The consideration of alternatives during incremental comprehension of counterfactuals

Ebru Evcen (UC San Diego) & Eva Wittenberg (UC San Diego)
eevcen@ucsd.edu

Counterfactuals such as "If there had been zebras, there would have been lions" force comprehenders to construe two alternative situations: (i) the *suppositional* but factually false state ([+ZEBRA,+LION]), and (ii) the *implied actual state* ([-ZEBRA,-LION], Byrne, 2002). The computational cost of considering both of these meanings, and the suppression of the suppositional interpretation have been subject to much research, but it still is an open question whether and at which point comprehenders actually consider each of these alternative states as the sentence unfolds, and specifically, how and when they reach the correct *actual state* interpretation. According to most incremental sentence processing theories, the suppositional state should be dismissed as soon as all morpho-syntactic cues are processed, but it has conversely been argued that the commitment to one of the alternatives is delayed until sentence end (Ferguson & Sanford, 2008).

In four preregistered studies, we test these competing predictions by replicating and extending an earlier study (Orenes et al., 2019), which presented displays such as in Fig.1 while participants heard sentences containing indicative conditionals or counterfactuals. Results showed that participants, while hearing indicatives, only looked at the suppositional state. For counterfactuals, roughly half of participants looked only at the suppositional state, whereas the other half either looked only at the factual state or at both states. In **Exp.1a** (N=57/48, Fig.2), we conducted a loyal replication this study, translated into English, using participants' webcams. Unlike in the original, our participants only ever looked at the suppositional state during counterfactuals (from 550ms, $t(22)=-3.26$, $p=0.03$), just as during indicatives (from 400ms, $t(22)=-2.73$, $p=0.009$).

However, there are three problems with this design: First, it is unclear how participants actually interpreted these sentences, and specifically, whether they committed to the factual state, because interpretation was not measured; second, crossed-out images were unnatural (crossed-out images do not inherently represent the meaning of the negation: the negation of [+ZEBRA,+LION] would be a blank page); and third, it fails to capture the temporal course of counterfactual comprehension, since there were only two pictures containing a zebra at all; as soon as 'zebras' is encountered, participants can dismiss the other set of pictures. In the next three studies, we address these problems through step-wise changes in the design.

In **Exp.1b** (N=26/48, Fig.3), we explicitly shift the QUD to the *actual state* interpretation, by asking participants explicitly what the actual state would look like. Crucially, now 9 participants considered all pictures equally until hearing 'zebras', and only then both fixated and clicked on the actual state ([-ZEBRA,-LION]). 17 participants looked only at the suppositional state ([+ZEBRA,+LION] and the alternative state [+GIRAFFE,+ELEPHANT]), and only then both fixated and clicked on the suppositional state*.* This pattern confirms that most participants did not consider the crossed-out images as alternatives, and that they interpreted the counterfactual as describing the suppositional, rather than the actual state. It is therefore difficult to draw inferences from this design about how alternatives are considered in incremental processing of counterfactuals.

Therefore, **Exp.2a and 2b** (N=48ea) create a visual world (Fig.4) that consists of images depicting a) the actual state [-ZEBRA,-LION], b) the suppositional state [+ZEBRA,+LION], c) an antecedent-only state [+ZEBRA,-LION], and d) a consequent-only state [-ZEBRA,+LION]. We predict that as the sentence unfolds, participants will look at the pictures containing the referents mentioned (ZEBRA, then LION). Crucially, however, if participants take counterfactual cues into consideration as soon as possible, then they should disregard both images containing ZEBRA as soon as the antecedent is over, and focus on the two remaining pictures to settle on the picture representing the actual state after hearing the next referent (LION). Conversely, if commitment to one of the alternatives is delayed until sentence end, then participants should consider the suppositional world throughout, and focus on its opposite, actual state only at sentence end.
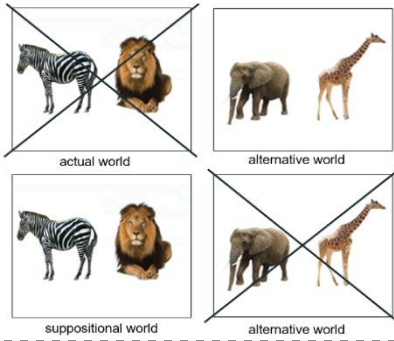
Condition	Example Sentence
Indicative (Control Cond.)	Si *hay* habido cebras, entonces *hay* leonas. *If there are zebras, then there are lions.*
Counterfactual	Si *hubiera habido* cebras, entonces *habria habido* leonas. *If there had been zebras, then there would have been lions.*

Fig. 1. Experimental stimuli used in Experiment 1a & 1b.
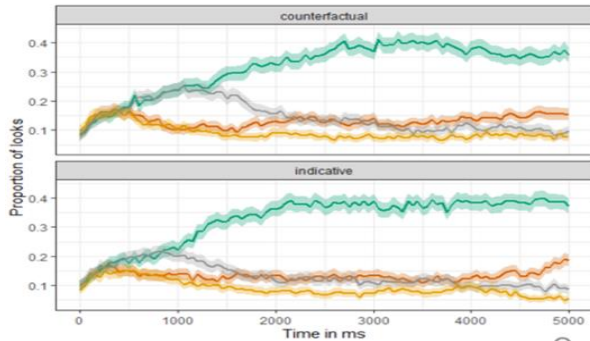

Fig. 2. Probabilities of fixations for counterfactuals and indicatives (Exp. 1a). All parameters are the same as in the original study, leading to indicatives (the control items) being shorter than counterfactuals, and therefore an earlier divergence of looks to suppositional (green) vs. alternative (gray) state. Otherwise, patterns of looks are similar between conditionals and counterfactuals.
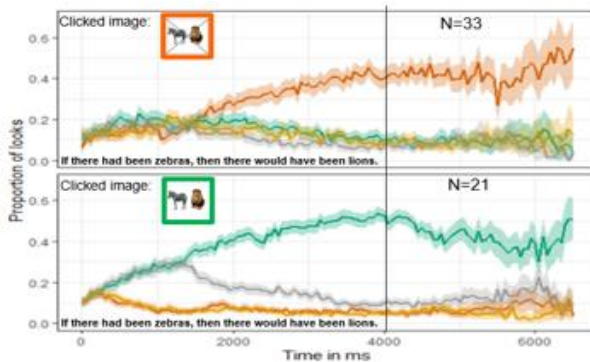

Fig. 3. Probabilities of fixations for counterfactuals by subgroups (Exp. 1b). Here, the additional task, to click on the actual state, created two subgroups: one which considered only non-crossed-out pictures (above), and one that considered all pictures equally and then clicked on the factual state (below). Indicative results are not shown.

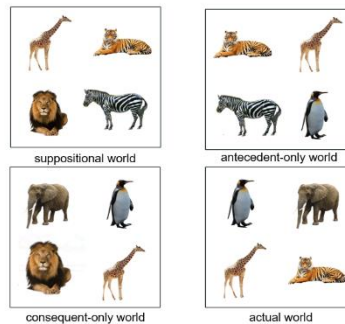| Condition	Example Sentence
Counterfactual	If there had (not) been zebras, then there would (not) have been lions.
Declarative (Control Cond.)	There were (no) zebras and there were (no) lions.

Fig. 4. Stimuli in Exp. 2a&b, to better track the incremental construal of counterfactuals by individually and jointly presenting antecedent and consequent.

References: Byrne, 2002, Mental models and counterfactual thoughts about what might have been, *TiCS*; Orenes et al., 2019, The comprehension of counterfactual conditionals: Evidence from eye-tracking in the visual world paradigm. *Frontiers in Psychology*; Ferguson & Sanford, 2008, Anomalies in real and counterfactual worlds: An eye-movement investigation. *JML*.