Is the monotonicity effect due to covert negation or pragmatic bias?

Schlotterbeck, F.S.; Ramotowska, S.; van Maanen, L.; Szymanik, J.K.

Publication date
2020

Is the monotonicity effect due to covert negation or pragmatic bias?

Fabian Schlotterbeck (University of Tübingen), Sonia Ramotowska (University of Amsterdam), Leendert van Maanen (Utrecht University) & Jakub Szymanik (University of Amsterdam)
fabian.schlotterbeck@uni-tuebingen.de

Downward entailing (DE) quantifiers (Qs) are more difficult to process than upward entailing (UE) ones [e.g. 1, 2]. What ***cognitive processes*** cause this monotonicity effect is, however, a matter of current debate. We test predictions about the underlying processes that are derived from two broad classes of competing theoretical proposals: two-step and pragmatic processing models. To this end, we model data from two verification experiments, in particular, reaction times (RT) and accuracy, using the ***diffusion decision model*** (DDM).

The DDM (Fig. 1) has been applied successfully to a large variety of decision tasks [for review, see 3; 4]. One of its strengths is that it concurrently models both accuracies and entire RT distributions. Moreover, its free parameters correspond to distinct ***components of the underlying cognitive processes***. Their estimation, therefore, allows inferences about the processing components involved in the experimental task. For the present purpose, the most important parameters are *drift rate* and *non-decision time* (for explanation see Fig. 1)*.*

Two-step processing models are based on an additional processing step in the verification of DE vs. UEQs [e.g. 1, 2, 5, 6]. One way to link such models to components of the DDM is to assume that ***monotonicity affects non-decision time*** in verification tasks. We could, e.g., think of the verification of DEQs as falsification of a suitable UE counterpart followed by a subsequent, time-consuming step of truth-value reversal. A radically different view is taken by ***pragmatic processing models*** [e.g. 7], which implement the assumption that contextual fit is a major determinant of processing difficulty. Under this view, DEQs cause processing difficulties because they are systematically dispreferred to suitable UE alternatives in various contexts [cf. 8; and also 9; 10] due to violation of pragmatic principles [cf. 11]. In such models, verification is often treated as production: Participants, in fact, judge whether they would utter the sentence to describe the context [12]. Taking into account what accumulator models have revealed in closely related domains, e.g. picture naming [13; 14], pragmatic models let us expect that ***monotonicity affects drift rate***: Slower accumulation is expected for DE vs. UEQs. This expectation is further motivated by theoretical considerations [15; 16] that link Bayesian models [cf. 7] to the DDM.

We compared verification of UE '*more than half*' (*mth*) to DE '*fewer than half*' (*fth*) in ***two web-based experiments***. The first used a visual (i.e. sentence-picture), the second a purely linguistic (i.e. sentence-sentence) verification task. In both, the participants' task was to verify the first sentence, which was read self-paced, against the information given afterwards (either in a dot picture or a sentence). Exp. 1 (N=56, 240 trials per Q) employed a factorial within-design with the two factors monotonicity (*mth* vs. *fth*) and ratio of the colored dots (28:20, 26:22, 22:26 and 20:28). The first sentence of Exp. 2 (N=72, 50 trials per Q) also contained either *mth* or *fth* and the second sentence specified a random percentage drawn from the range 1-99% (excluding 50%). Both experiments replicated the monotonicity effect (Fig. 2). When applying the DDM to the data of both experiments, we assumed monotonic relations between drift rates and (log-) ratios or percentages, respectively (specifically: a generalized logistic function). First, we fitted the DDM to Exp. 2 and used model comparisons to determine which parameters differ between Qs. Then, we applied the constrained model to Exp. 1. Model fit was good in both cases. In line with previous results [17], decision processes differed between the two tasks: Drift rate increased gradually with log-ratio in Exp. 1, whereas a step-like relation was found in Exp. 2. Apart from this difference, we found ***consistent results across the two tasks***: Non-decision times were longer for DE than UEQs (ling.: 34 ms mean diff.; $t(71) = 5.53$, $p < .001$; vis.: 43 ms; $t(55) = 5.74$, $p < .001$) and drift rates were attenuated for DE vs. UEQs (ling.: $t(71) = 9.10$, $p < .001$; vis.: $t(55) = 8.46$, $p < .001$).

Thus, our results support both two-step and pragmatic processing models. More generally, they indicate ***two separate sources of the monotonicity effect*** that map onto different DDM parameters. We will relate our results to recent studies [2, 18] that reached similar conclusions.

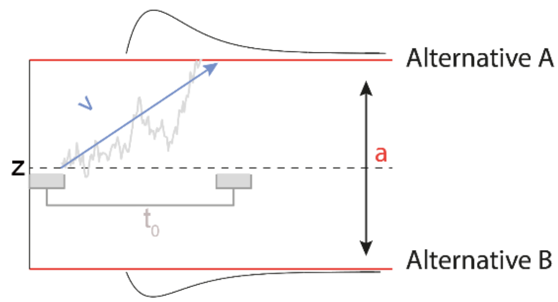

Fig. 1: Sketch of the DDM [3]. Decision processes, such as true/false judgments, are described as the accumulation of a noisy signal over time until a decision boundary is reached and a response initiated. Drift rate, *v*, determines how much information is accumulated per time and non-decision time, t_0, measures RT components that are not part of the decision process itself, e.g. stimulus encoding or response execution. In addition, the standard DDM has also a parameter, *a*, which specifies the separation between the two decision boundaries and a parameter, *z*, which determines where decision processes will start.

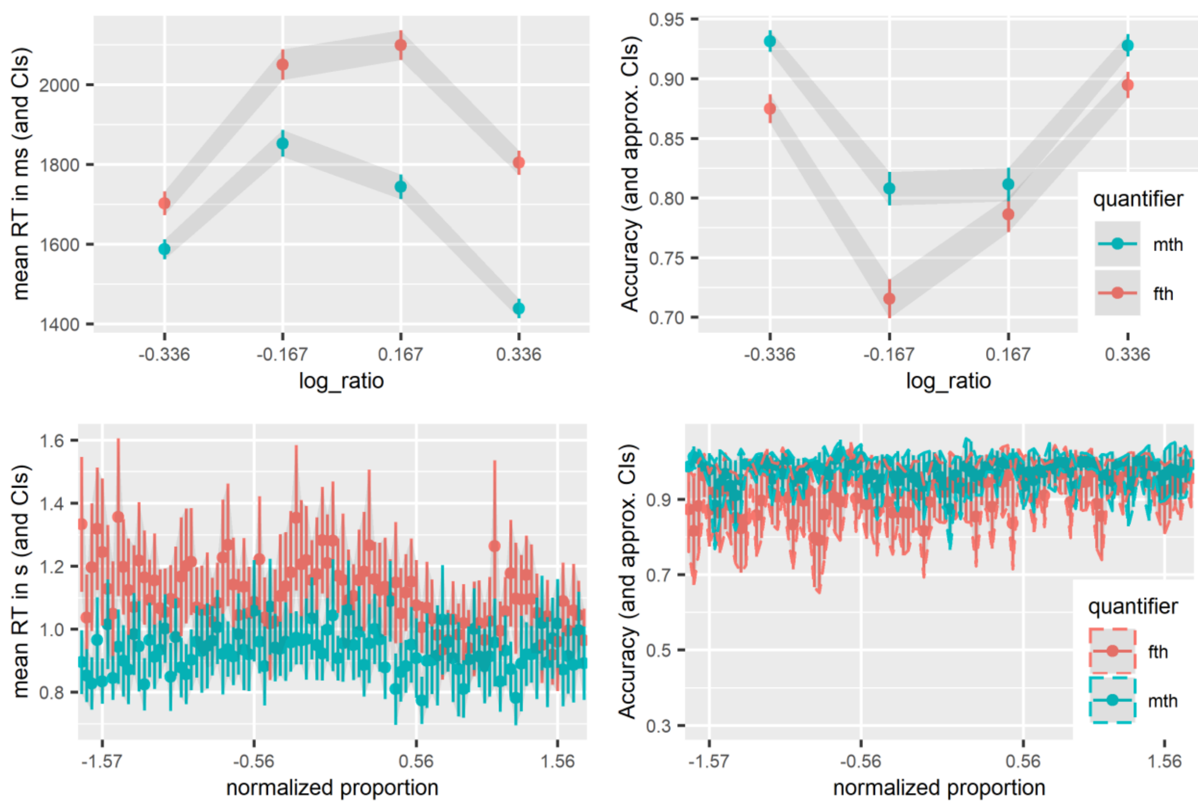

Fig. 2: RT and accuracy. Upper: visual task; lower: linguistic task. Across tasks mean RT for UE (ling.: 926 ms; vis.: 1655 ms) were faster than for DEQs (ling.: 1110 ms, vis.: 1913 ms; mixed-effects model: ling.: $\beta = 136$, $t = 5.55$, $p < .001$; vis.: $\beta = 231$, $t = 15.32$, $p < .001$) and accuracy was higher for UE than for DEQs (ling.: 97.7% vs. 92.3%; $\beta = 1.01$, $z = 5.09$, $p < .001$; vis.: 86.9% vs. 81.8%; $\beta = 0.19$, $z = 3.28$, $p < .001$).

References: [1] Deschamps et al *Cognition* (2015); [2] Agmon et al *Glossa: A Journal of General Linguistics* (2019); [3] Ratcliff & McKoon, *Neural Comput* (2008); [4] Mulder et al *Neuroscience* (2014); Just & Carpenter, *J Verb Learn Verb Be* (1971); [5] Kaup et al *J Exp Psychol* (2007); [6] Schlotterbeck *Uni Tübingen* (2017); [7] Degen & Tanenhaus in Cummins & Katsos (eds.), *Handbook of Experimental Semantics and Pragmatics* (2019)*;* [8] Nieuwland *J Exp Psychol Learn Mem Cogn* (2017); [9] Nieuwland & Kuperberg *Psychol Sci* (2008); [10] Nordemeyer & Frank *Proceedings of Cog Sci 36* (2014); [11] Grice *Logic and conversation* (1975); [12] Degen & Goodman *Proceedings of Cog Sci 36* (2014); [13] Anders et al *Cogn Psychol* (2015); [14] Anders et al *Cogn Neuropsychol* (2019); [15] Bitzer et al *Front Hum Neurosci* (2014); [16] Bogacz et al *Psychol Rev* (2006); [17] Dehaene in Haggard (ed.), *Sensorimotor Foundations of Higher Cognition* (2007); [18] Grodzinsky et al *Proceedings of Sinn und Bedeutung 22* (2018).