

The Role of Task Framing and Context Source in Scalar Implicature Detection

Michael Vrazitulis

University of Potsdam

vrazitulis@uni-potsdam.de

Scalar implicatures (SIs) are commonly studied using sentence verification tasks. In these tasks, participants judge underinformative sentences such as *Some trees are plants*. A standard linking assumption holds that rejecting such sentences as false reflects a pragmatic interpretation (but see [7]). Accepting them as true reflects a literal interpretation [1, 4].

Literal vs. pragmatic response rates for the English scalar *some* vary widely across previous studies (e.g., from 38% literal responses [9] up to 77% literal responses [6]). One underexplored source of variation is task framing—whether judgments are about truth or felicity (as suggested by [7]). Another is the source of context provided, i.e., either world knowledge (WK) or accompanying pictures (sentence–picture verification, [2]). We tested the impact of these two factors on SI response behavior.

Across three web-based experiments (N = 252, 252, and 576), we used a 2×2 between-subjects design. We manipulated framing (truth vs. felicity) and context source (WK vs. picture). Participants judged underinformative targets and control items, responding with either *true/false* or *good/bad* judgments. Figure 1 illustrates an example item. There were 24 items overall.

Experiments 1 and 2 included 24 trials per participant, evenly distributed across three conditions: target, control true, and control false. Experiment 3 also included 24 trials per participant, but featured three additional control types using the stronger scalemate *all*. Further, Experiment 2 differed from the other two in response format: In Experiments 1 and 3, participants chose between explicit polarity pairs (e.g., *Is this a good or a bad statement?*). In Experiment 2, they answered yes/no questions instead (e.g., *Is this a good statement?*).

All three experiments showed the same basic result pattern. Participants gave more pragmatic responses in the felicity condition than in the truth condition (95% credible interval: Exp. 1 [+6.5%, +30.3%], Exp. 2 [+0.6%, +15.5%], Exp. 3 [+25.1%, +47.9%]). We also observed a consistent main effect of context source: Participants responded more pragmatically in the picture-based condition than in the WK-based one (95% credible interval: Exp. 1 [+4.6%, +29.3%], Exp. 2 [−0.9%, +13.5%], Exp. 3 [+18.1%, +44.5%]). Literal response rates in the WK condition were especially high, sometimes even reaching over 80%. Figure 2 displays the average literal response rates across conditions and experiments. These results suggest that SI processing is indeed modulated by task framing and context source.

Taken together, our findings indicate that task design can systematically influence how underinformative sentences are interpreted. Future work should identify which design factors most reliably affect response rates. Crosslinguistic differences may also play a role: For example, all else being equal, English *some* (e.g., [6, 7]) seems to elicit more literal responses than French *certain*s (e.g., [1, 3, 8]) or Dutch *sommige* (e.g., [4, 5]). A clearer understanding of how contextual features shape response rates will help clarify which aspects of SI processing are stable and which depend on particularities of the experimental setup.

Experiment scripts for PCIBex [10], the collected response data, and R code for statistical analysis can be accessed here (anonymized link):

https://osf.io/86r4u/files/osfstorage?view_only=61773f9ca6e84e5db09e317d57617947

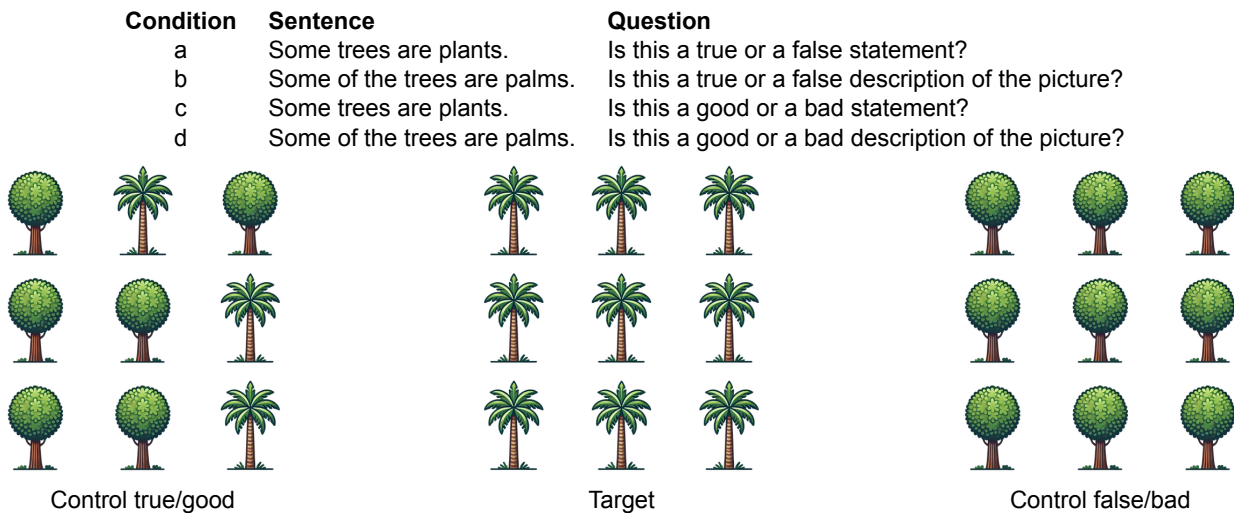


Figure 1: Above: Example item shown in each of the four conditions in the 2×2 design. Conditions a and b use world knowledge as context; c and d use pictures. Conditions a and c use truth-based framing; b and d use felicity-based framing. Below: Pictures for the example item. For the sentence *Some of the trees are palms*, the control picture on the left requires an unambiguous *true/good* verification. The target picture in the middle is pragmatically ambiguous. The control picture on the right requires an unambiguous *false/bad* verification.

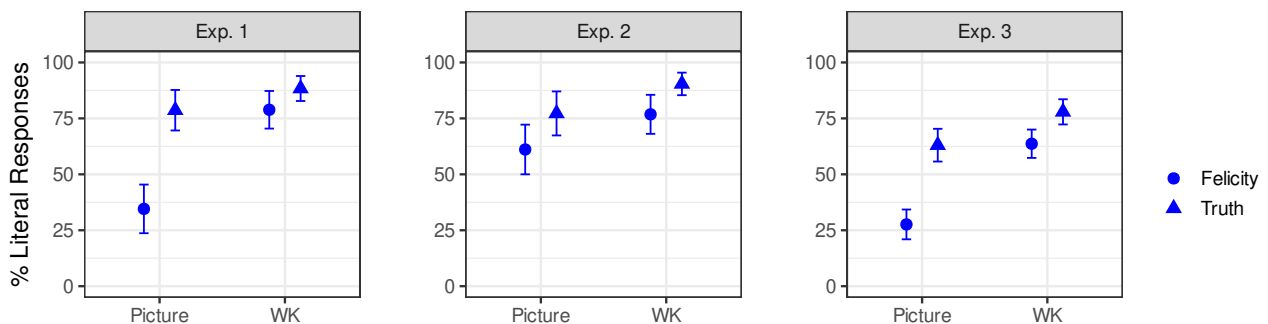


Figure 2: Average literal response rates across the four conditions in Experiments 1–3. Error bars represent 95% confidence intervals computed over subject-level means.

References

- [1] L. Bott and I. A. Noveck. "Some utterances are underinformative: The onset and time course of scalar inferences". In: *Journal of Memory and Language* 51.3 (2004), pp. 437–457. [2] H. H. Clark and W. G. Chase. "On the process of comparing sentences against pictures". In: *Cognitive Psychology* 3.3 (1972), pp. 472–517. [3] A. Cremers and E. Chemla. "Direct and indirect scalar implicatures share the same processing signature". In: *Pragmatics, semantics and the case of SIs*. Springer, 2014, pp. 201–227. [4] W. De Neys and W. Schaeken. "When people are more logical under cognitive load: Dual task impact on scalar implicature". In: *Experimental Psychology* 54.2 (2007), p. 128. [5] K. Dieussaert et al. "Some effort for some: Further evidence that scalar implicatures are effortful". In: *Quarterly Journal of Experimental Psychology* 64.12 (2011), pp. 2352–2367. [6] A. Feeney et al. "The story of some: Everyday pragmatic inference by children and adults". In: *Canadian Journal of Experimental Psychology* 58.2 (2004), pp. 121–132. [7] M. Kissine and P. De Brabanter. "Pragmatic responses to under-informative some-statements are not scalar implicatures". In: *Cognition* 237 (2023), p. 105463. [8] I. A. Noveck and A. Posada. "Characterizing the time course of an implicature: An evoked potentials study". In: *Brain and Language* 85.2 (2003), pp. 203–210. [9] B. van Tiel, E. Pankratz, and C. Sun. "Scales and scalarity: Processing scalar inferences". In: *Journal of Memory and Language* 105 (2019), pp. 93–107. [10] J. Zehr and F. Schwarz. *PennController for Internet Based Experiments (IBEX)*. 2018.