

## Fixing De Morgan's laws in counterfactual antecedents

Ciardelli, Zhang, and Champollion 2018 (CZC) have argued that De Morgan logical equivalencies fail in counterfactual antecedents, on the basis of a series of experiments in which counterfactuals of the form in (1a) and (1b) generate different endorsement rates in the same scenario. We report two experiments where we investigate the same type of sentences against simpler scenarios, probing both participants' truth-value judgments and their understanding of the counterfactual scenario. We find no difference between connectives, which suggests that De Morgan equivalences hold even in counterfactual antecedents. In addition, we argue that the difference in endorsement rates we found is to be attributed to other factors—e.g., the presence of negation (cf. Schulz 2018, Bar-Lev 2018).

- (1)    a.  $(\neg p \vee \neg q) \Box \rightarrow r$                       b.  $\neg(p \wedge q) \Box \rightarrow r$

**Background.** Standard semantics for counterfactuals (see a.o. Lewis 1973 and Kratzer 1981) treat them as involving universal quantification over a set of closest worlds (2). This analysis makes two predictions: (i) logically equivalent clauses are substitutable in antecedents; (ii) the inference from  $(p \vee q) \Box \rightarrow r$  to the 'simplified'  $p \Box \rightarrow r$  and  $q \Box \rightarrow r$  is invalid (Fine 1975, Nute 1975 a.o.).

- (2)     $\llbracket p \Box \rightarrow r \rrbracket^w = 1$  in  $w$  iff for every closest  $p$ -world  $w'$  to  $w$  s.t.  $\llbracket p \rrbracket^{w'} = 1$ ,  $\llbracket r \rrbracket^{w'} = 1$

SUBSTITUTION OF LOGICAL EQUIVALENTS (SLE)	SIMPLIFICATION
$p \Box \rightarrow q \models p' \Box \rightarrow q$ (with $p, p'$ equivalent)	$(p \vee q) \Box \rightarrow r \not\models p \Box \rightarrow r, q \Box \rightarrow r$

CZC conducted a series of experiments, whose results challenge this view. They presented a scenario involving two switches and a lightbulb; the light is on iff the switches are both up or both down. Participants were asked to evaluate (3a) and (3b) in a Truth Value Judgment Task (TVJ):

- (3)    a. If switch A or switch B was down, the light would be off.  
       b. If switch A and switch B were not both up, the light would be off.

In CZC's main experiment, (3a) is judged true by 69.33% of the participants included in the analysis, and (3b) by only 22.04% of them. On these grounds, CZC develop an account of counterfactuals which invalidates SLE and validates SIMPLIFICATION and can, in turn, account for the difference in TVJ ratings between (3a) and (3b). Crucially, their account appeals to an inquisitive semantics for disjunction (see a.o. Ciardelli et al. 2018), and in fact can be seen as an empirical argument for it.

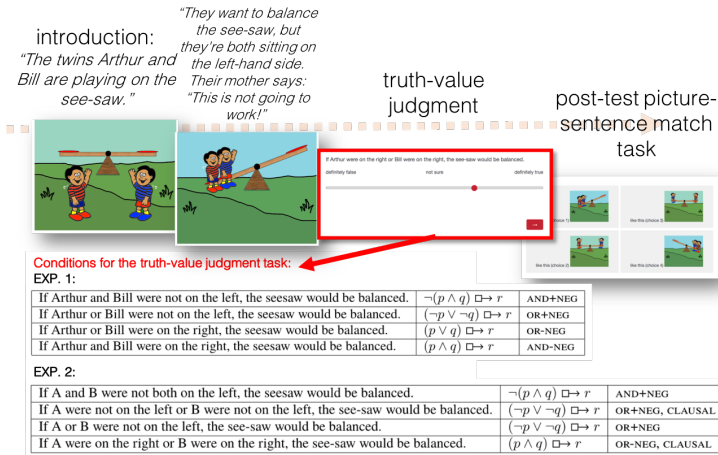
**Goals.** Our motivation for investigating this further is twofold. First, CZC's participant rejection rates ranged from 38% to 71% (rejected participants gave a 'wrong' judgment on an uncontroversial filler sentence). This suggests substantial comprehension difficulties of their scenario, casting doubts on how reliable and generalisable their results are. We probed judgments about the same type of sentences in similar but considerably simpler scenarios. Second, given the complexity of these sentences and their possible interpretations, we added a further measure to probe how speakers imagine the scenario and which alternatives they consider, in order to interpret their TVJs.

**Experiments.** **MATERIALS & PROCEDURE:** In each study, 200 adult English native speakers judged a counterfactual sentence against a simple, intuitive scenario – two children trying to balance a see-saw – in a TVJ task (see Fig.1a). Each participant was randomly assigned to one of the four conditions in each experiment. The only difference between experiments was that Exp.1 used NP conjunction and disjunction, and Exp.2 also used clausal conjunction and disjunction, to better control for the scope of negation. After the TVJ, participants were asked to select pictures that would match the counterfactual supposition corresponding to the antecedent of the sentence they just evaluated, e.g., "What would it look like if Arthur or Bill were on the right?".

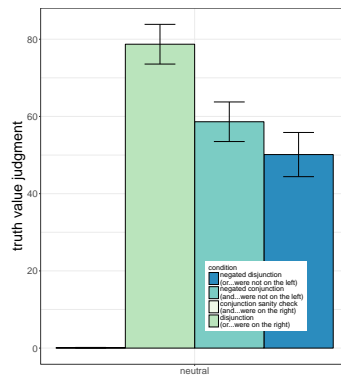
**RESULTS:** Bonferroni-corrected Mann-Whitney tests on contrast-coded conditions showed that **Exp.1** resulted in significant differences between all comparisons (all  $ps < .006$ ), except between the conditions that contained negation ( $\neg(p \wedge q)$  vs.  $\neg(p \vee q)$ ;  $p > .8$ ). In **Exp.2**, only the effect of negation was significant ( $p < .02$ ), but we did not find an effect of clausality or disjunction vs. conjunction ( $ps > .13$ .) Importantly, the picture-matching task revealed that in both studies, people considered two alternatives (A and B on different sides, red in Figs.1d & 1e) at a similar rate as three alternatives (A and B on different sides and A and B on the right, green in Fig.1d & 1e). Moreover, *only* in negated sentences, but regardless of whether the sentence contained a conjunction or a disjunction, did participants consider all three alternatives; and considering all three alternatives was correlated with lower truth-value judgments (y-axis in Figs.1d & 1e).

**Discussion.** Our results, unlike those of CZC, show no difference between negated conjunctions and disjunctions in the antecedent of counterfactuals, with high endorsement rates for both. Additionally, we found lower endorsement rates for those cases involving overt negation, regardless of the connective involved. Finally, the additional picture-matching measure allows us to better understand the participants' interpretation of these sentences. Overall, our results suggest three conclusions. (i) In at least some scenarios,  $\neg(p \wedge q)$  and  $(\neg p \vee \neg q)$  are equivalent in counterfactual antecedents. (ii) Even if  $\neg(p \wedge q)$  and  $(\neg p \vee \neg q)$  were not equivalent in some other scenarios, as CZC maintain, our results challenge the idea that this is due to the lexical meaning of *or* (since on this hypothesis it is not clear why we would observe variation). (iii) Negation affects endorsement rates. At this stage, this is consistent with negation simply increasing processing complexity or with it playing a more specific role in these configurations (see e.g. Schulz 2018 and Bar-Lev 2018).

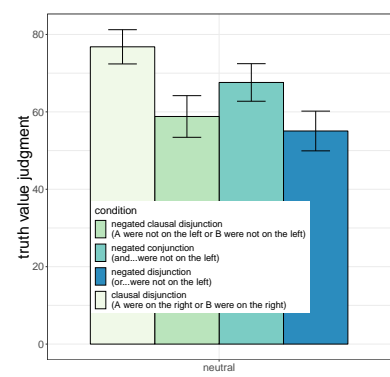
**References.** Bar-Lev, *Free Choice, Homogeneity, and Innocent Inclusion*, 2018 · Ciardelli, Groenendijk and Roelofsen, "Inquisitive Semantics", 2018 · Ciardelli, Zhang and Champollion, "Two switches in the Theory of Counterfactuals", 2018 · Schulz 2018. "The similarity approach strikes back - negation in counterfactual conditionals".



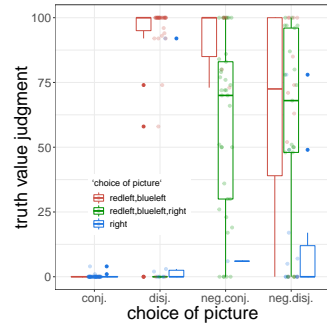
(a) Trial structure and conditions.



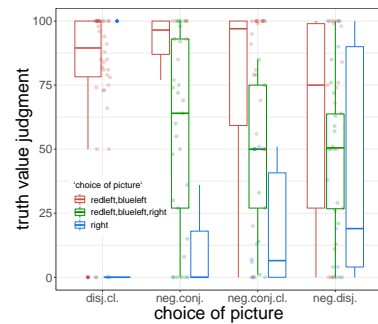
(b) Results of Exp.1 (TVJ).



(c) Results of Exp.2 (TVJ).



(d) Picture Choices Exp.1.



(e) Picture Choices Exp.2.

Figure 1: Design and results.