Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–19

DOI: 10.1177/17470218211044085
qjep.sagepub.com
SAGE

Similarities and differences in understanding negative and affirmative counterfactuals and causal assertions: Evidence from eye-tracking

Isabel Orenes[1], Orlando Espino[2] and Ruth MJ Byrne[3]

Abstract

Two eye-tracking experiments compared affirmative and negative counterfactuals, "if she had (not) arrived early, she would (not) have bought roses" and affirmative and negative causal assertions, "Because she arrived (did not arrive) early, she bought (did not buy) roses." When participants heard a counterfactual, they looked on screen at words corresponding to its conjecture ("roses"), and its presupposed facts ("no roses"), whereas for a causal assertion, they looked only at words corresponding to the facts. For counterfactuals, they looked at the conjecture first, and later the presupposed facts, and at the latter more than the former. The effect was more pronounced for negative counterfactuals than affirmative ones because the negative counterfactual's presupposed facts identify a specific item ("she bought *roses*"), whereas the affirmative counterfactual's presupposed facts do not ("she did not buy roses"). Hence, when participants were given a binary context, "she did not know whether to buy roses or carnations," they looked primarily at the presupposed facts for both sorts of counterfactuals. We discuss the implications for theories of negation, the dual meaning of counterfactuals, and their relation to causal assertions.

Keywords

Counterfactuals; causals; conditionals; negation; eye-tracking

Received: 4 December 2020; revised: 30 June 2021; accepted: 5 July 2021

In everyday life people often think about how things could have turned out differently (e.g., De Brigard et al., 2013; Kahneman & Tversky, 1982; Markman et al., 2008; Roese & Epstude, 2017). They frequently communicate their imagined alternatives in counterfactual conditionals such as, "if she had arrived early, she would have bought roses." What possibilities do listeners immediately envisage in the first few seconds after they hear a counterfactual? For an affirmative counterfactual, they simulate two possibilities, its conjecture, "she bought roses," and its presupposed facts, "she did not buy roses" (for a review, see Byrne, 2016). The dual meaning of counterfactuals is reflected in the inferences reasoners make (e.g., Byrne & Tasso, 1999; Thompson & Byrne, 2002), in their reading times (e.g., Santamaria, Espino, & Byrne, 2005), and in brain activity (e.g., Kulakova et al., 2013; Nieuwland, 2012; Urrutia et al., 2012). Online comprehension studies show that people rapidly process the conjecture and the presupposed fact in the first few seconds (Ferguson, 2012; Ferguson & Cane, 2015), notwithstanding individual differences (Orenes et al., 2019). However, although both possibilities are rapidly activated, after a short time the conjecture tends to be focused on somewhat less than the presupposed facts (de Vega et al., 2007; de Vega & Urrutia, 2011, 2012; Ferguson & Jayes, 2018; Urrutia et al., 2012; for a review, see Kulakova & Nieuwland, 2016). It is plausible that people try to avoid maintaining mental representations that correspond to multiple possibilities—it is cognitively difficult to do so given the limitations of working memory (e.g., Johnson-Laird & Byrne, 2002). Our aim in this article is to examine two cases in which people focus readily on the possibility corresponding to the presupposed facts when they understand a counterfactual.

[1]Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain
[2]Universidad de La Laguna, Tenerife, Spain
[3]Trinity College Dublin, University of Dublin, Dublin, Ireland

Corresponding author:
Isabel Orenes, Universidad Nacional de Educación a Distancia (UNED), Psicología, C/Juan del Rosal, 10, 28040 Madrid, Spain.
Email: iorenes@psi.uned.es

The first case is a negative counterfactual, such as, "If she had not arrived early, she would not have bought roses." We expect that when people understand the negative counterfactual, they think about the conjecture, e.g., "she did not arrive early and she did not buy roses," and they also think about the presupposed facts, e.g., "she arrived early and she bought roses," just as they think about dual possibilities when they understand an affirmative counterfactual (see Byrne, 2016). We test the novel hypothesis that people focus more on the presupposed facts of a negative counterfactual, compared with an affirmative counterfactual, in the first few seconds after they have heard the counterfactual. The presupposed fact for the negative counterfactual corresponds to "she bought roses," and it enables people to mentally represent the *specific* flower she bought, "roses." In contrast, when they understand an affirmative counterfactual, "if she had arrived early, she would have bought roses" they think about the conjecture, e.g., "she arrived early and she bought roses," and the presupposed facts, e.g., "she did not arrive early and she did not buy roses." The presupposed fact for the affirmative counterfactual corresponds to, "she did not buy roses," and it does not enable people to mentally represent the specific flower she bought, if any, from the potentially large set of other possible flowers, e.g., she could have bought lilies, or carnations, or tulips, or any of a large set of possible alternatives. We suggest that this difference in the specificity of the presupposed fact for a negative counterfactual compared with an affirmative counterfactual leads to a difference in participants' maintenance of the mental representations corresponding to the conjecture and the presupposed fact. The presupposed facts of a negative counterfactual can be grasped in the first few seconds ("roses") without needing to continue to allocate as much attention to the conjecture ("no roses"). But for the affirmative counterfactual, people need to keep in mind both possibilities because the presupposed facts ("no roses") are not specified, and so continued allocation of attention to the conjecture ("roses") is required to ensure the meaning of the counterfactual is grasped (see Espino & Byrne, 2018; Khemlani et al., 2012; Orenes et al., 2014). Our aim in the experiments we report is to examine, using eye-tracking methods, what words people look at in the immediate few seconds after they hear a negative counterfactual, compared with an affirmative one. We examine whether people focus more on the presupposed facts for a negative counterfactual than the affirmative one.

A related question is, are negative counterfactuals easier or more difficult to understand than affirmative ones? Perhaps the most obvious hypothesis is that they should have longer processing times than affirmative ones. When readers understand a negative sentence, such as, "she did not arrive early," they first represent what is mentioned, "she arrived early" and then the operation of negation is applied to represent what actually happened "she did not arrive early." Hence, negative sentences have longer processing times and higher error rates compared with affirmative ones (Carpenter & Just, 1975; Clark & Chase, 1972; Farshchi et al., 2019; Greco, 2020; Hasson & Glucksberg, 2006; Kaup & Dudschig, 2020; Trabasso et al., 1971; Wason & Johnson-Laird, 1972; Wason & Jones, 1963). Moreover, reasoners have difficulty in making inferences from negated compound assertions, such as negated conditionals, conjunctions, or disjunctions (Espino & Byrne, 2012; Khemlani et al., 2012, 2014; Macbeth et al., 2014, 2015, 2017; Moreno-Rios & Byrne, 2018; Orenes & Santamaría, 2014). However, an alternative, intriguing hypothesis is that negative counterfactuals should have the same or even faster processing times than affirmative ones. Negative sentences are more difficult than their corresponding affirmative ones because negative sentences require two mental representations whereas affirmative ones require only one. But this distinction does not apply to counterfactuals. Even affirmative counterfactuals require people to think about two possibilities and so the difference in the required number of representations for affirmative and negative sentences does not hold for counterfactuals. Indeed, negative counterfactuals could have even faster processing times than affirmative ones if they are understood by immediately focusing on the presupposed facts, as we have suggested. Hence, we also test the hypothesis that negative counterfactuals are processed at least as quickly as affirmative ones.

The second case we examine in which people may focus immediately, and even only, on the possibility corresponding to the presupposed facts concerns counterfactuals presented in a *binary* context, such as "she did not know whether to buy roses or carnations." A distinction is commonly made between binary (complementary) predicates, such as "odd" and "even," which have just a single alternative possibility, and multiple (contrastive) predicates, such as "red," "green," and "blue,", which have many alternatives (Wason, 1959, 1961). For an affirmative counterfactual, e.g., "if she had arrived early, she would have bought roses," the presupposed facts are "she did not buy roses," and the binary context allows people to mentally represent the specific flower she bought, "she bought *carnations*." Several cognitive steps are required to understand that "she did not buy roses" means "she bought carnations" in the binary context (Johnson-Laird & Tridgell, 1972). For a negative counterfactual in the binary context, e.g., "if she had not arrived early, she would not have bought roses," the presupposed facts are "she bought *roses*." An additional cognitive step is required to infer the specific flower for the affirmative counterfactual compared with the negative one, in the binary context. It follows that the representation of the presupposed fact ("no roses") by the alternate for affirmative counterfactuals ("carnations") will be somewhat slower than the representation of the presupposed fact for negative counterfactuals ("roses"). But nonetheless, overall, people can represent an alternate for binary predicates for which an alternate is accessible (Beltrán et al., 2008; Kaup et al., 2006; Kaup & Zwaan, 2003; Mayo et al., 2004; Nordmeyer & Frank, 2014; Orenes et al., 2014). For example, the inferences people make from a counterfactual, such as "If she had arrived early, she would have bought roses," when it is presented after a binary context statement, such as "she did not know whether to buy roses or carnations" indicate that they have represented the presupposed facts by an alternate, e.g., "she bought carnations," rather than by an explicit negation "she did not buy roses" (Espino & Byrne, 2018). Hence, in a binary context even the affirmative counterfactual may be understood by people increasing their focus on the presupposed facts, because it corresponds to a specific item, e.g., "carnations." We test the hypothesis that in a binary context, people rapidly update their attention in the first few seconds after they have heard a counterfactual, to increase focus on the presupposed facts, whether affirmative or negative.

In both of the experiments we report, we also compare counterfactuals to causal assertions. It has long been assumed that a counterfactual such as "If A had not happened, B would not have happened" is closely related to the causal assertion, "A caused B" (Gerstenberg et al., 2017; Lagnado et al., 2013; Lucas & Kemp, 2015; McEleney & Byrne, 2006; Meder et al., 2009, 2010; but see Mandel & Lehman, 1996). We examine the comparability (in terms of what listeners look at) of a negative counterfactual "if she had not arrived early, she would not have bought roses" to an affirmative causal assertion, such as "because she arrived early, she bought roses," and of an affirmative counterfactual "if she had arrived early, she would have bought roses" to a negative causal assertion, "because she did not arrive early, she did not buy roses."

The experiments rely on the study of fixations through eye-tracking which is generally considered to provide a measure directly related to the mental representations that are activated in working memory (Cooper, 1974; Huettig & Altmann, 2011; Just & Carpenter, 1976; for reviews, see Huettig et al., 2011; Tanenhaus et al., 1995). Eye movements are an especially useful online tool to study when people represent the conjecture and when they represent the presupposed fact for counterfactuals (e.g., Ferguson, 2012; Ferguson et al., 2008, 2010; Ferguson & Sanford, 2008; Orenes et al., 2019; Stewart et al., 2009). They are also especially useful to assess whether the presupposed fact is represented by the explicitly negated words ("no roses") or by the alternate ("carnations"). The "visual world paradigm" provides a listener with four images on screen while they hear an assertion, as eye-tracking cameras register what they look at when the verbal input is presented simultaneously (on computer speakers) with the visual image (on screen). Most visual world studies have presented images (e.g., Allopenna et al., 1998; Duñabeitia et al., 2009; Orenes et al., 2015; Rayner, 1998), but some have used printed words (Huettig & McQueen, 2007, 2011; Ito, 2019; McQueen & Viebahn, 2007; Primativo et al., 2016; Salverda & Tanenhaus, 2010), including printed words corresponding to heard counterfactuals (Orenes et al., 2019; Experiment 3). The two experiments we report examine affirmative and negative counterfactuals and causal assertions using eye-tracking in a visual world paradigm that presents printed words.

Experiments 1a and 1b: affirmative and negative counterfactual and causal assertions

The first experiment compared participants' understanding of negative counterfactual conditionals to their understanding of affirmative ones; it also compared their understanding of counterfactuals to affirmative and negative causal assertions. Participants heard an assertion on a computer speaker, e.g., "if she had arrived early, she would have bought roses," while they looked at printed words on a computer screen. They saw four printed words or phrases, the affirmative mentioned word, e.g., "roses," the negative phrase, e.g., "no roses," and two distractors—an affirmative distractor, e.g., "flowers," and a negative distractor, e.g., "no flowers." We measured where participants looked on screen when they heard the last word in the sentence, "roses," e.g., "if she had arrived early, she would have bought roses." We made the following predictions:

1. *Affirmative counterfactuals*: For a counterfactual, such as "if she had arrived early, she would have bought roses," we expect to replicate previous observations that participants maintain a constant tendency from the outset to look at the written word corresponding to the *conjecture*, e.g., "roses," throughout the period of eye-tracking measurements, and they *increase* looking at the negative phrase corresponding to the *presupposed facts*, e.g., "no roses" (Orenes et al., 2019, Experiment 3, for printed words). This pattern has been interpreted as indicating that participants think very rapidly about the conjecture and the presupposed facts.
2. *Negative counterfactuals*: One interpretation of the results for affirmative counterfactuals is that participants rapidly focus on the affirmative word "roses" and increase their focus on the negative phrase "no roses" merely because of differences in preferences for looking at affirmative words compared with negative phrases. If this interpretation is correct, then the opposite pattern should be observed for negative counterfactuals, such as "if she had *not* arrived early, she would *not* have bought roses." On this interpretation, participants

from the outset will maintain a steady tendency to look at the word corresponding instead to the *presupposed facts* for a negative counterfactual, e.g., "roses," because it is an affirmative word; and they will show an increasing tendency to look at the words corresponding to the *conjecture*, e.g., "no roses," because it is a negative phrase. But an alternative interpretation is that the result for affirmative counterfactuals reflects a genuine preference for thinking about the *conjecture first* and then the *presupposed facts second*. If so, then the *same* pattern will be observed for negative counterfactuals as for affirmative ones: Participants will maintain a steady tendency to look at the words corresponding to the conjecture, this time "no roses," even though it is a negative phrase, and they will show an increasing tendency to look at the word corresponding to the presupposed facts, this time "roses."

3. *Affirmative and negative causal assertions*. For an affirmative causal assertion, such as "because she arrived early, she bought roses" we expect that participants will look at the affirmative mentioned word, e.g., "roses," and they will not look at any of the other three words or phrases. For a negative causal assertion such as, "because she did not arrive early, she did not buy roses," we expect that they will look at the negative phrase "no roses" and they will not look at any of the other three words. The predictions in each case are based on the proposal that when people understand a causal assertion, they think initially only about the facts it describes (e.g., Byrne, 2005; Frosch & Byrne, 2012; Goldvarg & Johnson-Laird, 2001; see Johnson-Laird & Khemlani, 2017). We also predict that participants will look at the appropriate words in each case very rapidly, but that they will do so more rapidly for affirmative causals than for negative causals, given the long-standing observations of increased difficulty in understanding negation. However, we expect to observe no such advantage for affirmative counterfactuals over negative ones, given that counterfactuals require participants to consider both the affirmative and negative possibilities.

Method

Participants. The participants in Experiment 1a were 45 volunteers (35 women and 10 men) who were students at the University of La Laguna in Tenerife, Spain. Their average age was 19 years, with a range from 17 to 53 years. In Experiment 1b, the participants were 26 volunteers (20 women and 6 men) who had not taken part in Experiment 1a. Their average age was 22 years, with a range from 19 to 59 years. The sample size of 45 participants in Experiment 1a provides 95% power at alpha .05 to detect a medium-sized effect for a one-tailed test (91% power for a two-tailed test), and the sample size in Experiment 1b of 26 participants provides 80% power at alpha .05 to detect a medium-sized effect for a one-tailed test (69% for a two-tailed test), according to G*power (Faul et al., 2009). The participants in both experiments were native Spanish speakers and they all reported normal or corrected to normal vision (glasses or contact lenses). They participated in the experiment in exchange for course credits. These experiments and the subsequent one received ethical approval prior to their commencement from the Committee on Ethics in Research and Animal Welfare (Comité de Ética de la Investigación y Bienestar Animal) at the University of Laguna.

Materials and design. In both experiments, participants received 36 vignettes with different contents in their native Spanish language (adapted from Orenes et al., 2019, see the Online Supplementary Materials). In each experiment, polarity (affirmative and negative) of the target sentence (counterfactuals in Experiment 1a and causal assertions in Experiment 1b) and polarity of the last sentence in the vignette were manipulated within participants. Each vignette contained three sentences, the first sentence provided an opening scene, e.g., "Carmen went to the flower shop" (translated from "Carmen fue a la floristería"). In Experiment 1a, the second sentence contained either an affirmative counterfactual, e.g., "If she had arrived early, she would have bought roses/flowers" ("Si hubiera llegado pronto, habría comprado rosas/flores"), or a negative one, e.g., "If she had *not* arrived early, she would *not* have bought roses/flowers" ("Si no hubiera llegado pronto, no habría comprado rosas/flores"). In Experiment 1b, the second sentence contained either an affirmative causal assertion, "Because she arrived early, she bought roses/flowers" ("Como llegó pronto, compró rosas/flores"), or a negative one, "Because she did *not* arrive early, she did *not* buy roses/flowers" ("Como no llegó pronto, no compró rosas/flores"). These sentences were prerecorded and presented via a computer speaker. When the second sentence was presented, four printed words or phrases were shown on a computer screen, e.g., "roses," "no roses," "flowers," and "no flowers." The third sentence concluded the trial and comprised either an affirmative assertion, e.g., "Carmen bought roses/flowers" ("Carmen compró rosas/flores"), or a negative one, e.g., "Carmen did not buy roses/flowers" ("Carmen no compró rosas/flores"). This third sentence replaced the four words on the screen for participants to read (see Figure 1).

We constructed eight versions of each vignette that varied in the words that were presented in the second assertion (roses, no roses, flowers, no flowers) and in whether the last sentence was affirmative or negative. We

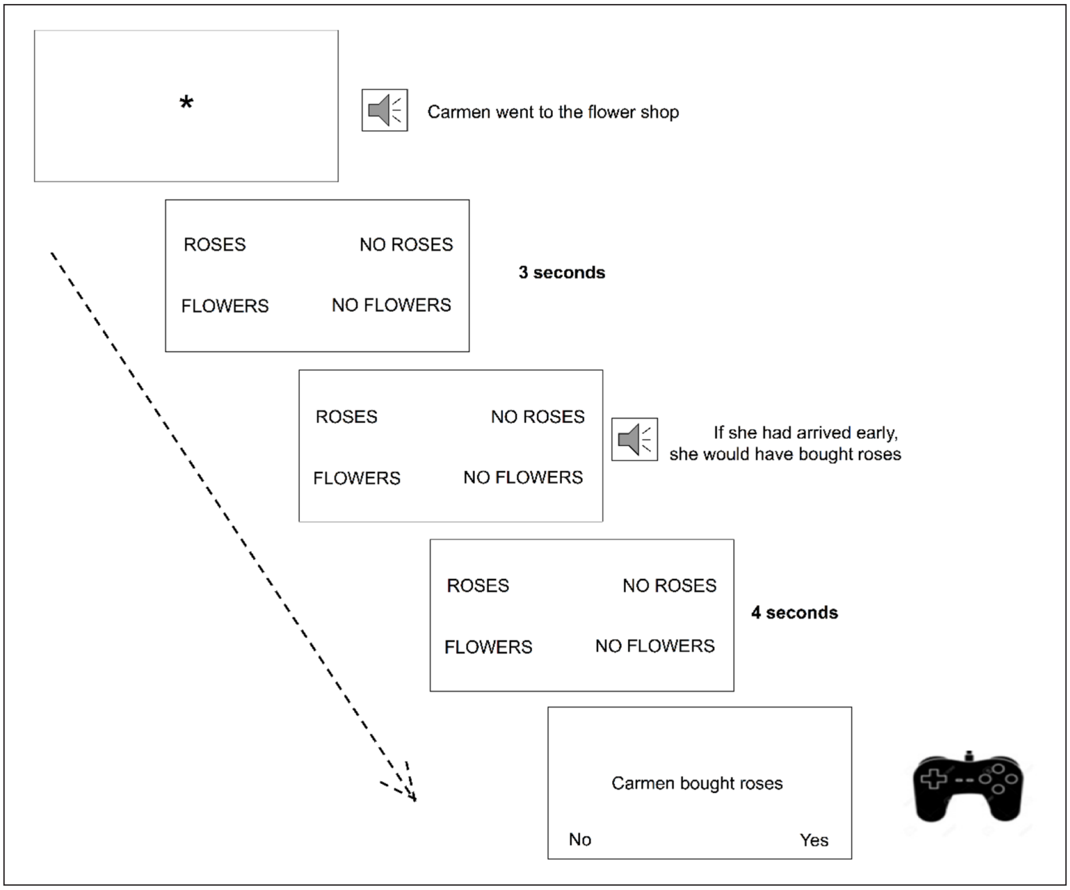

Figure 1. Illustration of the procedure in Experiment 1a. Participants heard a first sentence as an orientating asterisk appeared in the centre of the screen. Next they saw four printed words or phrases on screen for 3 s, then they heard a second sentence as the printed words remained on screen, and then they continued to see the printed words for a further 4 s. Finally, they saw a third sentence on screen, to which they responded yes or no.

constructed eight sets in a Latin-square design to ensure that each participant received only one of the eight possible versions for each content (see Table 1). Given the variety and balance of the materials in the four distinct conditions, no fillers were required. The order of the trials was randomised for each participant. Participants were asked to indicate whether the final sentence was consistent with the previous sentences or not.

They answered by pressing either a "yes" or a "no" button on a gamepad, which was counterbalanced by participant. They received feedback for each answer, as correct or incorrect. The effects of the polarity of the last sentence were tested by the response accuracy and latency of the behavioural data; the effects of the polarity of the target sentence were tested by the fixation probabilities of eyes in the eye-tracking data.

Procedure. Participants listened to the 36 trials on a computer speaker while looking at a computer screen and at the end of each vignette they answered a simple question about it. They were instructed that they should listen to the sentences carefully and that they should not take their eyes off the screen throughout the experiment. Their eye movements were recorded at a rate of 500 Hz using an SR Research EyeLink II head-mounted eye-tracker connected to a 21 colour CRT for visual stimulus presentation. Procedures were implemented in SR Research Experiment Builder. Calibration and validation procedures were carried out at the beginning of the experiment and were repeated several times per session. Trials started with the presentation of a central fixation dot for drift correction while participants listened to the opening-scene sentence (e.g., "Carmen went to the flower shop"). After that, a visual display with the four printed words or phrases appeared (e.g., roses, no roses, flowers, no flowers). After the visual display had been on screen for 3 s, the target sentence was heard (e.g., "if she had arrived early, she would have bought roses"). The trial concluded with the final sentence (e.g., "Carmen bought roses") and their task was to read it and answer whether or not it was consistent with the previous sentences (see Figure 1). There was a practice block of four trials before the experiment proper started. Participants were tested individually in a quiet testing room and each experimental session lasted approximately 30 min.

Table 1. Examples of the eight versions of the verbal description of each content in Experiment 1a, illustrated for the roses and flowers content. Participants heard a first sentence, e.g., "Carmen went to the flower shop," then they heard one of the four second sentences described below, and then read one of the eight third sentences described below.

Second sentence	Third sentence
Affirmative—roses	
If she had arrived early, she would have bought roses	1. Carmen bought roses
	2. Carmen did not buy roses
Negative—roses	
If she had *not* arrived early, she would *not* have bought roses	3. Carmen bought roses
	4. Carmen did not buy roses
Affirmative—flowers	
If she had arrived early, she would have bought flowers	5. Carmen bought flowers
	6. Carmen did not buy flowers
Negative—flowers	
If she had *not* arrived early, she would *not* have bought flowers	7. Carmen bought flowers
	8. Carmen did not buy flowers

Results and discussion

Eye-tracking data coding. The data for these experiments and the subsequent one is available at https://osf.io/79kcr/. The analysis of fixations was time-locked to the onset of the last word of the target sentences (counterfactuals or causal assertions), e.g., the onset of "roses," to 3,500 ms after that word. The periods were divided into 50 ms time slots and for each time slot, the number of fixations on each rectangle quadrant corresponding with each word or phrase was counted and converted into fixation probabilities. The number of fixations on each word or phrase was divided by the sum of the fixation on the four words or phrases. We excluded measures of saccades from the analyses. To avoid problems inherent in proportional data, participant averages were arcsin-transformed prior to *t*-test comparisons. Given that 180–200 ms are usually assumed to account for saccade programming (Martin et al., 1993), the mean of the first time-region (0–100 ms) was considered to be the baseline and was used to conduct statistical comparisons against means on each time point until 3,500 ms later (for a similar method, see Huettig & Altmann, 2011). This correction to baseline allowed us to control for any bias in the pattern of fixations on printed words caused by the type of assertion. A false discovery rate (FDR) thresholding procedure was used to effectively control for Type 1 error due to multiple comparisons (70 for each condition; see Genovese et al., 2002). This analysis allows us to consider in the most straightforward manner the temporal point at which the proportion of fixation changes with regard to the baseline. The method has been the conventional one for eye movement analyses (since Huettig & Altmann, 2011). To counteract any statistical deficiency of this simple analysis, we also used the growth curve analysis, which is frequently used for visual world paradigm data (more so than the similar method of beta regression using a Dirichlet distribution). Growth curve analysis (Mirman, 2014) analyses the whole period through the trajectory of the curve, not point by point as *t*-tests against baseline. The curve has four elements that we report: the intercept (total fixations), the linear or slope (how fast the curve increases), the quadratic (rise and fall rate of fixation ratios around the central inflection point), and the cubic (the sharpness of the two peaks). The overall time course of target fixations was captured with third-order (cubic) orthogonal polynomial terms and fixed effects of condition within participants on all time terms. The model also included participant and participant-by-condition random effects on all time terms. Statistical significance (*p*-values) for individual parameter estimates was assessed using the normal approximation (treating the *t*-value as a *z*-value). All analyses were carried out in R version 3.5.3 using the lme4 package. Following our hypothesis, we compared the fixations on the words corresponding to the conjecture and the presupposed facts for affirmative and negative counterfactuals, and we also examined whether there are differences on fixations of the presupposed facts/facts between affirmative and negative counterfactual and causal assertions.

T*-tests against the baseline of 0–100 ms and growth curve analysis*

Affirmative counterfactuals. In Experiment 1a, for an affirmative counterfactual, e.g., "If she had arrived early, she would have bought roses," participants focused at the outset on the four words or phrases on the screen (e.g., roses, no roses, flowers, and no flowers) with probabilities of fixation of .2 to .3 as shown in Figure 2a. The probabilities of fixation on the affirmative mentioned word (e.g., "roses") remained relatively constant throughout: they increased from 750 ms (*FDRcorr* = .0158); and then decreased from around 1,500 ms, with no significant differences from the baseline from 1,850 ms onwards (*FDRcorr* = .0546). But fixations on the negative phrase (e.g., "no roses") increased from 1,600 ms (*FDRcorr* = .0236) and remained elevated compared with the baseline thereafter. Fixations decreased on the distractors (affirmative-distractor, e.g., "flowers,"

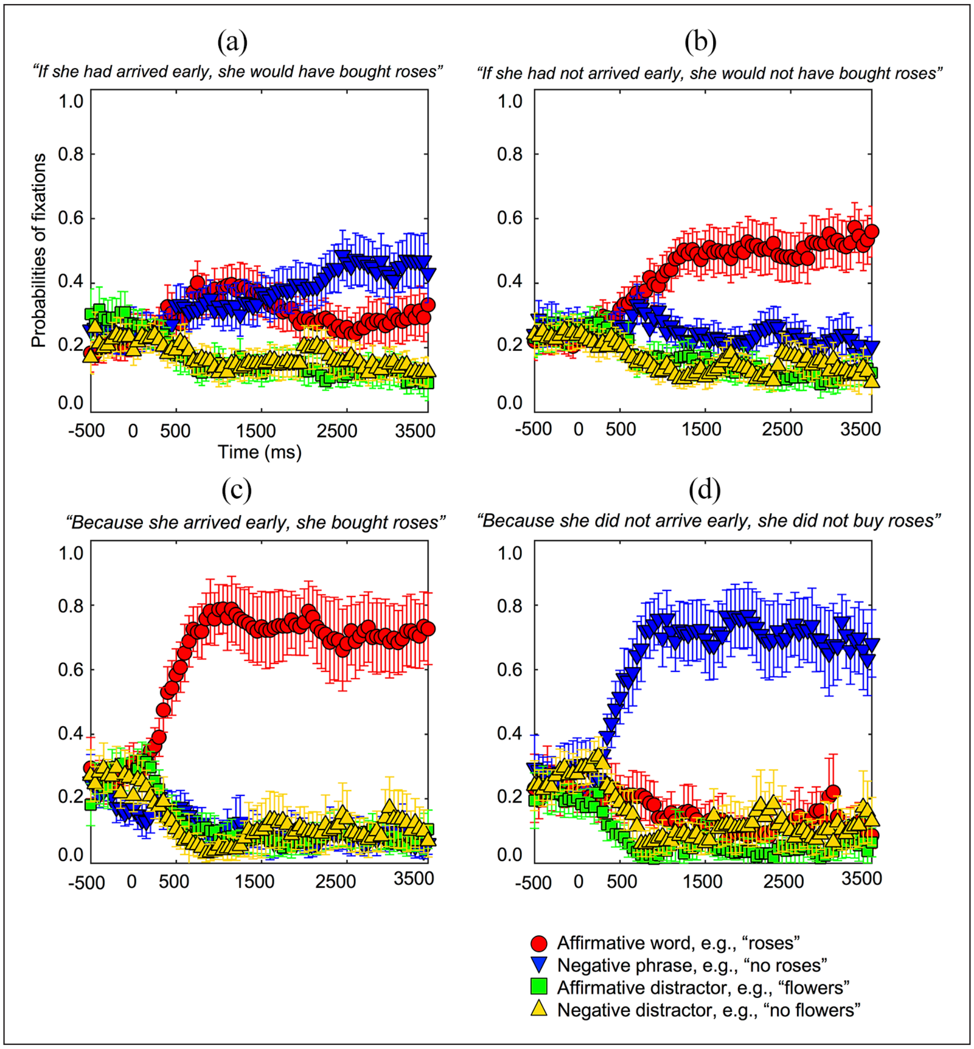

Figure 2. Probabilities of fixations for (a) affirmative counterfactuals, (b) negative counterfactuals, (c) affirmative causal assertions, and (d) negative causal assertions in Experiment 1 (without any context). Error bars are 95% confidence intervals within participants (see Morey, 2008; O'Brien & Cousineau, 2014).

from 450 ms, *FDRcorr* = .0342; negative-distractor, e.g., "no flowers," from 800 ms, *FDRcorr* = .0210). These results show that participants maintained a constant tendency to look at the word corresponding to the affirmative counterfactual's conjecture, e.g., "roses" at a steady rate from the outset; they showed an increasing tendency to look at the negative phrase corresponding to its presupposed fact, e.g., "no roses"; they did not look at the two distractors. The results replicate earlier findings for counterfactuals, with a printed words display (e.g., Orenes et al., 2019, Experiment 3).

To test whether there are differences between the affirmative mentioned word and the negative phrase for affirmative counterfactuals, we carried out a growth curve analysis between 2,000 and 3,500 ms, that is, the time period during which differences between the affirmative mentioned word and negative phrase could be observed, as shown in Figure 2a. There is a significant effect of the type of possibility, that is, the differences between the conjecture (e.g., "roses") and the presupposed facts (e.g., "no roses") on the intercept and the quadratic term. The largest effect is on the intercept term (*Estimate* = .162, *SE* = 0.043, $p < .001$) which means that people looked more at the words corresponding to the presupposed facts "no roses" than the conjecture "roses" (see Table 2).

Negative counterfactuals: For a negative counterfactual conditional, e.g., "If she had *not* arrived early, she would *not* have bought roses," participants focused on the four words on the screen at the outset with probabilities of fixation of .2 to .3 as shown in Figure 2b. The probabilities of fixation on the affirmative mentioned word (e.g., "roses") started to increase early on in the process, 300 ms after the mentioned word onset (*FDRcorr* = .0203) and remained elevated compared with the baseline thereafter. Fixations on the negative phrase (e.g., "no roses") remained relatively constant and showed no differences compared with

Table 2. Parameter estimates for analysis of effect of the type of possibility (the affirmative mentioned word versus the negative phrase) on the affirmative and negative counterfactuals.

	B	*SE*	*t*	*p*
Affirmative counterfactual				
Intercept	0.280	0.030	9.211	<.001
Linear	0.054	0.056	0.956	.339
Quadratic	0.070	0.038	1.836	.066
Cubic	–0.014	0.034	–0.421	.674
POSSIBILITIES	0.162	0.043	3.761	<.001
Linear: POSSIBILITIES	0.039	0.079	0.493	.622
Quadratic: POSSIBILITIES	–0.146	0.054	–2.720	.007
Cubic: POSSIBILITIES	0.071	0.048	1.489	.137
Negative counterfactual				
Intercept	0.444	0.021	20.874	<.001
Linear	0.309	0.054	5.701	<.001
Quadratic	–0.121	0.039	–3.096	.001
Cubic	–0.022	0.031	–0.728	.467
POSSIBILITIES	–0.192	0.030	–6.364	<.001
Linear: POSSIBILITIES	–0.466	0.077	–6.076	<.001
Quadratic: POSSIBILITIES	0.117	0.055	2.124	.034
Cubic: POSSIBILITIES	0.087	0.043	1.999	.046

SE: standard error.

Table 3. Parameter estimates for analysis of effect of the type of polarity (affirmative versus negative) on the presupposed facts for counterfactuals and causal assertions.

	b	*SE*	*t*	*p*
Counterfactuals				
Intercept	0.454	0.025	18.079	<.001
Linear	0.440	0.089	4.933	<.001
Quadratic	–0.259	0.063	–4.136	<.001
Cubic	0.076	0.047	1.602	.109
POLARITY	–0.083	0.026	–3.181	.001
Linear: POLARITY	–0.017	0.094	–0.176	.860
Quadratic: POLARITY	0.291	0.080	3.653	<.001
Cubic: POLARITY	–0.069	0.066	–1.055	.292
Causal assertions				
Intercept	0.705	0.041	16.966	<.001
Linear	0.195	0.102	1.909	.056
Quadratic	–0.379	0.070	–5.419	<.001
Cubic	0.290	0.058	4.962	<.001
POLARITY	–0.023	0.026	–0.912	.361
Linear: POLARITY	0.195	0.111	1.765	.077
Quadratic: POLARITY	–0.047	0.094	–0.501	.616
Cubic: POLARITY	–0.041	0.082	–0.490	.623

SE: standard error.

the baseline throughout. Fixations decreased on the distractor words (affirmative-distractor, e.g., "flowers," from 600 ms, *FDRcorr*=.0088; negative-distractor, e.g., "no flowers," from 700 ms, *FDRcorr*=.0238). The results show that participants maintained a constant tendency to look at the words corresponding to the negative counterfactual's conjecture, e.g., "no roses"; they showed an increasing tendency to look at the word corresponding to its presupposed fact, e.g., "roses"; they did not look at the two distractors. The findings are similar to the results for affirmative counterfactuals. Hence, they support the interpretation that the pattern of results for affirmative counterfactuals reflects a genuine preference for thinking about the *conjecture first* and then the *presupposed facts second.* They rule out the interpretation that participants merely have a preference for looking at affirmative words rather than negative phrases: for negative counterfactuals, participants maintained a constant tendency from the outset to look at the words corresponding to the conjecture (this time the negative phrase "no roses"); they showed an increasing tendency to look at the word corresponding to the presupposed facts (this time the affirmative word "roses").

To test whether there are significant differences between the affirmative mentioned word and the negative phrase for negative counterfactuals, we carried out a growth curve analysis between 500 and 2,000 ms, that is, the time period at which differences between the affirmative mentioned word and negative phrase could be observed as shown in Figure 2b. There is a significant effect of the type of possibility on all terms, which means that the differences between the conjecture and the presupposed facts are bigger for negative counterfactuals than affirmatives. The largest effect is on the intercept term (*Estimate*=−.192, *SE*=0.030, $p<.001$) which indicates that people looked more at the words corresponding to the presupposed facts "roses" than the conjecture "no roses" (see Table 2).

Finally, we carried out a growth curve analysis to compare the probability of fixations on the presupposed facts between affirmative and negative counterfactuals from 300 to 3,000 ms and they differed significantly in the intercept (*Estimate*=−.083, *SE*=0.026, $p<.001$), that is, the average height of the curve was higher for negative counterfactuals than affirmatives. There is a large effect on the quadratic term (*Estimate*=.291, *SE*=0.080, $p<.001$). This effect resulted from the fact that the rise is earlier for the negative counterfactuals (300 ms) than affirmative counterfactuals (1,600 ms; see Table 3). Hence, people looked at the words corresponding to the presupposed facts more for negative counterfactuals than affirmative ones, and at an earlier point.

Overall, for both affirmative and negative counterfactuals, the tendency to look at the presupposed facts (whether an affirmative word or negative phrase) occurred at a higher rate than the tendency to look at the conjecture, that is, the two possibilities were not equally activated. The differences between participants' constant maintenance of a tendency to look at the conjecture and their increased tendency to look at the presupposed facts are more pronounced for negative counterfactuals than for affirmative ones. It also occurs earlier: participants showed a tendency to increasingly look at the presupposed facts for negative counterfactuals (the affirmative word, e.g., "roses," 300 ms) earlier than they showed a tendency to increasingly look at the presupposed facts for affirmative counterfactuals (the negative phrase, e.g., "no roses," 1,600 ms).

Affirmative and negative causal assertions: In Experiment 1b, for affirmative causal assertions, e.g., "Because she arrived early, she bought roses," participants focused on the four words or phrases on the screen (e.g., "roses," "no roses," "flowers," and "no flowers") at the outset with probabilities of fixation of .2 to .3 as shown in Figure 2c. The probabilities of fixation on the affirmative mentioned word (e.g., "roses") started to increase very early from 400 ms (*FDRcorr* = .0300). Fixations decreased on the negative phrase (e.g., "no roses") from 1,400 ms (*FDRcorr* = .0333), and fixations also decreased on the distractors (affirmative distractor, e.g., "flowers," from 500 ms, *FDRcorr* = .0255; negative distractor, e.g., "no flowers," from 500 ms, *FDRcorr* = .0307). Hence, participants showed an increasing tendency to look at the word corresponding to what was mentioned in the affirmative causal assertion, e.g., "roses," from very early in the process; and they did not tend to look at the other three words.

Likewise, for negative causal assertions, e.g., "Because she did *not* arrive early, she did *not* buy roses," participants focused on the four words or phrases on the screen with probabilities of fixation of .2 to .3 as shown in Figure 2d. The probabilities of fixation on the negative phrase (e.g., "no roses") increased from 600 ms (*FDRcorr* = .0130). Fixations on the affirmative mentioned word (e.g., "roses") started to decrease from 1,550 ms (*FDRcorr* = .0244), and fixations also decreased on the distractor words (affirmative distractor, e.g., "flowers," from 450 ms, *FDRcorr* = .0091; negative distractor, e.g., "no flowers," from 650 ms, *FDRcorr* = .0295). Hence, participants showed an increasing tendency to look at the words corresponding to what was mentioned in the negative causal assertion, e.g., "no roses," from very early in the process, and they did not tend to look at the other three words.

Finally, we carried out a growth curve analysis to compare the probability of fixations on the words corresponding to the facts between affirmative causal assertions (e.g., "roses") and negative causal assertions (e.g., "no roses") from 300 to 3,000 ms, and although people looked increasingly at "no roses" for negative causal assertions (600 ms) later than at "roses" for affirmative causal assertions (400 ms) the difference is not significant (see Table 3). However, negative causal assertions were slower to process than affirmative ones in the latency data from the behavioural results, to which we now turn.

Behavioural data. In Experiment 1a, participants made an inference about whether the final affirmative assertion, e.g., "Carmen bought roses" or final negative assertion, e.g., "Carmen did not buy roses" was consistent or inconsistent with the previous assertion. Response accuracy and latency were both analysed separately using two analysis of variance (ANOVA) tests with the design of 2 (polarity of counterfactual: affirmative or negative) × 2 (polarity of final assertion: affirmative or negative) repeated measures. Response accuracy showed no main effect of polarity of the counterfactual, $F(1,44) = 0.538$, $p = .467$, or polarity of the final assertion, $F(1,44) = 2.037$, $p = .161$, and no interaction, $F(1,44) = 0.800$, $p = .376$. Latency showed a main effect of polarity of the final assertion, $F(1,44) = 52.343$, $p < .001$, $\eta_p^2 = .543$, no main effect of polarity of the counterfactual, $F(1,44) = 0.640$, $p = .428$, and an interaction of the two, $F(1,44) = 21.393$, $p < .001$, $\eta_p^2 = .327$. The decomposition of the interaction showed there were no differences between affirmative and negative final assertions after affirmative counterfactuals, $M = 3{,}569$, $SD = 697$ vs $M = 3{,}773$, $SD = 897$, $t(44) = -1.781$, $p = .082$, a result consistent with previous findings that both possibilities, the conjecture and the presupposed facts, are equally accessible (Santamaría et al., 2005). But the affirmative final assertions were responded to more quickly than the negative ones after negative counterfactuals, $M = 3{,}252$, $SD = 790$ vs $M = 4{,}253$, $SD = 1{,}200$; $t(44) = -8.005$, $p < .001$, $d = -1.193$. This result is inconsistent with the idea that both possibilities are equally accessible. However, it may arise from an artefact of the behavioural task. A negative counterfactual, "If she had not arrived early, she would not have bought roses," followed by a negative assertion "Carmen did not buy roses" requires a negative response "no": processing the triple negation may slow participants' responses.

In Experiment 1b, response accuracy showed no main effect of polarity of the causal assertion, $F(1,25) = 1.000$, $p = .327$, or polarity of the final assertion, $F(1,25) = 0.074$, $p = .788$, and no interaction, $F(1,25) = 0.107$, $p = .746$. Latency showed a main effect of polarity of the causal assertion, $F(1,25) = 6.569$, $p = .017$, $\eta_p^2 = .208$, and of polarity of the final assertion, $F(1,25) = 43.975$, $p < .001$, $\eta_p^2 = .638$. The interaction was not significant, $F(1,25) = 3.022$, $p = .094$. Affirmative final assertions were responded to more quickly than negative ones after affirmative causal assertions, $M = 1{,}347$, $SD = 316$ vs $M = 1{,}556$, $SD = 376$, $t(25) = -4.605$, $p < .001$, $d = -0.903$, and after negative causal assertions, $M = 1{,}481$, $SD = 307$ vs $M = 1{,}581$, $SD = 331$, $t(25) = -3.188$, $p = .004$, $d = -0.625$. Overall, the analyses of the latency measures show the standard effect of negation for causal assertions, whereas no such effect is observed for counterfactuals.

In summary, the results of Experiments 1a and 1b show that for causal assertions people look at a single word or phrase: for affirmative causal assertions, they look at the affirmative word (e.g., "roses"), and for negative causal assertions they look at the negative phrase (e.g., "no roses"). But for counterfactual conditionals, the pattern is different. For affirmative counterfactuals, they look at the affirmative mentioned word (e.g., "roses," the conjecture) from the outset and constantly throughout the duration of processing, and they increasingly look at the negative phrase (e.g., "no roses," the presupposed facts). Likewise, for negative counterfactuals, they look at the words that correspond to the conjecture (this time the negative phrase, e.g., "no roses") constantly from the outset, and they increasingly look at the word that corresponds to the presupposed facts (this time the affirmative word, e.g., "roses"). The results indicate that negative counterfactuals are interpreted in a similar way to affirmative ones: participants maintain from the outset a steady tendency to look at the words corresponding to the conjecture in both cases regardless of whether it is an affirmative word or a negative phrase. They show an increased tendency to look at the words corresponding to the presupposed facts throughout the measured duration, again regardless of whether it is an affirmative word or a negative phrase. Their tendency to do so is more pronounced for negative counterfactuals than affirmative ones, as shown in Figure 2.

One concern is that for the affirmative counterfactual, the word corresponding to its conjecture is affirmative (e.g., "roses") whereas the words corresponding to its presupposed facts is a negative phrase (e.g., "no roses"). The opposite is the case for the negative counterfactual—the words corresponding to its conjecture are a negative phrase (e.g., "no roses"), and the word corresponding to its presupposed facts is affirmative (e.g., "roses"). Hence, the tendency to look at the presupposed facts may be more pronounced for a negative counterfactual than for an affirmative counterfactual merely because of this difference: it reflects a difference in looking at an affirmative word (for the negative counterfactual) and at a negative phrase (for the affirmative counterfactual). The next experiment removes this potential confound in the materials by presenting the same form of counterfactuals (affirmative and negative counterfactuals) in a binary context, e.g., "she did not know whether to buy roses or carnations."

Experiment 2: affirmative and negative counterfactual and causal assertions in a binary context

The aim was to examine how people understand a counterfactual such as "If she had arrived early, she would have bought roses" in a binary context, such as "she did not know whether to buy roses or carnations." Participants saw four printed words or phrases on the screen corresponding to the mentioned word and its negation: "roses," "no roses," and to the alternate word and its negation, "carnations," "no carnations." For the affirmative counterfactual, the binary context provides the option for participants to look at an affirmative mentioned word (e.g., "roses") corresponding to its conjecture and to look at an affirmative alternate word (e.g., "carnations") corresponding to its presupposed facts rather than a negative phrase (e.g., "no roses"). For the negative counterfactual, the binary context provides the option for participants to look at an affirmative alternate word (e.g., "carnations") corresponding to its conjecture rather than a negative phrase (e.g., "no roses"), and to look at an affirmative mentioned word (e.g., "roses") corresponding to its presupposed facts.

We made the following set of predictions:

1. *Affirmative counterfactuals*: For an affirmative counterfactual, we expect that participants will look at the word corresponding to the conjecture mentioned in the counterfactual, e.g., "roses"; but this time in the binary context, the word they will look at corresponding to the presupposed facts will be the alternate "carnations," rather than the negative phrase "no roses." We also expect that although they will look at the word corresponding to the conjecture, e.g., "roses," this time their tendency to look at the word corresponding to the presupposed facts, e.g., "carnations," will increase throughout the processing duration, in as pronounced a manner as the tendency observed for negative counterfactuals in the previous experiment, given that the word corresponding to the presupposed facts "carnations" is affirmative. Of course, the tendency to do so will occur during a later period of time given the extra cognitive steps to infer that "no roses" means "carnations" as described earlier.
2. *Negative counterfactuals.* For a negative counterfactual such as "if she had *not* arrived early, she would *not* have bought roses," we predict that this time given the binary context, the word participants will look at corresponding to the conjecture will be the alternate, "carnations"; and they will look at the word corresponding to the presupposed facts, "roses." We also predict that although they will look at the word corresponding to the conjecture "carnations," their tendency to look at the word corresponding to the presupposed facts, e.g., "roses" will increase throughout the processing duration in a pronounced manner, replicating the previous experiment.
3. *Affirmative and negative causal assertions.* For an affirmative causal assertion, such as "because she arrived early, she bought roses" in a binary context, "she did not know whether to buy roses or carnations," we expect that participants will look at the affirmative mentioned word, e.g., "roses," and they will not look at any of the other three words or phrases ("no roses," "carnations," "no carnations"). For a negative causal assertion such as "because she did not arrive early, she did not buy roses" in this binary context, we expect that they will look at the alternate "carnations" rather than the negative phrase, "no roses," and they will not look at any of

the other three words or phrases. In the previous experiment, participants looked at the words corresponding to the facts somewhat more quickly for the affirmative causal assertion ("roses") than for the negative causal assertion ("no roses"), and although the difference was not significant, one interpretation is that it reflects merely the added difficulty of processing the negative phrase "no roses" compared with the affirmative word "roses." If so, then the difference should be eliminated in this binary context for "carnations" and "roses." An alternative interpretation is that there is an additional conceptual difficulty in understanding a negative causal assertion compared with an affirmative one. If so, then participants will look at the word corresponding to the facts more rapidly for affirmative causals ("roses") than for negative causals ("carnations"), even when the words are both affirmative.

Method

Participants. The participants were 32 volunteers (28 women and 4 men) whose average age was 21 years, with a range from 19 to 44 years. They participated in exchange for course credits and none of them had taken part in the previous experiments. A sample size of 32 participants provides 87% power at alpha .05 to detect a medium-sized effect for a one-tailed test (78% for a two-tailed test).

Materials, design, and procedure. The same materials, design, and procedure were used as in the previous experiments with the following three exceptions:

1. We added a binary context at the start of each vignette, e.g., "Carmen liked to buy flowers in the market to decorate her house. She did not know whether to buy roses or carnations" ("A Carmen le gustaba comprar flores en el mercado para decorar su casa. No sabía si comprar rosas o claveles").
2. Participants heard the counterfactuals or causal assertions in the same experiment.
3. Participants all read the same final assertion, e.g., "Carmen also bought a vase for the roses" ("Carmen también compró un jarrón para las rosas").

Their task was the same as the previous experiment, to judge whether the final sentence was consistent with the earlier ones. Half of the final assertions were consistent with the earlier sentences and the other half were not. Participants saw on screen four words or phrases, the affirmative mentioned word "roses," its explicit negation, "no roses," the alternate mentioned in the binary context, "carnations," and its negation, "no carnations." The design included two types of assertion: causal or counterfactual, and two types of polarity of those assertions, affirmative or negative, in a within-participant design. We constructed eight versions of each vignette, four conditions: causal-affirmative, causal-negative, counterfactual-affirmative, and counterfactual-negative, and two objects (e.g., roses or carnations). We constructed eight sets and each participant was assigned one set in a Latin-square design to ensure that each participant received only one of the eight possible versions for each content. There were 36 vignettes, 9 trials in each of the 4 conditions (9 counterfactual-affirmative, 9 counterfactual-negative, 9 causal-affirmative, and 9 causal-negative), presented in a randomised order.

Results and discussion

Prior to any data analysis, one participant was eliminated from the analysis because his eye movements explored the screen continuously without any systematic fixations on any point.

T-*tests against the baseline of 0–100 ms and growth curve analysis*

Affirmative counterfactuals. For an affirmative counterfactual, e.g., "If she had arrived early, she would have bought roses" in a binary context "she did not know whether to buy roses or carnations," before the onset of the target word, i.e., the last word in the conditional (e.g., "roses"), participants focused on the two possibilities mentioned in the binary context (e.g., "roses" and "carnations"), with probabilities of fixation of .3 to .4 as shown in Figure 3a. The probabilities of fixation on the affirmative mentioned word (e.g., "roses") started to increase from 450 ms (FDRcorr = .012), but then decreased from 2,350 ms (FDRcorr = .029). For the affirmative alternate word (e.g., "carnations"), fixations first decreased from 450 ms (FDRcorr = .029), but then increased from 2,600 ms (FDRcorr = .027). For the negative phrases ("no roses," "no carnations") fixations were low and remained unchanged. The results show that participants looked at the affirmative alternate word ("carnations") rather than at the negative phrase ("no roses") in this binary context; the result extends earlier findings from inferential studies that the presupposed facts are represented by an alternate in a binary context (Espino & Byrne, 2018). The results also show that in the binary context, participants tended to look initially at the affirmative mentioned word corresponding to the conjecture (e.g., "roses"), but their tendency to look at it diminished and was replaced instead by a tendency to look at the affirmative alternate word corresponding to the presupposed facts ("carnations"). The finding indicates that for an affirmative counterfactual in a binary context, participants look at the conjecture ("roses") only during the earliest period of measurement, and then they look instead at the presupposed facts ("carnations") during the remaining period. This fixation pattern is different from the pattern observed in Experiment 1, in which fixations on the conjecture were maintained throughout the duration of measurement. It suggests that in a binary context (e.g., "roses or carnations"), participants focus on the possibility corresponding to the presupposed facts ("carnations") rather than the one corresponding to the conjecture ("roses") for an affirmative counterfactual.

To test whether there are differences between the affirmative mentioned word and the alternate word for affirmative counterfactuals, we carried out a growth curve analysis between 300 and 3,000 ms, which is the time period during which differences between the affirmative mentioned word and the alternate word could be observed in Figure 3a. Table 4 shows that there is an effect of the type of possibility on the slope (*Estimate* = 1.464, *SE* = 0.161, $p < .001$), indicating a steeper slope for the affirmative mentioned word than the alternate. This result means that the increase for the affirmative mentioned word was faster (e.g., "roses" which corresponds to the conjecture; 450 ms) than the alternate word (e.g., "carnations" which corresponds to the presupposed facts; 2,600 ms).

Negative counterfactuals. For a negative counterfactual, e.g., "If she had *not* arrived early, she would *not* have bought roses" in a binary context "she did not know whether to buy roses or carnations," before the onset of the target word in the conditional, e.g., "roses," participants focused on the possibilities mentioned in the binary context ("roses" and "carnations") equally frequently with probabilities of fixations of about .4, as shown in Figure 3b. They increased their fixations on the affirmative mentioned word ("roses") from 800 ms after the target word onset (FDRcorr = .030). They decreased their fixations on the affirmative alternate word ("carnations") from 650 ms (FDRcorr = .041). Fixations were low and remained unchanged for the two negative phrases ("no roses" and "no carnations"). The results show that for a negative counterfactual in a binary context, participants tended to look at the word corresponding to the presupposed facts (e.g., "roses"), and their tendency to look at the word corresponding to the conjecture ("carnations") rapidly diminished.

To test whether there are significant differences between the affirmative mentioned word and the alternate word for negative counterfactuals, we carried out a growth curve analysis between 300 and 3,000 ms, which is the time period during which differences between the affirmative mentioned word and the alternate word could be observed in Figure 3b. There is a significant effect of the type of possibility on all terms, which means that the differences between the conjecture and the presupposed facts are bigger for negative counterfactuals than affirmatives. This finding replicates Experiment 1. The largest effect is on the intercept term (*Estimate* = −0.494, *SE* = 0.044, $p < .001$), which means that people looked more at the presupposed fact "roses" than the conjecture "carnations" (see Table 4).

Finally, we carried out a growth curve analysis to compare the probability of fixations on the presupposed facts between affirmative counterfactuals (for which the presupposed facts corresponds to the alternate word) and negative counterfactuals (for which the presupposed facts corresponds to the affirmative mentioned word) from 300 to 3,000 ms. They differ significantly in the intercept (*Estimate* = 0.334, *SE* = 0.046, $p < .001$), that is, the average height of the curve was higher for negative counterfactuals than affirmatives. There was also an effect of polarity on the slope (*Estimate* = −0.389, *SE* = 0.151, $p = .009$), indicating a steeper slope for the negative counterfactual (e.g., "roses," 800 ms) than for an affirmative counterfactual (e.g., "carnations," 2,600 ms); and the quadratic (*Estimate* = −0.256, *SE* = 0.084, $p = .002$), indicating that the affirmative and negative counterfactuals differed in the processing of the presupposed fact (see Table 5). In sum, people looked more at the presupposed fact—and at an earlier time—for negative counterfactuals than affirmative ones in the binary context too.

We can rule out an alternative explanation of the results, that they arise because the first step in constructing a representation of a counterfactual requires a participant to focus their attention on different words or phrases at three distinct stages: (a) the mentioned word or phrase, (b) the word or phrase corresponding to the conjecture, and (c) the word or phrase corresponding to the presupposed facts. For the affirmative counterfactual, these stages correspond to: (a) "roses" (the mentioned word), (b) "roses" (the word corresponding to the conjecture), and (c) "no roses"/"carnations" (the negative phrase or alternate corresponding to the presupposed facts), whereas for the negative counterfactual, they correspond to: (a) "roses" (the mentioned word), (b) "no roses"/"carnations" (the negative phrase or alternate corresponding to the conjecture), and (c) "roses" (the word corresponding to the presupposed facts). On this explanation, the downward datapoint for "roses" around 2,000 ms in Figure 3b corresponds to the "middle step" of this process. This alternative explanation may appear attractive, because it is surprising that people understand negative counterfactuals so early compared with affirmative counterfactuals, given previous findings that negation is generally slower to process. However, it is important to recall that such a downward datapoint did not occur for negative counterfactuals in Experiment 1 (see Figure 2b). Moreover, it is noteworthy that people looked directly at the negative phrase "no roses" for negative causal assertions in Experiment 1 (see Figure 2d), and so it seems that negation is indeed integrated immediately in that case.

Affirmative and negative causal assertions. For affirmative causal assertions, e.g., "Because she arrived early,

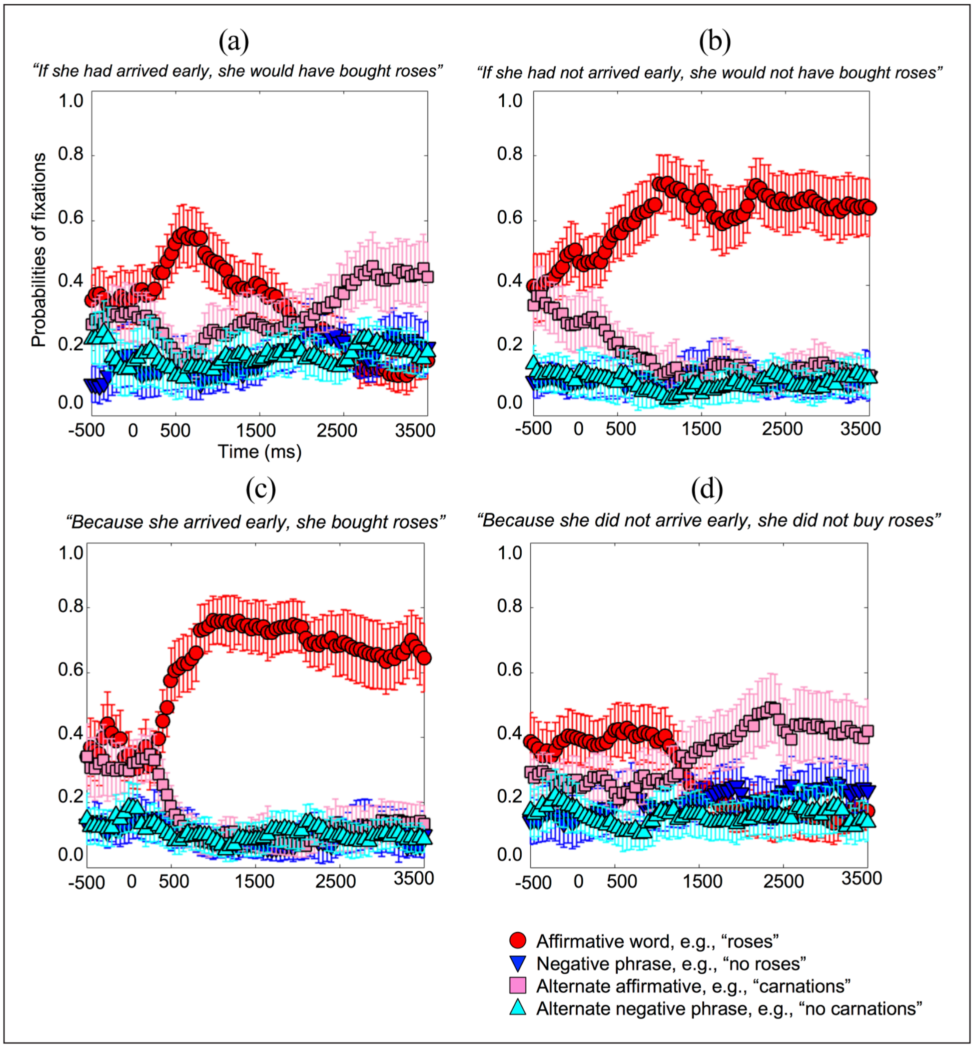

Figure 3. Probabilities of fixations on the affirmative mentioned word (e.g., "roses"), negative phrase (e.g., "no roses"), affirmative alternate word (e.g., "carnations"), and alternate negative phrase (e.g., "no carnations") for (a) affirmative counterfactuals, (b) negative counterfactuals, (c) affirmative causal assertions, and (d) negative causal assertions, presented in a binary context (e.g., "roses or carnations") in Experiment 2, time-locked to the onset of the first object word, e.g., "roses."
Error bars are 95% confidence intervals within participants (see Morey, 2008; O'Brien & Cousineau, 2014).

she bought roses" in a binary context "She did not know whether to buy roses or carnations," before the onset of the target word in the causal assertion (e.g., "roses"), participants focused on the two possibilities mentioned in the binary context ("roses" and "carnations"), with probabilities of fixation of .3 to .4 as shown in Figure 3c. The probabilities of fixation on the affirmative mentioned word (e.g., "roses") started to increase early in the process, 450 ms after the target word onset (FDRcorr = .009). Fixations decreased on the affirmative alternate word (e.g., "carnations") from 500 ms (FDRcorr = .015). Fixations remained low and with no change for the negative phrases ("no roses" and "no carnations"). The results show that people look at the affirmative mentioned word (e.g., "roses") for an affirmative causal assertion, as shown in Figure 3c.

For negative causal assertions, e.g., "Because she did *not* arrive early, she did *not* buy roses" in a binary context, before the onset of the target word, e.g., "roses," participants focused on the two possibilities mentioned in the binary context ("roses" and "carnations") equally frequently with probabilities of fixations of .3 to .4, as shown in Figure 3d. Participants decreased their fixations on the affirmative mentioned word (e.g., "roses") from 1,450 ms after the target word onset (FDRcorr = .010) and increased their fixations on the affirmative alternate word (e.g., "carnations") from 1,750 ms (FDRcorr = .006). Fixations on the negative phrases remained low with no changes ("no roses" and "no carnations"). The results show that people look at the affirmative alternate word (e.g., "carnations") for a negative causal assertion, as shown in Figure 3d. It is noteworthy that people look at the affirmative alternate

Table 4. Parameter estimates for analysis of effect of the type of possibility (the affirmative mentioned word versus the alternate word) on the affirmative and negative counterfactuals.

	b	*SE*	*t*	*p*
Affirmative counterfactual				
Intercept	0.347	0.028	12.325	<.001
Linear	–0.930	0.114	–8.174	<.001
Quadratic	–0.059	0.089	–0.653	.513
Cubic	0.081	0.069	1.180	.238
POSSIBILITIES	–0.049	0.040	–1.237	.216
Linear: POSSIBILITIES	1.464	0.161	9.102	<.001
Quadratic: POSSIBILITIES	0.167	0.127	1.315	.188
Cubic: POSSIBILITIES	–0.044	0.097	–0.449	.653
Negative counterfactual				
Intercept	0.642	0.031	20.800	<.001
Linear	0.144	0.082	1.759	.078
Quadratic	–0.148	0.074	–1.991	.046
Cubic	0.142	0.061	2.334	.019
POSSIBILITIES	–0.494	0.044	–11.317	<.001
Linear: POSSIBILITIES	–0.282	0.116	–2.432	.015
Quadratic: POSSIBILITIES	0.323	0.105	3.086	.002
Cubic: POSSIBILITIES	–0.195	0.086	–2.268	.023

SE: standard error.

Table 5. Parameter estimates for analysis of effect of the type of polarity (affirmative versus negative) on the presupposed facts for counterfactuals and causal assertions.

	b	*SE*	*t*	*p*
Counterfactuals				
Intercept	0.297	0.035	8.437	<.001
Linear	0.534	0.111	4.786	<.001
Quadratic	0.108	0.074	1.470	.141
Cubic	0.037	0.062	0.595	.551
POLARITY	0.344	0.046	7.491	<.001
Linear: POLARITY	–0.389	0.151	–2.589	.009
Quadratic: POLARITY	–0.256	0.084	–3.044	.002
Cubic: POLARITY	0.104	0.084	1.246	.212
Causal assertions				
Intercept	0.688	0.034	20.009	<.001
Linear	0.153	0.118	1.298	.194
Quadratic	–0.438	0.095	–4.573	<.001
Cubic	0.243	0.067	3.631	<.001
POLARITY	–0.329	0.041	–8.078	<.001
Linear: POLARITY	0.411	0.153	2.687	.007
Quadratic: POLARITY	0.331	0.126	2.622	.008
Cubic: POLARITY	–0.415	0.091	–4.611	<.001

SE: standard error.

(e.g., "carnations") rather than the negative phrase (e.g., "no roses") for a negative causal assertion, although they looked at it late in the measurement period (1,750 ms). This late processing contrasts with the much earlier processing of the negative phrase, e.g., "no roses," in Experiment 1b (600 ms) for negative causal assertions, again perhaps reflecting the additional cognitive steps required to infer that, e.g., "no roses" means, e.g., "carnations," as described earlier.

Finally, we carried out a growth curve analysis to compare the probability of fixations on the facts between affirmative causal assertions (e.g., "roses") and negative causal assertions (e.g., "carnations") from 300 to 3,000 ms and they differ significantly in the intercept (*Estimate* = −0.329, *SE* = 0.041, $p < .001$), that is, the average height of the curve was higher for affirmative causal assertions than negatives. There was also an effect of polarity on the slope (*Estimate* = 0.411, *SE* = 0.153, $p = .007$), indicating a steeper slope for the affirmative causal assertions (e.g., "roses," 450 ms) than for negatives (e.g., "carnations," 1,750 ms); and the quadratic (*Estimate* = 0.331, *SE* = 0.126, $p = .008$), indicating that the affirmative and negative causal assertions differed in the processing of the word corresponding to the fact (see Table 5). The results are consistent with the well-known effects of negation, that is, people need more time to process negative sentences than affirmative sentences, as their search for an alternate takes some time. The behavioural results for response accuracy and latency to respond to the comprehension questions, to which we now turn, are consistent with the eye-tracking data.

Behavioural data. Response accuracy and latency were both analysed separately using two 2 (sentence: causal or counterfactual) × 2 (polarity: affirmative or negative) repeated-measures ANOVAs. Response accuracy showed a main effect of sentence, $F(1,30) = 4.258$, $p = .048$, $\eta_p^2 = .124$, a main effect of polarity, $F(1,30) = 6.378$, $p = .017$, $\eta_p^2 = .175$, and an interaction, $F(1,30) = 4.235$, $p = .048$, $\eta_p^2 = .124$. Accuracy was higher after affirmative causal ($M = 8.81$, $SD = 0.40$) than negative causal assertions, $M = 8.06$, $SD = 0.96$; $t(30) = 4.004$, $p < .001$, $d = 0.719$, but there was no difference after affirmative counterfactuals ($M = 8.19$, $SD = 0.96$) compared with negative counterfactuals, $M = 8.13$, $SD = 0.99$; $t(30) = 0.24$, $p = .810$. Latency showed an interaction, $F(1,30) = 4.765$, $p = .037$, $\eta_p^2 = .137$, with no main effects of sentence, $F(1,30) = 2.144$, $p = .153$, or polarity, $F(1,30) = 0.083$, $p = .775$. Latency was faster after affirmative causal ($M = 2{,}169.88$, $SD = 460.12$) than negative causal assertions, $M = 2{,}331.96$, $SD = 498.12$; $t(30) = -2.548$, $p = .016$, $d = -0.458$, but there was no difference between affirmative counterfactuals ($M = 2{,}394.68$, $SD = 655.69$) and negative counterfactuals, $M = 2{,}267.59$, $SD = 639.99$; $t(30) = 1.155$, $p = .257$. Again, the behavioural analyses show the standard effect of negation for causal assertions, whereas this effect is removed for counterfactuals.

General discussion

There are notable similarities and differences in how people understand a negative counterfactual and an affirmative one, in the immediate few seconds after they first hear it. First, there are similarities between affirmative and negative counterfactuals. When participants heard an affirmative counterfactual, such as "if she had arrived early, she would have bought roses," they looked at the word corresponding to the conjecture ("roses") from the outset, and they then also looked at the negative phrase corresponding to its presupposed fact ("no roses"), as shown in Experiment 1a. When they were also given a binary context, such as "she did not know whether to buy roses or carnations," they looked at the word corresponding to the conjecture ("roses") initially, but then they looked instead at the word corresponding to the presupposed facts ("carnations"), as shown in Experiment 2. The results are consistent with previous evidence that participants understand an affirmative counterfactual by thinking about two possibilities (e.g., Orenes et al., 2019) and that they represent the presupposed facts by an alternate in a binary context (e.g., Espino & Byrne, 2018).

Similarly, when participants heard a negative counterfactual such as "if she had not arrived early, she would not have bought roses," they did something quite comparable. They maintained a steady tendency to look at the negative phrase corresponding to its conjecture ("no roses"), and their tendency to look at the word corresponding to its presupposed facts ("roses") increased in the first few seconds after hearing it, as shown in Experiment 1a. And in a binary context, such as "she did not know whether to buy roses or carnations," they tended to look only at the word corresponding to the presupposed facts ("roses"), and they did not tend to look at the word corresponding to the conjecture ("carnations"), as shown in Experiment 2. The implication of these similarities is that they support the idea that people have a genuine preference for thinking about the *conjecture first* and then the *presupposed facts second*, for counterfactuals, whether affirmative or negative. They rule out the suggestion that the tendency, observed originally for affirmative counterfactuals, arises merely because of differences in preferences for looking at affirmative words compared with negative phrases.

Second, there are differences between affirmative and negative counterfactuals. For both sorts of counterfactuals, people tended to look at the presupposed facts more than the conjecture, and the two possibilities were not equally activated. But the effect was far more pronounced for negative counterfactuals than affirmative ones. Participants looked at the words corresponding to the presupposed facts *more* for negative counterfactuals than affirmative ones, and at an *earlier* time, in both experiments. Negative counterfactuals allow people to make a well-specified presumption about the facts, e.g., "in fact she bought roses," which identifies a specific item "roses"; affirmative counterfactuals allow people to make a less well-specified presumption about the facts, e.g., "in fact she did not buy roses," which does not distinguish whether she bought carnations, lilies, tulips, and so on. As a result, given the constraints of working memory, participants may tend to envisage just a single possibility—the presupposed facts for a negative counterfactual, even more so than for an affirmative one, in the immediate few seconds after hearing it.

Hence, although participants consider two possibilities for counterfactuals, their focus on the conjecture diminishes after some time (de Vega et al., 2007; de Vega & Urrutia, 2011, 2012; Ferguson & Jayes, 2018; Urrutia et al., 2012). Of course, the evidence relates to what participants focus on in just the first 2 or 3 s after they have heard a counterfactual. Given task demands, such as to make an inference, they can bring both possibilities back into the focus of attention, and hence make similar inferences from affirmative and negative counterfactuals (Thompson & Byrne, 2002). Indeed, our behavioural results for judgements of the consistency of the final sentence, which have similar task demands to inference tasks, also confirm that people make similar judgements about affirmative and negative counterfactuals, notwithstanding the initial differences in their immediate comprehension revealed by our eye-tracking results.

The experiments show that negative counterfactuals do not require longer processing times than affirmative ones. The increase in the tendency to look at the word corresponding to the presupposed fact occurred more quickly for negative counterfactuals than affirmative ones. In contrast, causal assertions show the standard pattern for negation: participants increased their tendency to look at the affirmative word for affirmative causal assertions more than the negative words for negative causal assertions. This finding suggests that the difficulty of negation does not depend merely on the syntactic operator "not" but instead is related to the conceptual meaning of negation (Miller, 1962).

These two ways to represent negation, by an explicit negation ("no roses") or an alternate ("carnations"), have been extensively studied (see Espino & Byrne, 2018; Khemlani et al., 2012; Orenes et al., 2014; Wason, 1959, 1961). Our results are consistent with theories that admit at least two types of representations: iconic and symbolic (e.g., Johnson-Laird & Byrne, 2002; Khemlani et al., 2012, 2014; Orenes et al., 2014). People can construct a mental representation of a negative assertion such as "there were no roses," either by envisaging "roses" and including a symbolic annotation, such as a propositional-like tag to capture negation explicitly, e.g., "no roses," or they may envisage an alternate, e.g., "carnations." Our results may pose some problems for theories based solely on experiential simulation, which assume that negation can only be represented by an alternate (see Kaup & Zwaan, 2003; Kaup et al., 2006).

One issue is whether the advantage for negative counterfactuals compared with affirmative counterfactuals reflects the importance of "pragmatic felicity," that is, that negative sentences are easier to process when they are presented in a context (Glenberg et al., 1999; Nieuwland & Kuperberg, 2008; Orenes et al., 2015; Wason, 1965; Xiang et al., 2020). The slower processing of negation is eliminated for cleft assertions such as "It was Miguel who did not iron the shirt," and concessive sentences, such as, "although" (Lyu et al., 2020; Tian et al., 2010; see also Johnson-Laird & Tridgell, 1972). Participants can process sentences such as "with proper equipment, scuba-diving is not very dangerous and often good fun" more quickly than sentences such as "Bulletproof vests are not very dangerous and used worldwide for security" because of the pragmatic felicity of the context (Nieuwland & Kuperberg, 2008), or its effects on predictability (Darley et al., 2020; Nieuwland, 2016). In fact, the counterfactual can be considered a context that does not merely eliminate the effects of negation but can reverse them. An affirmative counterfactual such as "if she had arrived early, she would have bought roses" can be considered a sort of implicit negation, in so far as it mentions an affirmative conjecture (she arrived early and bought roses) but communicates implicitly the opposite of what it mentions, the negative presupposed facts (she did not arrive early and did not buy roses). In that case, a negative counterfactual such as "if she had not arrived early, she would not have bought roses" can be considered a sort of double negation, in so far as it mentions a negative conjecture (she did not arrive early and did not buy roses) but communicates implicitly the opposite of what it mentions, the affirmative presupposed facts (she arrived early and bought roses).

The third finding of note is that there are differences between counterfactuals and causal assertions. What people look at when they hear affirmative and negative counterfactuals is different from what they look at when they hear affirmative causal assertions such as "because she arrived early, she bought roses" and negative causal assertions such as "because she did not arrive early, she did not buy roses." For causal assertions people tend to look at only one possibility, the affirmative mentioned word for affirmative causal assertions ("roses") and the negative phrase for negative causal assertions ("no roses") as shown in Experiment 1b. They tend to look at the affirmative alternative for negative causal assertions ("carnations") in a binary context, as shown in Experiment 2. Nonetheless, a glance at Figure 3 reveals a striking comparison: the pattern of fixations for an affirmative counterfactual "if she had arrived early, she would have bought roses" in a binary context is similar to that for a negative causal, "because she did not arrive early, she did not buy roses"—in both cases participants tend to look at the affirmative mentioned word "roses" in the early period of processing, and then they tend to look at the affirmative alternate "carnations" in the later period. The pattern for a negative counterfactual, "if she had not arrived early, she would not have bought roses" is similar to that for an affirmative causal, "because she arrived early, she bought roses"—participants tend to look at the affirmative mentioned word "roses" throughout the period of processing. However, although there are close links between counterfactual and causal assertions, a glance at Figure 2 shows that in non-binary contexts, it is clear that counterfactuals make different sorts of possibilities most readily accessible compared with causal assertions.

In summary, we have examined two cases for counterfactuals in which people focus readily on the presupposed facts as well as the conjecture. One case concerns negative counterfactuals. People can focus readily on the presupposed fact when it refers to a specific instance, as is the case for a negative counterfactual, such as, "If she had not arrived early, she would not have bought roses" in which the presupposed fact is that "she bought roses." The presupposed fact can be represented by a single instance, roses. In contrast in the case of affirmative counterfactuals, such as, "If she had arrived early, she would have bought roses" the presupposed fact is "she did not buy roses" which cannot be represented by a specific instance because it can refer to a large set of instances (carnations, lilies, tulips, etc.). A second case is the case of binary contexts. In a binary context, such as "she did not know whether to buy roses or carnations," the presupposed facts for an affirmative counterfactual, just as for a negative counterfactual, have a specific instance, e.g., "she bought carnations." In both cases, the evidence from where participants looked in a visual world paradigm measured by eye-tracking shows that they focus readily on the presupposed facts as well as the conjecture, particularly for negative counterfactuals. Overall, the results of our experiments are consistent with the theory that the mental representation of a counterfactual requires dual possibilities. Eye-tracking studies have established that both possibilities are represented very early in the processing of the counterfactual, but that attention settles on the presupposed facts and the conjecture gradually falls outside the focus of attention (e.g., de Vega et al., 2007; de Vega & Urrutia, 2011, 2012; Ferguson & Jayes, 2018; Urrutia et al., 2012). Our experiments clarify the factors that determine the process by which the presupposed facts occupy this advantaged position in the representation of counterfactuals.

Acknowledgements

The experiments were designed when the first author was a visiting fellow at Trinity College Dublin for 3 months in 2016. Some of the results were presented at the Meeting in Memory of Vittorio Girotto in London in 2017. We would like to thank David Beltrán, Juan A. García-Madruga, Isabel Gómez-Veiga, Phil Johnson-Laird, Sunny Khemlani, and Carlos Santamaría for their suggestions and comments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Isabel Orenes https://orcid.org/0000-0002-8677-2399

Data accessibility statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/79kcr/

Supplementary material

The supplementary material is available at qjep.sagepub.com.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Beltrán, D., Orenes, I., & Santamaría, C. (2008). Context effects on the spontaneous production of negation. *Intercultural Pragmatics*, *5*(4), 409–419.

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. MIT Press.

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.

Byrne, R. M. J., & Tasso, A. (1999). Deductive reasoning with factual, possible, and counterfactual conditionals. *Memory & Cognition*, *27*, 726–740.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*, 45–73.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*, 472–517.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.

Darley, E. J., Kent, C., & Kazanina, N. (2020). A 'no' with a trace of 'yes': A mouse-tracking study of negative sentence processing. *Cognition*, *198*, Article 104084.

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*, *51*(12), 2401–2414.

de Vega, M., & Urrutia, M. (2011). Counterfactual sentences may activate motoric processes. An action-sentence compatibility effect study. *Journal of Cognitive Psychology*, *23*(8), 962–973.

de Vega, M., & Urrutia, M. (2012). Discourse updating after reading a counterfactual event. *Psicologica*, *33*, 157–173.

de Vega, M., Urrutia, M., & Riffo, B. (2007). Cancelling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition*, *35*, 1410–1421.

Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*, 284–292.

Espino, O., & Byrne, R. M. J. (2012). It is not the case that if you understand a conditional you know how to negate it. *Journal of Cognitive Psychology*, *243*(3), 329–334.

Espino, O., & Byrne, R. M. J. (2018). Thinking about the opposite of what is said: Counterfactual conditionals and symbolic or alternate simulations of negation. *Cognitive Science*, *42*, 2459–2501.

Farshchi, S., Andersson, R., van de Weijer, J., & Paradis, C. (2019). Processing negation in a miniature artificial language. *Cognitive Science*, *43*, Article e12720.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Ferguson, H., & Jayes, L. (2018). Plausibility and perspective influence the processing of counterfactual narratives. *Discourse Processes*, *55*, 166–186.

Ferguson, H. J. (2012). Eye movements reveal rapid concurrent access to factual and counterfactual interpretations of the world. *Quarterly Journal of Experimental Psychology*, *65*, 939–961.

Ferguson, H. J., & Cane, J. (2015). Examining the cognitive costs of counterfactual language comprehension: Evidence from ERPs. *Brain Research*, *1622*, 252–269.

Ferguson, H. J., & Sanford, A. J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language*, *58*, 609–626.

Ferguson, H. J., Sanford, A. J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. *Brain Research*, *1236*, 113–125.

Ferguson, H. J., Scheepers, C., & Sanford, A. J. (2010). Expectations in counterfactual and theory of mind reasoning. *Language, Cognition and Neuroscience*, *25*, 297–346.

Frosch, C. A., & Byrne, R. M. J. (2012). Causal conditionals and counterfactuals. *Acta Psychologica*, *141*(1), 54–66.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878.

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tennenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *2812*(12), 1731–1744.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research*, *1*, 19–33.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive Science*, *25*(4), 565–610.

Greco, M. (2020). On the syntax of surprise negation sentences: A case study on expletive negation. *Natural Language & Linguistic Theory*, *38*, 775–825.

Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *Journal of Pragmatics*, *38*, 1015–1032.

Huettig, F., & Altmann, G. T. (2011). Looking at anything that is green when hearing 'frog': How object surface colour and stored object colour knowledge influence language-mediated overt attention. *The Quarterly Journal of Experimental Psychology, I*, *64*, 122–145.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic, and shape information in language-mediated visual search. *Journal of Memory and Language*, *54*, 460–482.

Huettig, F., & McQueen, J. M. (2011). The nature of the visual environment induces implicit biases during language-mediated visual search. *Memory & Cognition*, *39*, 1068–1084.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171.

Ito, A. (2019). Prediction of orthographic information during listening comprehension: A printed-word visual world study. *The Quarterly Journal of Experimental Psychology*, *72*(11), 2584–2596.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

Johnson-Laird, P. N., & Khemlani, S. (2017). Mental models and causation. In M. Waldman (Ed.), *Oxford handbook of causal reasoning* (pp. 1–42). Oxford University Press.

Johnson-Laird, P. N., & Tridgell, J. M. (1972). When negation is easier than affirmation. *The Quarterly Journal of Experimental Psychology*, *24*(1), 87–91.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, *246*(1), 160–173.

Kaup, B., & Dudschig, C. (2020). Understanding negation: Issues in the processing of negation. In V. Deprez & T. Espinal (Eds.), *The Oxford handbook of negation* (pp. 635–655). Oxford University Press.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, *38*, 1033–1050.

Kaup, B., & Zwaan, R. A. (2003). Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 439–446.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2012). Negation: A theory of its meaning, representation and use. *Journal of Cognitive Psychology*, *24*, 541–559.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2014). The negations of conjunctions, conditionals, and disjunctions. *Acta Psychologica*, *151*, 1–7.

Kulakova, E., Aichhorn, M., Schurz, M., Kronbichler, M., & Perner, J. (2013). Processing counterfactual and hypothetical conditionals: An fMRI investigation. *NeuroImage*, *72*, 265–271.

Kulakova, E., & Nieuwland, M. S. (2016). Understanding counterfactuality: A review of experimental evidence for the dual meaning of counterfactuals. *Language and Linguistics Compass*, *102*(2), 49–65.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*(6), 1036–1073.

Lucas, C. G., & Kemp, C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review*, *122*(4), 700–734.

Lyu, S., Tu, J., & Lin, C. (2020). Processing plausibility in concessive and causal relations: Evidence from self-paced reading and eye-tracking. *Discourse Processes, 57 4*, *574*(4), 320–342.

Macbeth, G., Razumiejczyk, E., Crivello, M., del, C., Bolzán, C., Pereyra Girardi, C. I., & Campitelli, G. (2014). Mental models for the negation of conjunctions and disjunctions. *Europe's Journal of Psychology*, *10*(1), 135–149.

Macbeth, G., Razumiejczyk, E., Crivello, M., del, C., Fioramonti, M., & Pereyra Girardi, C. I. (2015). Prior affirmation facilitates the processing of compound negation. *International Journal of Psychology and Behavioral Sciences*, *5*, 205–215.

Macbeth, G. E., Crivello, M., del, C., Fioramonti, M. B., & Razumiejczyk, E. (2017). Chronometrical evidence supports the model theory of negation. *SAGE Open*, *72*(2), 1–8.

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, *71*(3), 450–463.

Markman, K. D., Mizoguchi, N., & McMullen, M. N. (2008). 'It would have been worse under Saddam': Implications of counterfactual thinking for beliefs regarding the ethical treatment of prisoners of war. *Journal of Experimental Social Psychology*, *44*(3), 650–654.

Martin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380.

Mayo, R., Schul, Y., & Burnstein, E. (2004). 'I am guilty' vs. 'I am innocent': Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*, 443–449.

McEleney, A., & Byrne, R. M. J. (2006). Spontaneous counterfactual thoughts and causal explanations. *Thinking and Reasoning*, *122*(2), 235–255.

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology*, *60*, 661–671.

Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal*, *3*, 119–135.

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition*, *37*, 249–264.

Miller, G. A. (1962). Some psychological studies of grammar. *American Psychologist*, *17*, 748–762.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Chapman & Hall.

Moreno-Rios, S., & Byrne, R. M. J. (2018). Inferences from disclosures about the truth and falsity of expert testimony. *Thinking and Reasoning*, *241*(1), 41–78.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Nieuwland, M. S. (2012). Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri. *NeuroImage*, *59*, 3433–3440.

Nieuwland, M. S. (2016). Quantification, prediction, and the online impact of sentence truth-value: Evidence from event-related potentials. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(2), 316–334.

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science*, *19*(12), 1213–1218.

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25–39.

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *Tutorials in Quantitative Methods for Psychology*, *10*, 56–67.

Orenes, I., Beltrán, D., & Santamaría, C. (2014). How negation is understood: Evidence from the visual world paradigm. *Journal of Memory and Language*, *74*, 36–45.

Orenes, I., García-Madruga, J. A., Gómez-Veiga, I., Espino, O., & Byrne, R. M. J. (2019). The comprehension of counterfactual conditionals: Evidence from eye-tracking in the visual world paradigm. *Frontiers in Psychology*, *10*, Article 1172.

Orenes, I., Moxey, L., Scheepers, C., & Santamaría, C. (2015). Negation in context evidence from the visual world paradigm. *The Quarterly Journal of Experimental Psychology*, *69*, 1082–1092.

Orenes, I., & Santamaría, C. (2014). Visual content of words delays negation. *Acta Psychologica*, *153*, 107–112.

Primativo, S., Sebastian, J. R., & Crutch, J. (2016). Conceptual feature ratings predict gaze within written word arrays: Evidence from a visual wor(l)d paradigm. *Cognitive Science*, *41*, 659–685.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Roese, N., & Epstude, K. (2017). The functional theory of counterfactual thinking. *Advances in Experimental Social Psychology*, *56*, 1–79.

Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1108–1117.

Santamaría, C., Espino, O., & Byrne, R. M. J. (2005). Counterfactual and semifactual conditionals prime alternative possibilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1149–1154.

Stewart, A. J., Haigh, M., & Kidd, E. (2009). An investigation into the online processing of counterfactual and indicative conditionals. *Quarterly Journal of Experimental Psychology*, *62*, 2113–2125.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Thompson, V., & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1154–1170.

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we stimulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology*, *63*, 2305–2312.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, *2*, 239–289.

Urrutia, M., de Vega, M., & Bastiaansen, M. (2012). Understanding counterfactuals in discourse modulates ERP and oscillatory gamma rhythms in the EEG. *Brain Research*, *1455*, 40–55.

Urrutia, M., Gennari, S., & de Vega, M. (2012). An fMRI study of counterfactual sentences describing physical effort. *Neuropsychologia*, *50*, 3663–3672.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experiment Psychology*, *11*, 92–107.

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, *52*, 133–142.

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*, 7–11.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Harvard University Press.

Wason, P. C., & Jones, S. (1963). Negatives: Denotation and connation. *British Journal of Psychology*, *54*, 299–307.

Xiang, M., Kramer, A., & Nordmeyer, A. (2020). An informativity-based account of negation complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), 1857–1867.