

Truth judgments vs. validity judgments

Recent work has suggested that *validity judgment tasks*, where subjects evaluate the validity of arguments, differ from *truth judgment tasks*, where subjects judge sentences as true or false vis-à-vis a depicted scenario (Coppock and Brochhagen, 2013). For example, while Geurts et al. (2010) found that only 50% of participants accepted ‘Berta had 3 beers; therefore Berta had at least 3 beers’ as a valid argument, Coppock and Brochhagen (2013) (henceforth C&B) found that ‘there are at least 3 beers in the picture’ was consistently judged as true with a picture of three beers. Coppock and Brochhagen (2013) conclude that “truth-value judgments are less sensitive to pragmatic infelicity than inference judgments,” and in particular, “ignorance implicatures do not affect truth-value judgments even though they do affect inference judgment tasks.” C&B show however that even truth judgments can be affected by particularly strong implicatures. With a picture of three beers, ‘There are at most 4 beers in the picture’ was not reliably judged true. (See top two graphs in Figure 2.) C&B attribute this to the Maxim of Depictive Sincerity, a maxim so strong that it “can cause true sentences to be judged as false”.

The present work addresses more carefully what truth and validity judgment tasks diagnose, given that there are multiple differences between the two above-mentioned settings that could potentially explain the difference in results. For example, the C&B experiments involved English rather than Dutch speakers, a presentational ‘there’ construction, was done online via Mechanical Turk, and displayed pictures, unlike the Geurts et al. ones. Such differences are eliminated here in order to determine whether truth vs. validity is really the source of the contrast, and more generally what the consequences of this methodological choice for diagnosing various types of content really are.

As a first step, validity judgment experiments matching C&B’s truth-judgment experiments were carried out, using parallel stimuli (all involving statements of the form ‘there are N Xs’ with N between 3 and 8), fillers, and design (latin-square). The results are summarized in Figure 2. (The top two graphs are C&B’s truth value study; the bottom four involve validity judgment tasks. ‘N’ is the number of objects either in the picture or mentioned in the premise of the argument.) In the ‘pictures’ experiments (e1v, e2v), an argument like ‘there are 4 bananas; therefore there are at least 4 bananas’ was accompanied by a picture of four bananas (Figure 1). The ‘without pictures’ experiments (e1vn, e2vn) showed no pictures. The results show that validity judgment tasks are not a failsafe method for detecting ignorance implicatures, as the Geurts et al. result was not replicated in these experiments, with or without pictures, although they do give more variable results overall. However, the factors that affect truth judgments also affected validity judgments: subjects were particularly reluctant to judge ‘There are 3 beers; therefore there are at most 4 beers’ as valid (e2v, e2vn). Overall, the results so far show that validity judgment tasks, like truth judgment tasks, may fail to detect ignorance implicatures, but can reliably detect strong implicatures.

Question 1/64



There are 3 penguins. Therefore there are fewer than 5 penguins.

- Agree
- Disagree

Figure 1: Example stimulus in the validity with pictures experiment

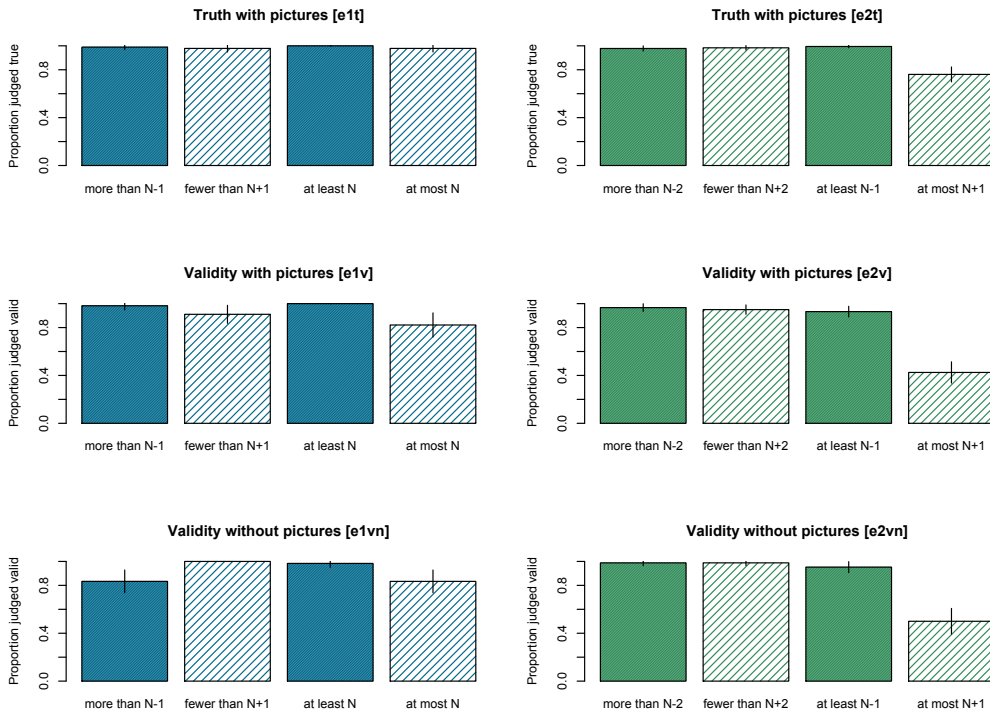


Figure 2: Results for six experiments

References

- Coppock, E. and Brochhagen, T. (2013). Diagnosing truth, interactive sincerity, and depictive sincerity. In *Proceedings of SALT 23*. eLanguage.
- Geurts, B., Katsos, N., Cummins, C., Moons, J., and Noordman, L. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes*, 25.