Self-Paced Counting
A New Experimental Technique to Study Verification Procedures for Quantified Statements

Ben Acland, Jorie Koster-Moeller, Jason Varvoutis, and Martin Hackl: Pomona College

Summary
This poster presents a novel experimental technique (“Self-paced Counting”) that allows us to gather fine grained timing information about how subjects gather information incrementally in verification tasks that involve counting. We show that this technique can detect different verification profiles for semantically equivalent quantified statements and that evidence of this sort can help distinguish between competing analyses of quantifiers that are said to be indistinguishable in their TC import and their compositional commitments.

Motivation
Formal semantic analyses aim to establish a systematic relation between the truth-conditional (TC) import of an expression and its syntactic/combinatorial properties. How TCs are used by other systems of the mind – for instance in verification tasks – is typically not seen as something that formal semantics needs to account for or that could help distinguish between competing semantic analyses. An area where this lack of interest yields a particularly wide gap that more complete theories eventually will have to bridge is quantification.

Generalized Quantifier Theory
The basic semantic building blocks for quantification in natural language according to Generalized Quantifier Theory are relations between sets of individuals (Barwise&Cooper 1981).

1. D NP is/sick
   VP
   Relations between sets

   Every/some/most student/s
   D                     NP                       is/are sick
   DP                             VP
   (1)                                             IP
   vidividuals (Barwise&Cooper 1981).
   Coarseness of GQT
   The internal make-up of Det. does not affect the DP-external semantics.

   (2) [Every] (A) = 1 if A ⊆ B
   (3) [Some] (A) = 1 if A ∩ B ≠ ∅
   (4) [Most] (A) = 1 if |A ∩ B| > |A - B|
   (5) [More than half] (A) = 1 if |A ∩ B| > ½ |A|

Verification of Quantified Statements
Form:
Most As are Bs
More than half As are Bs.

TCs:
[A ∩ B] > [A - B]          <=>
[A ∩ B] > ½ |A|

Verification:
“Vote Counting”          “Counting to a criterion”
• Compare # of As
  that are Bs to # of As that are not Bs
• Determine # of As.
  Divide by 2.
• Compare to # of A
  that are Bs.

• Can we find evidence that most and more than half trigger different verification procedures as indicated above – even though they are seen as semantically equivalent?

Self-Paced Counting
The basic idea behind SPC

Find out whether most/more than half of the marbles in the bag are black. You reach into the bag repeatedly until you have enough information. At each handful:

Most (Vote Counting):
• Are there more black than white?
• Does black lead overall?

More than half (Criterion Counting):
• Add # black marbles to count of black marbles.
• Is # black marbles > ½ # marbles?

Methods and Materials
Self-paced Counting Methodology:
- Subjects hear a sentence and determine as fast and as reliably as possible its truth/falsity relative to an array of dots.
- The dots in the array are initially empty and ...
- incrementally filled in as subjects press the space bar.
- Previously seen dots are masked.
- Subjects can answer as soon as they have enough information.

Experiment 1: Most / More than half
- 24 target items: 12 most, 12 more than half, 6 true/6 false each
- Dot arrays varied in length between 10 and 12
- Within the first 4 frames it is impossible to determine truth/falsity.
- Variied dot-size so that mass is not reliable predictor within frame 1-4.

Analysis:
- Only RTs from correct answers of subjects with ≥ 80% correct answers.
- Repeated measures ANOVA with factors “Determiner (Det)” and “Screen.”

Experiment 2

Experiment 1 Results (n=20):
• >75% correct: no sig diff
• No sig diff in total RT
• Main effect of Det
(p=0.005)
• Main effect of Screen
(p=0.14)

Experiment 2 Results (n=12):
• >90% correct: no sig diff
• No sig diff in total RT
• Main effect of Screen
(p=0.001)

Conclusions
• Most/More than half are treated as equivalent determiners.
• Verification strategies differ, suggesting (TC approximations):
  - More than half triggers criterion counting ([A ∩ B] > ½ |A|).
• GQT is too coarse to explain that. We need a theory that is based on more fine-grained semantic primitives.
• SPC faithfully tracks increases in the complexity of counting tasks and offers fine-grained timing information about verification procedures involving counting.