Question availability as a predictor of scalar diversity

Eszter Ronai (The University of Chicago) & Ming Xiang (The University of Chicago)
ronai@uchicago.edu

Background. Recent work has found considerable variation across different scales in the rates of scalar inference (SI) calculation; for instance, the SI in (1) arises much more robustly than the one in (2) (i.a. Van Tiel et al. 2014). This variation has been termed scalar diversity. The question arises, then, how to capture this observed variation: can we identify some properties of different scales that influence how robustly they lead to SI calculation? Van Tiel et al. (2014) found the distinctness of the stronger scalar term, while Sun et al. (2018) found local enrichability to be such a property. However, there is still a lot of variance unaccounted for in the empirical results.

Existing work on <*some, all*> SIs has shown that Question Under Discussion (QUD) has an effect on SI calculation rates (i.a. Degen & Tanenhaus 2014; Zondervan et al. 2008). But in previous work on scalar diversity, stimulus sentences were presented in the absence of any context —raising the possibility that there is variation across scalar terms in what QUD they most naturally bring to mind.

Hypothesis. Our hypothesis is that scalar diversity, in the absence of an explicit QUD, arises (in part) due to the differential availability of a polar question containing the stronger scalar term. That is, the more likely a question such as *Is the student brilliant?* is, the higher the rate of SI calculation from the corresponding statement *She is intelligent*. This predicts a correlation between the availability of the polar questions that contain the stronger scalar term and the relevant SI rates.

Exp. 1 (participant N=37) was a replication of van Tiel et al.: an inference task investigating SI rates from 43 different scales. Participants saw sentences such as "Mary: *The student is intelligent.*" and were asked the question "Would you conclude from this that, according to Mary, the student is not brilliant?". They responded by clicking "Yes" (= SI calculation) or "No" (= no SI calculation).

Exp. 2 (participant N=40) added a two-condition Question manipulation to the task of Exp. 1. Mary's statement was preceded either by a question containing the stronger scalar, or one containing the weaker scalar: "Sue: *Is the student brilliant/intelligent?*"; "Mary: *She is intelligent*".

Exp. 3 (participant N=35) was a forced choice task: participants had to choose which of the two polar questions (containing the stronger vs. the weaker scalar term) they would be more likely to ask. For example, given the instructions "Compare the following two questions about a student. Which one are you more likely to ask?", participants had to choose between *Is the student brilliant?* and *Is the student intelligent?*. Under our hypothesis, the results from Exp. 3 (henceforth Question Choice) should predict scalar diversity, i.e. the variation in SI calculation rates from Exp. 1 —the more preferred the stronger question is in Exp. 3, the higher the SI rate should be for that scale in Exp. 1.

Results and discussion. Exp. 1 replicated scalar diversity. Exp. 2 found a significant effect of the Question manipulation: across the board, more SIs were derived when the preceding question contained the stronger scalar term than when it contained the weaker one ($p < 0.001$; Fig. 2). Question Choice from Exp. 3 was not found to be an overall predictor of SI rates. The picture changes, however, once we take into account "boundedness". Van Tiel et al. define a scale as bounded if the stronger scalar denotes an endpoint: <*some, all*> is bounded, but <*intelligent, brilliant*> is not. They found that bounded scales produced significantly higher SI rates than unbounded ones—which our data replicated ($p < 0.001$). Crucially, we also found a significant interaction of Question Choice with Boundedness ($p < 0.05$). For unbounded scales, Question Choice (Exp. 3) showed a strong trend ($p < 0.08$) in predicting likelihood of SI calculation (Exp. 1). As Fig. 1 shows, for unbounded scales, the more likely participants were to choose the strong question, the higher the rate of calculating the relevant SI. For bounded scales, however, there was no effect of Question Choice ($p = 0.14$).

In unbounded scales, both scalar terms denote intervals whose values vary according to context. This leaves room for contextual support: the more available a QUD based on the stronger scalar, the more likely hearers are to reason about it as the stronger alternative, deriving the SI. In bounded scales, the stronger scalar instead denotes a fixed point, and is thus more salient as a stronger alternative to the vague, weaker scalar. This results in high SI rates, and the QUD makes no difference.

Conclusion. We investigated the role of QUD in explaining scalar diversity. **Explicit questions** (based on the stronger vs. weaker scalar) were shown to **influence SI calculation rates** for a large number of different scales. We also found that **the likelihood of a question based on the stronger scalar is a factor** contributing to scalar diversity, but only **for unbounded scales**.

(1) Mary ate some of the cookies. → SI: Mary ate some but not all of the cookies.

(2) The student is intelligent. → SI: The student is intelligent but not brilliant.

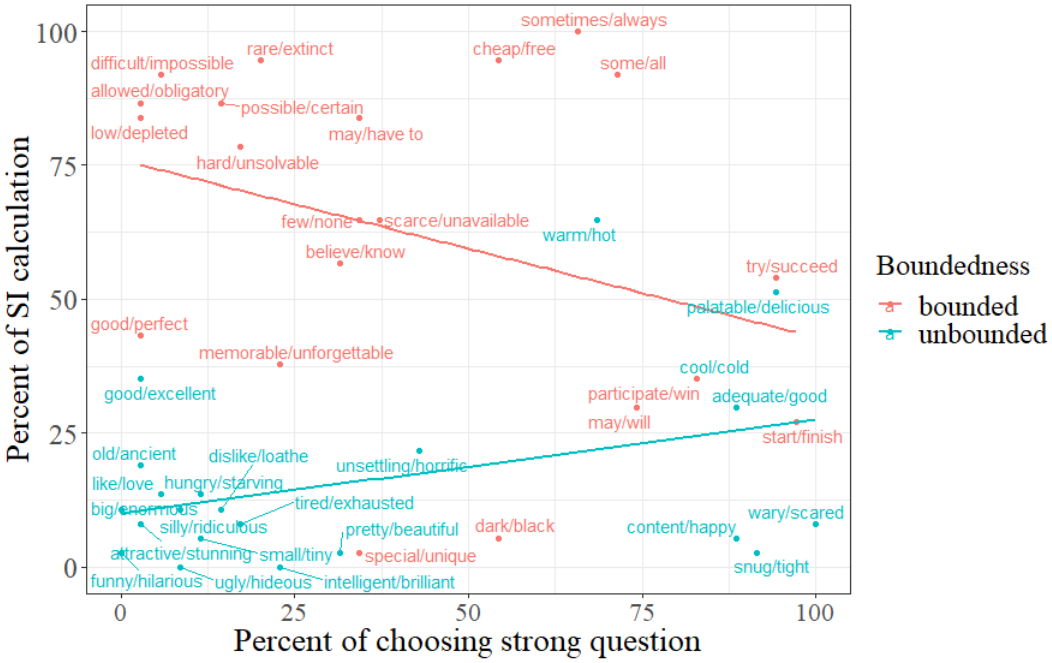

Figure 1: Experiment 1 (y axis) and Experiment 3 (x axis) results

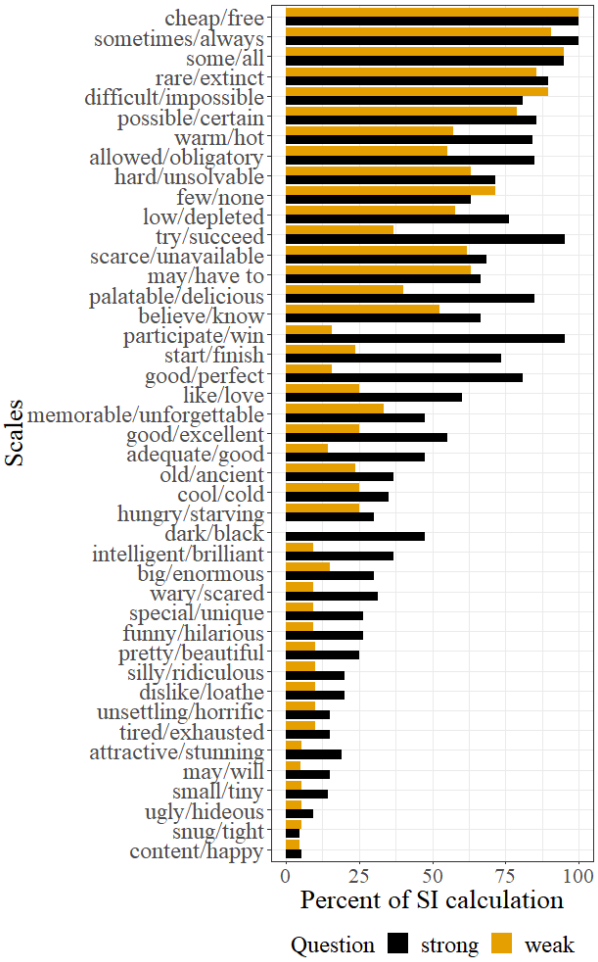

Figure 2: Experiment 2 results

References:

Degen, J., & Tanenhaus, M. K. (2014). Processing scalar implicature: A constraint-based approach. Cognitive Science.

Sun, C., Tian, Y., & Breheny, R. (2018). A link between local enrichment and scalar diversity. Frontiers in Psychology.

Van Tiel, B., Miltenburg, E. V., Zevakhina, N., & Geurts, B. (2014). Scalar diversity. Journal of Semantics.

Zondervan, A., Meroni, L., & Gualmini, A. (2008). Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. Proceedings of SALT 18.