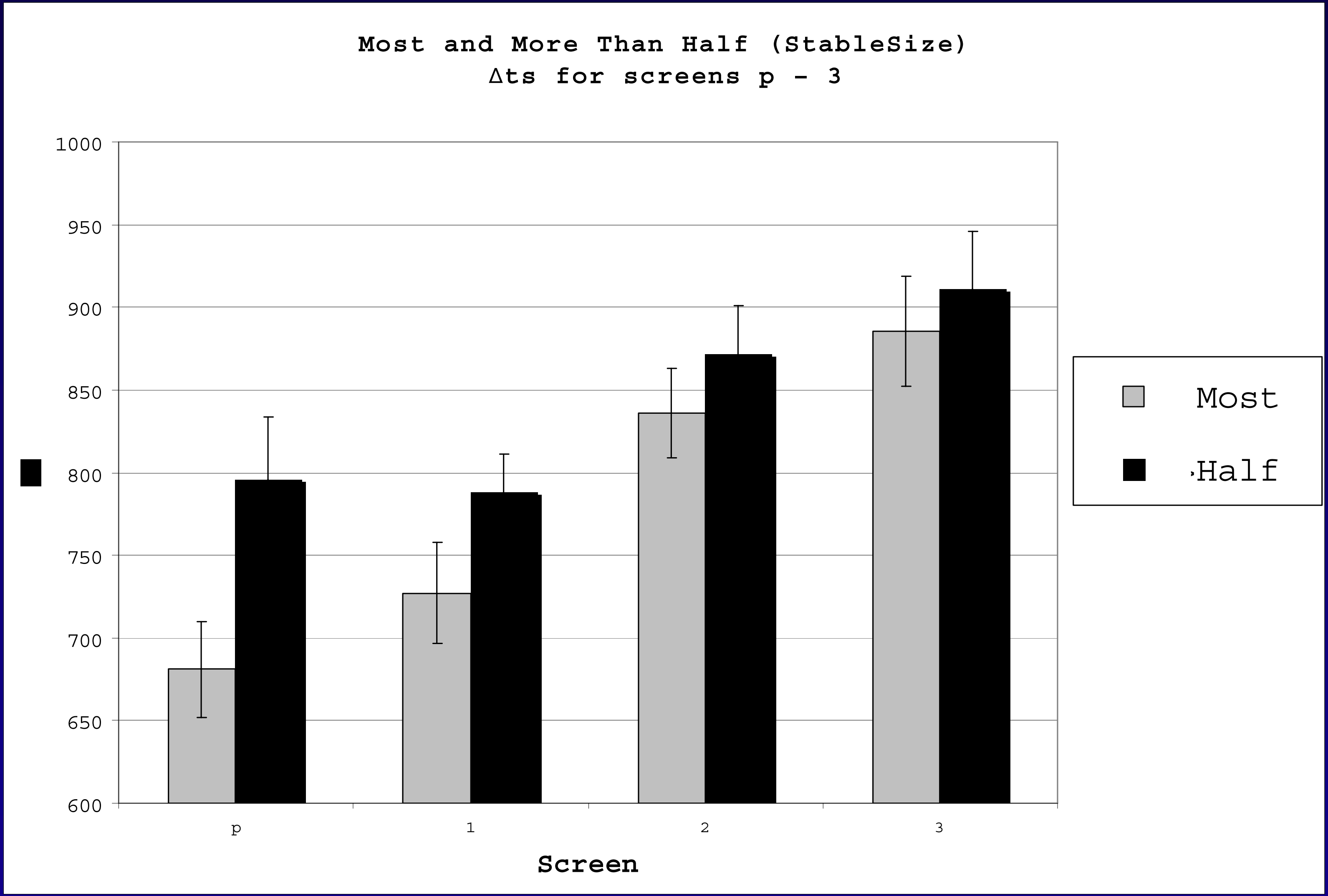
- We analyze only RT from correct answers.
- Subjects were excluded if the percentage if correct answers was below 80%
- We focus on RTs up to frame 3 when it is not yet decidable whether a target sentence is true or false.

Most vs. More than Half (n = 12)





Most vs. More than Half (n=12)



Most vs. More than Half

- Conclusions:
  - *Most* and *more than half* are treated as equivalent determiners
    - Total RTs are not significantly different
    - Percentage of correct responses is not sign. different
  - There is a main effect of Determiner over the first three frames indicating that the dot arrays are easier to process when the prompt is *most.*
  - The linear increase in RT over Frames approaches significance.

Most vs. More than Half

- Question:

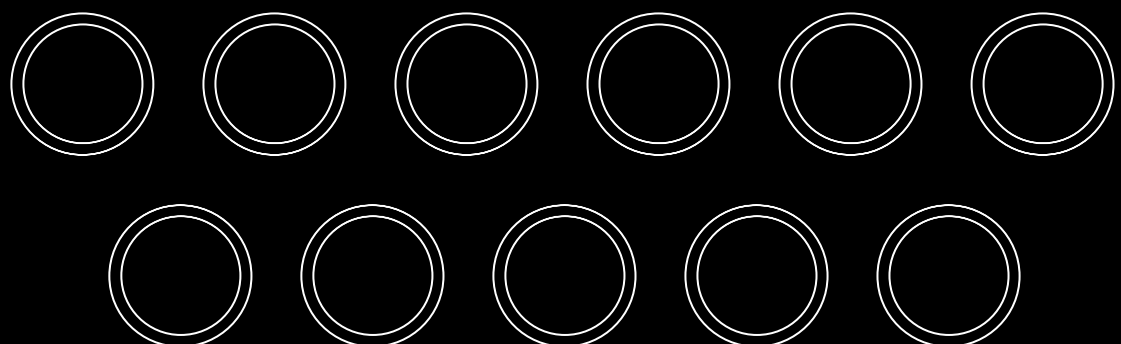
  Is the effect of most being easier in SPC driven by the fact that *most* lends itself more easily to being interpreted as mass quantifier?

  (1) Most dots are blue ≈ There is more blue stuff than there is non-blue stuff.

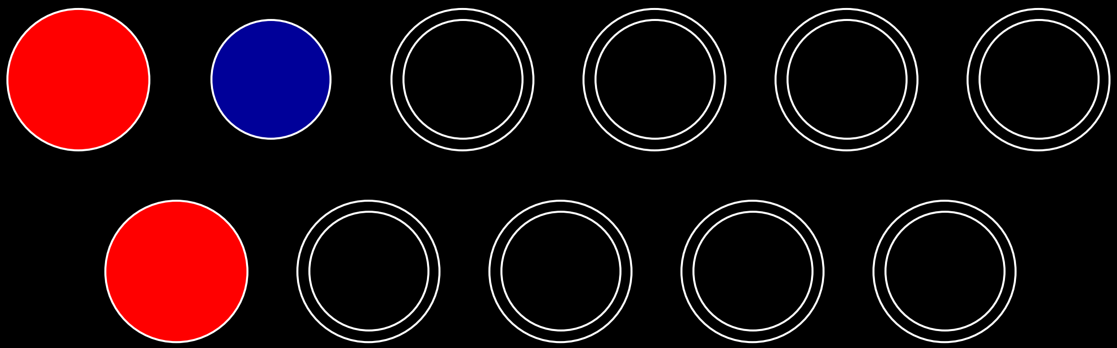
- “Size-Corrected Version of the Experiment”

  Vary the size of the dots pseudo-randomly so that mass is never a reliable predictor for truth/falsity.

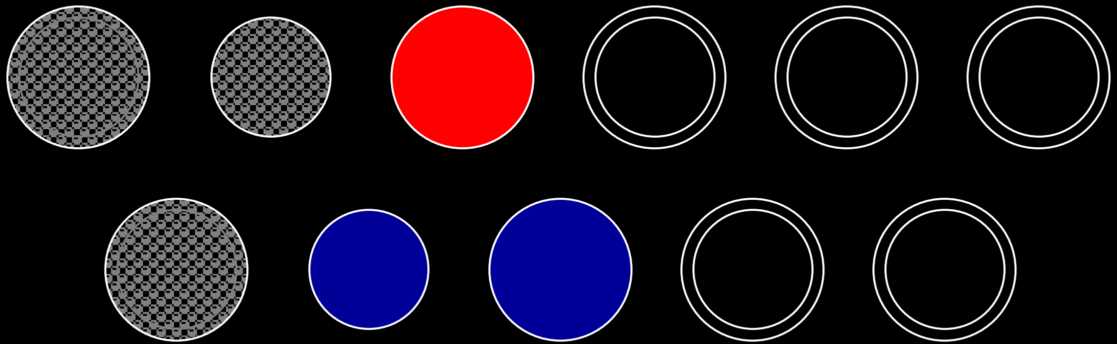
“Most/more than half of the dots are blue.”



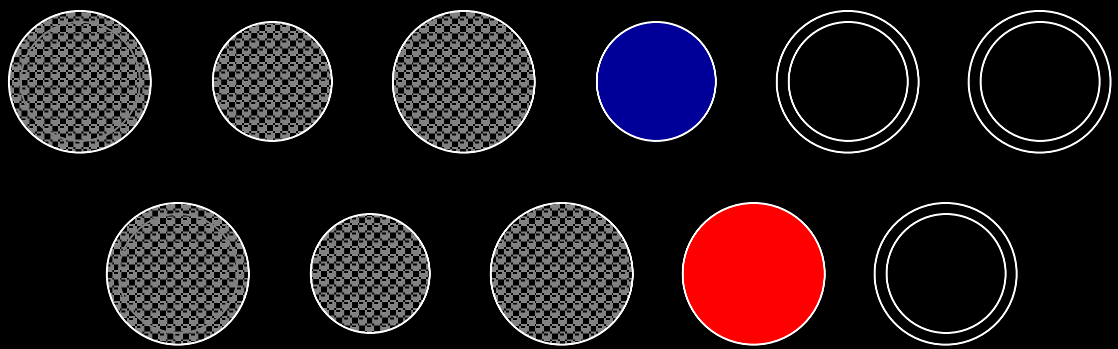
"Most/more than half of the dots are blue."



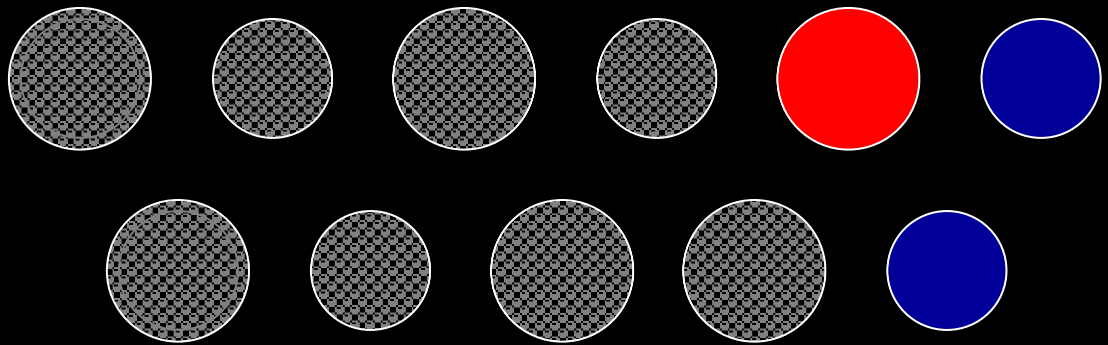
“Most/more than half of the dots are blue.”



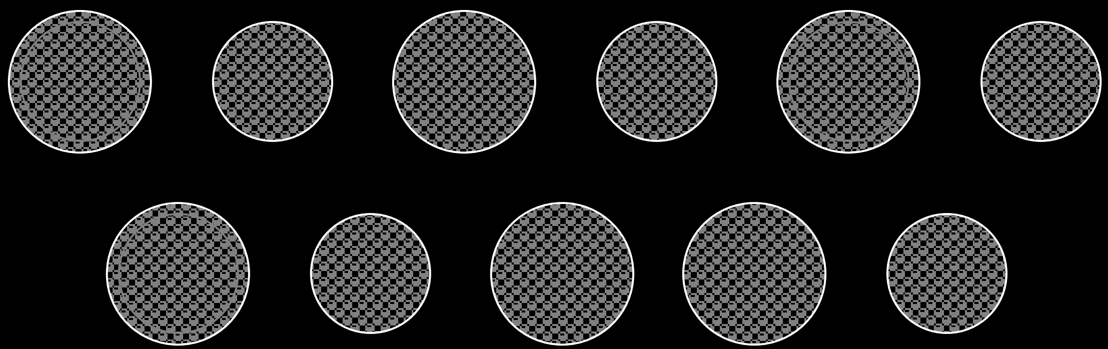
“Most/more than half of the dots are blue.”



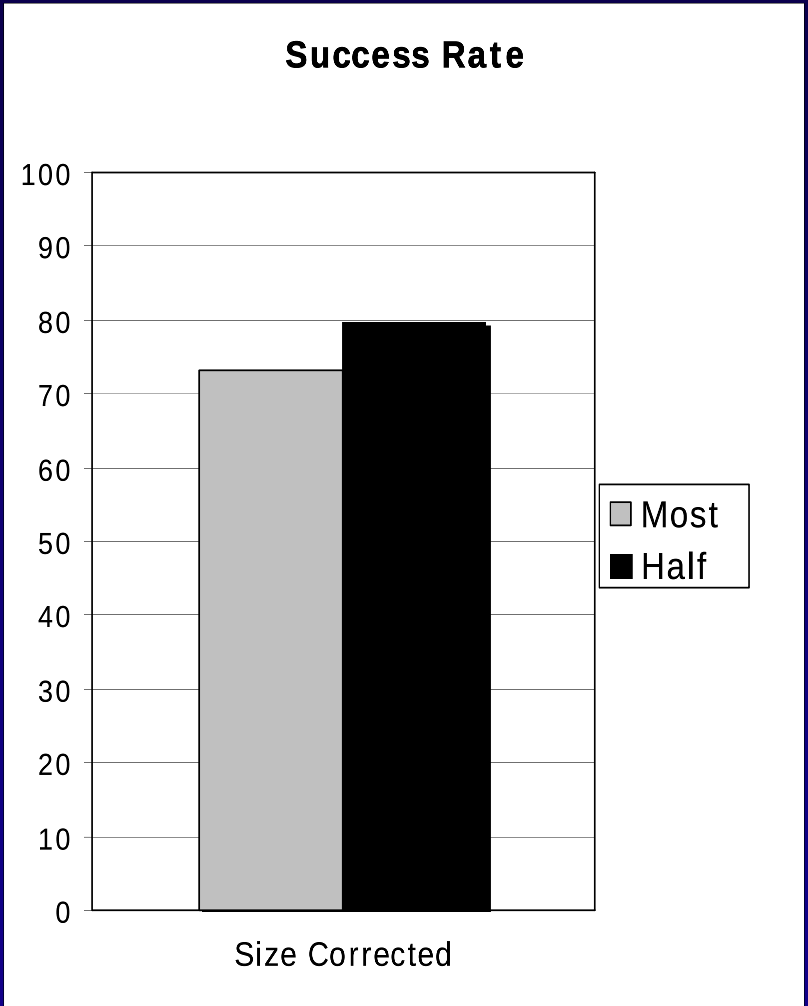
“Most/more than half of the dots are blue.”

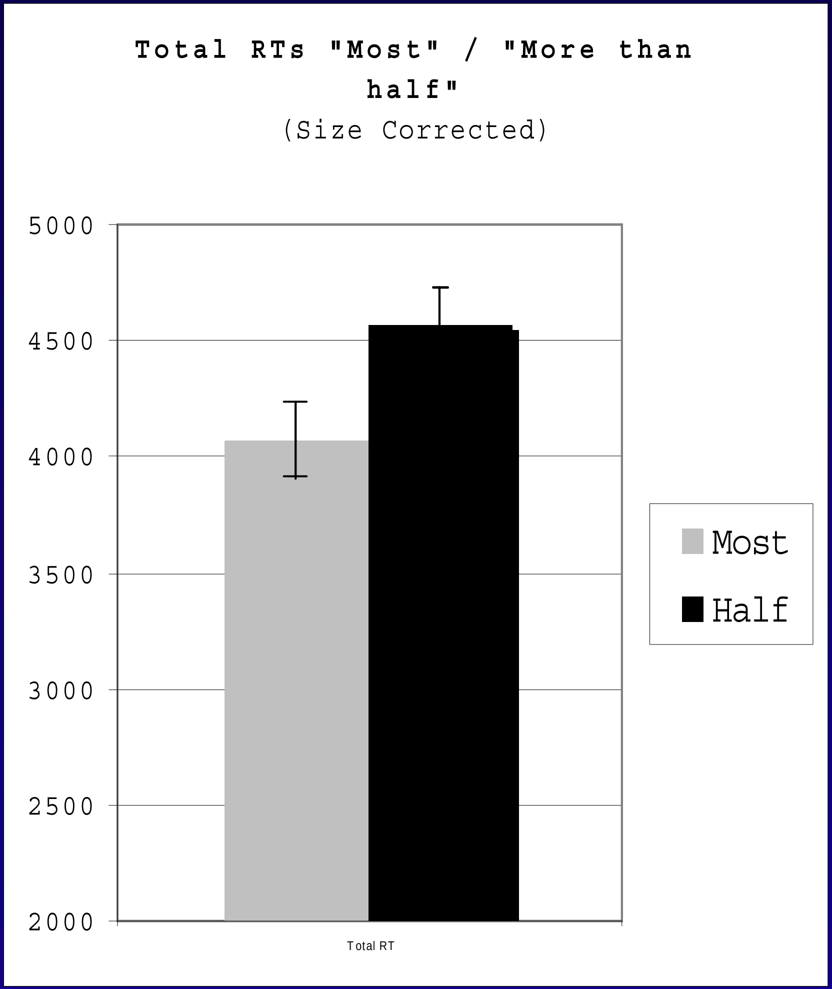


“Most/more than half of the dots are blue.”

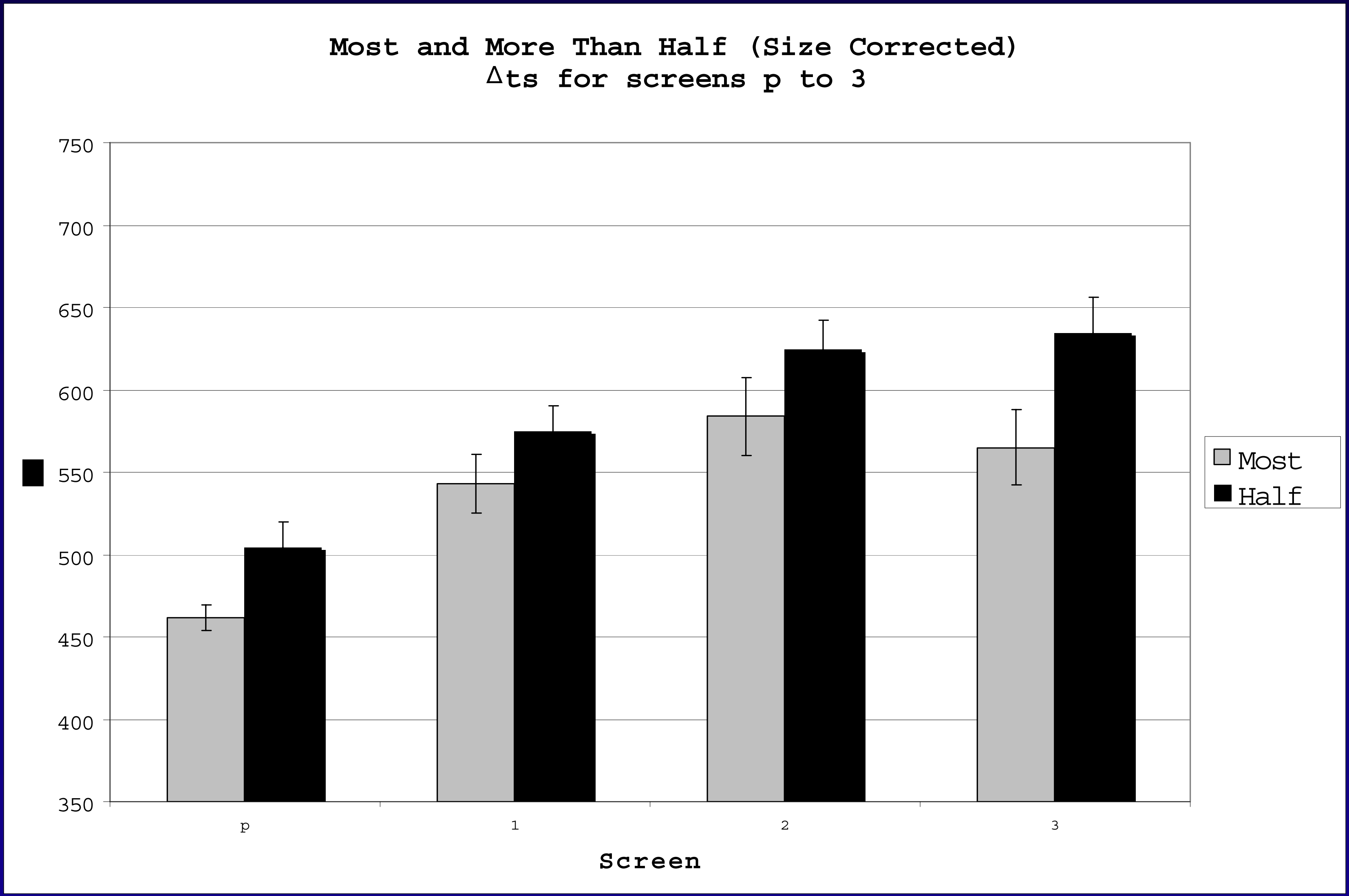


Most vs. More than Half – Size corrected (n=20)





Most vs. More than Half – Size Corrected



Most vs. More than Half – Size Corrected

- Conclusions:
  - *Most* and *more than half* are still treated as equivalent determiners in terms of
    - Total RTs: not sig. different
    - Percentage of correct responses: not sig. different
  - There is a main effect of Determiner over the first three frames indicating that the dot arrays are easier to process when the prompt is *most.*
  - There is a main effect of Frame indicating that RTs increase the farther you go into the array.

Counting to a Criterion in SPC

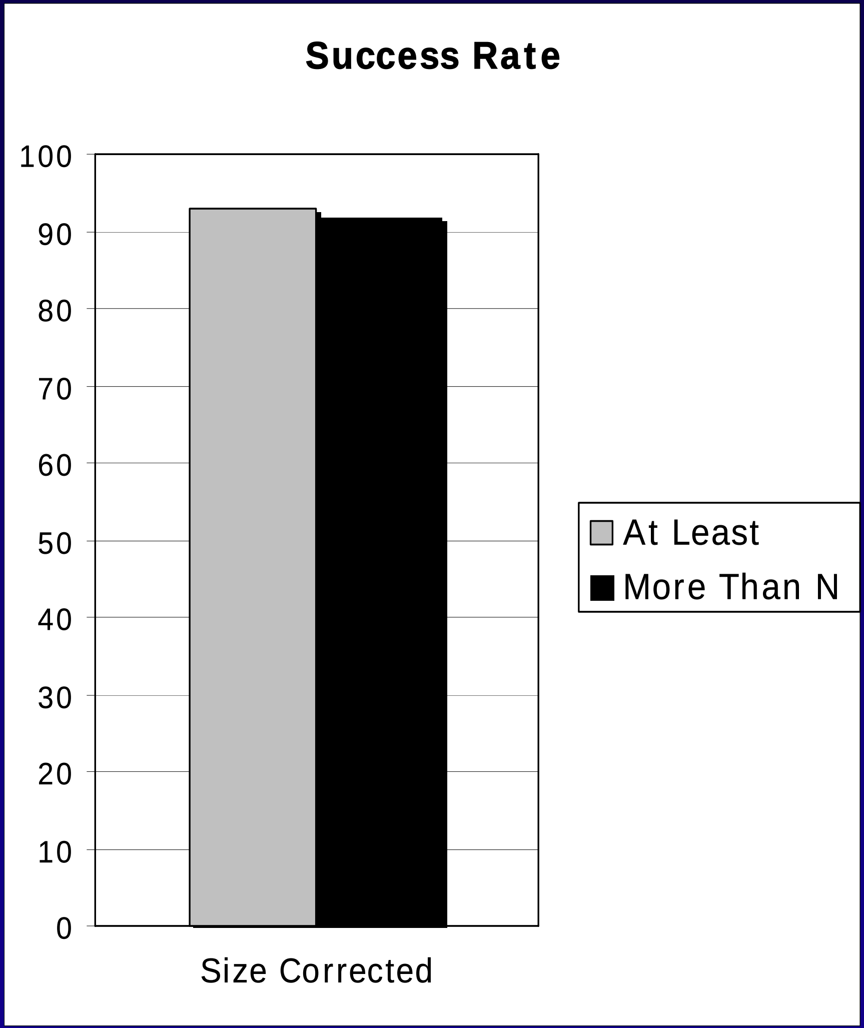
- Question: Is the linear increase by frames due to increased difficulty in counting? Is SPC a reliable methodology to tap into counting processes?

- More than n/At least n+1(Size-Corrected)

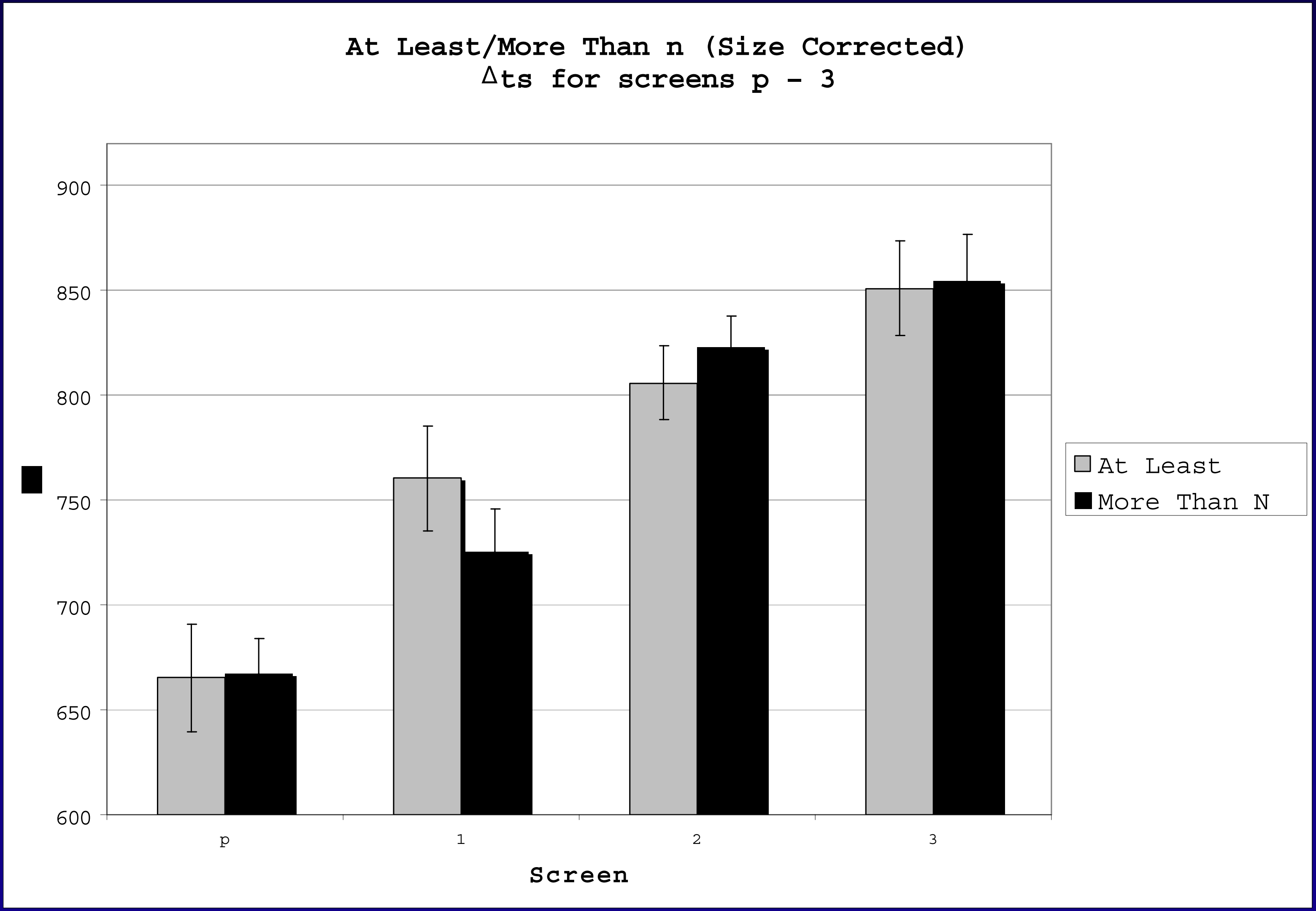
  - Replace the sound files in the *most/more than half* experiment with the corresponding *more than n or at least n+1* sound files.

  - Everything else stays the same.

More than n/At least n +1 (Size Corrected)



More than n/At least n +1 (Size Corrected)



More than n/ At least n+1 Size Corrected

- Conclusions:
  - There is a main effect of Frame indicating that RTs increase the farther you go into the array.
  - There is no effect of determiner by frame 3.

Conclusions

- The particular form in which the truth-conditional import of determiners like *most* and *more than half* is stated is less arbitrary than GQT would have it.
- Language external evidence: Verification procedures triggered by the equivalent determiners (*most* vs. *more than half*) are distinct. They differ in ways that support the following approximations:

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

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

Conclusions (cont.)

- To account for these facts, the difference in form of equivalent determiners needs to be taken into account. I.e. complex determiners like *most* and *more than half* need to be decomposed into smaller building blocks.

- GQT is too coarse because it assumes semantic primitives for natural language quantification – in particular relations between sets – that do not provide the means for systematic decomposition of determiners into smaller building blocks.

Thanks to the following current or former Pomona Students:

David Clausen

Jorie Koster-Moeller

Sandor Prater

Jason Varvoutis

End