Pilot results: extending the priming experiment to more scales (second edition)

Brandon Waldon

August 7, 2018

(A working version of the experiment is available at https://bwaldon.github.io/ad_qp/experiments/6scale/)

As with the first iteration of the multiple scale experiment, there are 6 different scales in this pilot, as detailed below (none of this changed from the first pilot). For every scale, I have developed 3 priming sentences which are changed according to experimental condition to include, a strong alternative (e.g. *all* to *some*), an exhaustive prime (e.g. *some but not all*), or no prime. In each of the priming sentences, participants are asked about the likelihood of an intended exhaustive interpretation (as illustrated below). This is also true for one critical sentence, developed for each scale, which includes the weak scalar item (e.g. *some*).

Scales:

'Some'

- Stronger alternative: 'all'
- Exhaustive prime: 'some but not all'
- Interpretation question: "not every..."
- Example priming sentence: **Sally saw some but not all of her former classmates at her high school reunion**
- Example sentence, no prime condition: **Sally saw her favorite teacher, Mr. Meyer, at her high school reunion.**
- Exhaustivity interpretation question: *Sally didn't see every former classmate of hers at the reunion*

'Or'

- Stronger alternative: 'and'
- Exhaustive prime: 'or but not both'
- Interpretation question: "only one of these things..."
- Example priming sentence: **Peter inherited the painting or the wardrobe from his grandmother, but not both**

- Example sentence, no prime condition: **Peter inherited the painting from his grandmother, whereas his aunt Jill inherited the wardrobe.**
- Exhaustivity interpretation question: *Peter inherited only one of these things from his grandmother*

'Looks like'

- Stronger alternative: 'is'
- Exhaustive prime: 'looks like, but isn't'
- Interpretation question: "something other than. . . "
- Example priming sentence: **It looks like it's raining outside, but it isn't.**
- Example sentence, no prime condition: **It's snowing outside.**
- Exhaustivity interpretation question: *The weather outside is something other than rain*

Cardinal numbers (i.e. 34)

- Stronger alternative: '35'
- Exhaustive prime: 'exactly 34'
- Interpretation question: "no more than 34. . . "
- Example priming sentence: **Exactly 34 books were stolen from the local library yesterday.**
- Example sentence, no prime condition: **There was a break-in at the local library yesterday.**
- Exhaustivity interpretation question: *No more than 34 books were stolen from the local library yesterday*

'Palatable'

- Stronger alternative: 'delicious'
- Exhaustive prime: 'palatable, but not delicious'
- Interpretation question: "not exceptionally tasty. . . "
- Example priming sentence: **Guillermo's brewery has just released a new beer that is palatable, but not delicious.**
- Example sentence, no prime condition: **Guillermo's brewery has just released a new beer.**
- Exhaustivity interpretation question: *Guillermo's brewery has just released a new beer that is not exceptionally tasty*

'Hard'

- Stronger alternative: 'unsolvable'

- Exhaustive prime: 'hard, but not unsolvable'
- Interpretation question: "not impossible to. . . "
- Example priming sentence: **The homework that Professor Bridges assigns is hard, but not unsolvable.**
- Example sentence, no prime condition: **The homework that Professor Bridges assigns is always due the Monday after it is assigned.**
- Exhaustivity interpretation question: *The homework that Professor Bridges assigns is not impossible to finish*

Changes from the first multi-scale pilot

More fillers

There's now a single filler trial in between each scale block, such that can crete clear exclusion criteria. For half of these fillers, we should reasonably expect that participants answer on the upper part of the scale; for the other half, they should answer on the lower end. The order in which these fillers is presented is randomized for every participant (but a single filler is always presented between scale blocks).

Six possible exclusion criteria

This is obviously not an exhaustive list, but for the purposes of analyzing these data, I explored six possible ways to exclude participants:

- The **strict** exclusion criterion: participants are excluded for making one or more mistakes on fillers in the experiment.
- The **lax** exclusion criterion: participants are excluded for making N or more mistakes, for $N > 1$.
- The **absolute-mean** exclusion criterion: if the mean of the good fillers is below 50, OR the mean of the bad fillers is above 50, you're out.
- The **relative-mean** exclusion criterion: one of the criteria you suggested - "if the mean of the good fillers is lower than the mean of the bad fillers, you're out".
- The **bad-better-than-good** exclusion criterion: another suggestion of yours - "if any of the good fillers receive lower ratings than any of the bad fillers, you're out".
- The **difference-in-means** exclusion criterion: your last suggestion - "if the difference in means is lower than X, you're out".

Below, I report the number of participants excluded given each of these criteria.

Coding changes

- I fixed the hard/palatable coding error.

Presentational changes

- Every scale block is associated with a different, named speaker (in this pilot, the speaker names were fixed for a given scale block; e.g. the speaker of the *looks like* block was always 'Christina').
- For every scale block, the (entire) critical trial is presented in red. Going back, I realize this isn't exactly what you wanted; in the next version of the experiment, the form will be colored uniquely from the rest of the sentence on critical trials.

Procedure

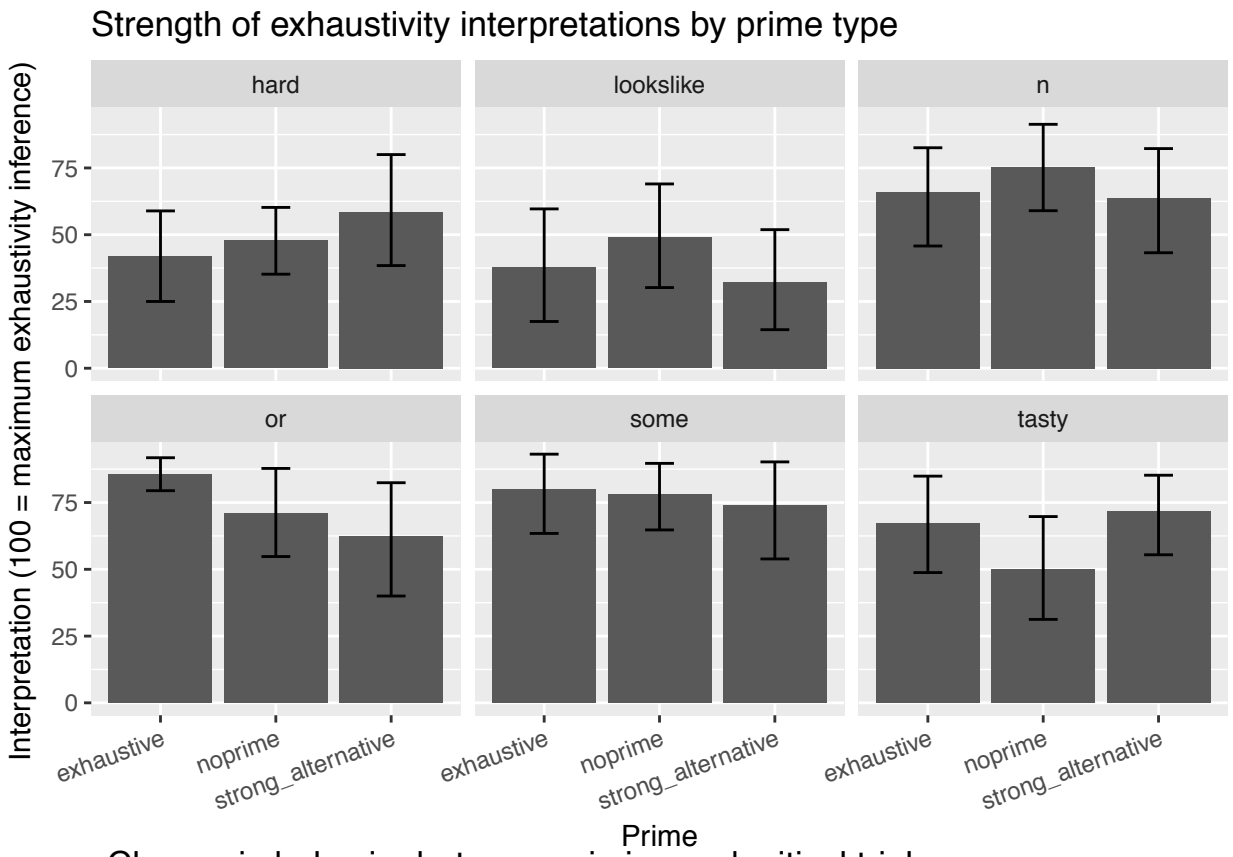
As with the last pilot, I created three experiment lists such that every list elicited interpretation of the weak scalar item from each scale above. That is, every participant interpreted *some*, *or*, *looks like*, etc. exactly once. Moreover, the lists were balanced such that data could be collected between-subjects for the effect of each type of prime on each scalar item. For example, in one list, participants were asked about their interpretation of *or* after seeing three exhaustive primes, *some* after seeing three strong primes, *looks like* after seeing three 'no prime' sentences, etc. For a given scale, the critical trial was shown immedately after the three corresponding prime trials, but the order of these groups of priming/critical trials was randomized within-subject. I recruited 30 participants to take this pilot.

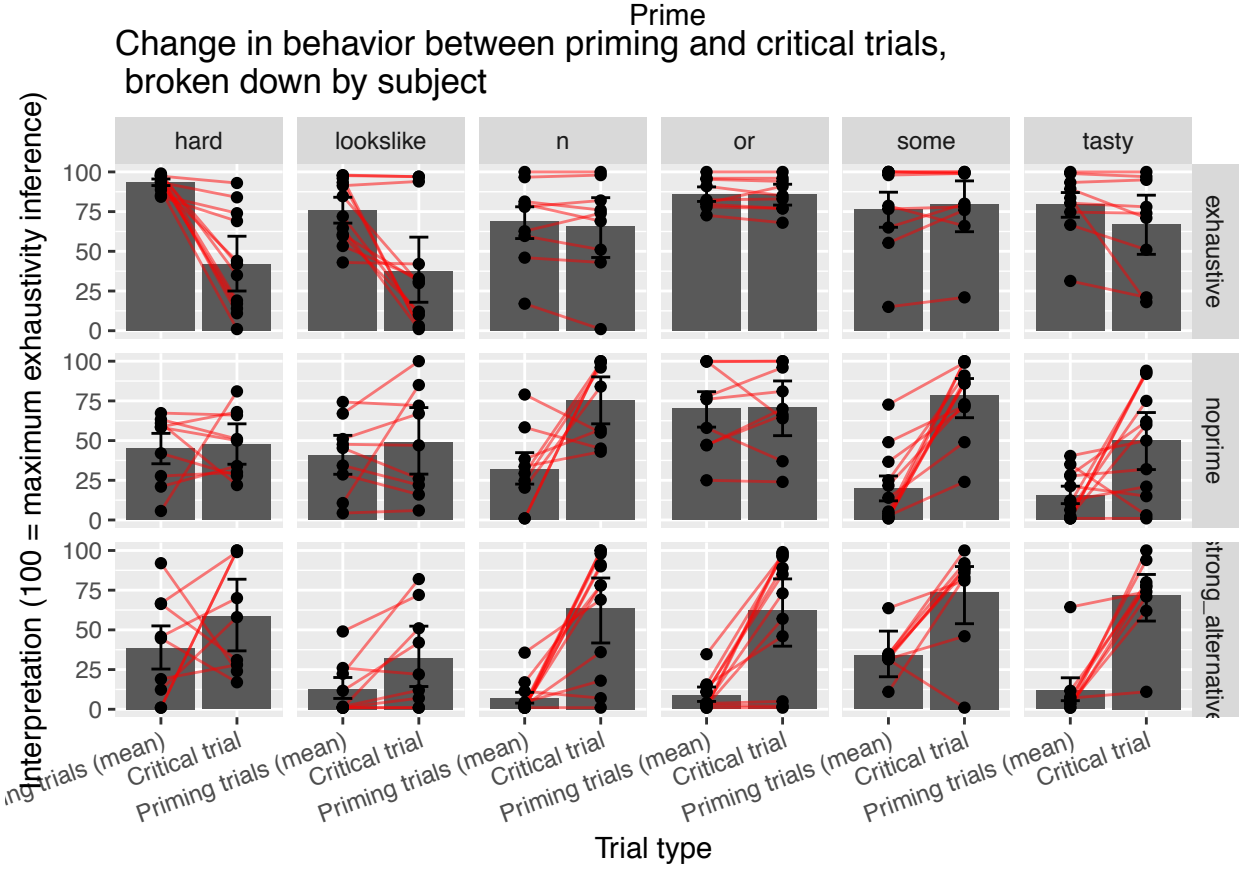
Results

Number of exclusions for every exclusion criterion (out of 30 participants total)

- The **strict** exclusion criterion: 10 exclusions
- The **lax** exclusion criterion: 3 exclusions (if we exclude participants who make more than 1 error).
- The **absolute-mean** exclusion criterion: 3 exclusions
- The **relative-mean** exclusion criterion: 1 exclusion
- The **bad-better-than-good** exclusion criterion: 4 exclusions
- The **difference-in-means** exclusion criterion: 3 exclusions (if we exclude participants for whom the difference in means is lower than 10, given a 0-100 scale).

Visualizations (without exclusions)





Visualizations (excluding participants using the strict criterion)