The Conjunction Fallacy: Confirmation or Relevance?*

WooJin Chung[a], Kevin Dorst[b], Matthew Mandelkern[c], and Salvador Mascarenhas[d]
[a]*Department of Linguistics, Seoul National University*
[b]*Department of Linguistics and Philosophy, MIT*
[c]*Department of Philosophy, NYU*
[d]*Institut Jean Nicod, Département d'études cognitives, Ecole Normale Supérieure, EHESS, CNRS, PSL University*

Abstract

The conjunction fallacy is the well-documented empirical finding that people sometimes rate a conjunction $A \wedge B$ as more probable than one of its conjuncts, A. Many explanations appeal to the fact that B has a high probability in the scenarios at hand, but Tentori et al. (2013) have challenged such approaches. They report experiments suggesting that *degree of confirmation*—rather than probability—is the central determinant of the conjunction fallacy. In this paper, we have two goals. First, we address a confound in Tentori et al.'s (2013) experiments: they failed to control for the fact that in their stimuli where B is confirmed, it is also *conversationally relevant* in the sense that it fits with the topic or *question under discussion* (Roberts, 2012). Conversely, when B has a high probability but isn't confirmed, it is conversationally *irrelevant*. Consequently, it is possible that conversational relevance, rather than confirmation, is responsible for the differences they found between confirmed and probable hypotheses. Second, inspired by recent theoretical work, we aim to give the first empirical investigation of the hypothesis that this type of conversational relevance on its own—independently of degree of confirmation—can be an important factor in the conjunction fallacy. We report on two experiments that vary Tentori et al.'s (2013) design by making B relevant without changing its degree of probability or confirmation. We found that doing so increases the rate of the conjunction fallacy, suggesting that relevance plays an important role in the conjunction fallacy.

Keywords: conjunction fallacy, confirmation, conversational relevance

1 Introduction

The *conjunction fallacy* is the well-documented empirical finding that people sometimes rate a conjunction $A \wedge B$ as more probable than one of its conjuncts, A—contrary to the laws of probability, which entail that $P(A \wedge B) \leq P(A)$. Here is the classic example from Tversky and Kahneman (1983):

> **Linda**: Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. [e]

*Dorst and Mandelkern contributed equally. Thanks to Sam Mitchell for help constructing the materials, and the LANG-REASON team at Ecole Normale Supérieure for invaluable feedback. This work was supported by Agence Nationale de la Recherche grants ANR-18-CE28-0008 (LANG-REASON; PI: Mascarenhas) and ANR-17-EURE-0017 (FrontCog; Department of Cognitive Studies, Ecole Normale Supérieure), and by the New Faculty Startup Fund from Seoul National University (Chung).

Given this, which is more probable?

(a) Linda is a bank teller.

(b) Linda is a bank teller and is active in the feminist movement.

A large majority chose the conjunction (b) rather than its conjunct (a). The literature has since found structurally parallel phenomena in a wide variety of circumstances (see e.g. Moro, 2009; Tentori et al., 2013, for summaries).

Though the details vary, many explanations for the conjunction fallacy are based on the fact that, given the description, there is a relatively high posterior probability that Linda is active in the feminist movement (Costello, 2009a,b; Nilsson et al., 2009). Of course, a story is needed about how this high posterior probability influences subjects to commit the conjunction fallacy. For an example, one approach (among many) posits that subjects calculate the probabilities of conjunctions by averaging the probabilities of the conjuncts, in which case, if the probability of *feminist* is high and *bankteller* low, the probability of *feminist and bankteller* will be (mistakenly) calculated as being higher than that of *bankteller* (Fantino et al. 1997; Juslin et al. 2009).

Recently, Tentori et al. (2013) have argued against any kind of explanation based on high posteriors (see also Tenenbaum and Griffiths 2001; Crupi et al. 2008; Tentori and Crupi 2012; Mangiarulo et al. 2021; Chung and Mascarenhas 2023). Tentori et al. argue instead that what drives the conjunction fallacy is the fact that the description of Linda in the vignette (e) *confirms* the claim that she's active in the feminist movement, i.e. *raises* its probability. Specifically, where P is the subject's probability function and e is the evidence they get from the vignette, Tentori and collaborators argue that what drives the fallacy is the fact that the perceived value of $P(\textit{feminist} \mid e) - P(\textit{feminist})$ is high (henceforce the *confirmatory value* of e for *feminist*), *not* the fact that $P(\textit{feminist} \mid e)$ is high. As they point out, this distinction is obscured in a case like **Linda** because the vignette (e) *both* confirms that Linda is a feminist, *and* naturally leads subjects to assign high probability to that claim.

To argue for their claim, Tentori et al. present a series of experiments that dissociate high posteriors from confirmation. Here is one of their cases:

Violinist: O. has a degree in violin performance. $[e]$

Which of the following hypotheses do you think is the most probable?

($h1$) O. is an expert mountaineer

($h1 \wedge h2$) O. is an expert mountaineer and gives music lessons

($h1 \wedge h3$) O. is an expert mountaineer and owns an umbrella

Intuitively, 'O. owns an umbrella' is very likely (high posterior), but not at all confirmed by the vignette e; while 'O. gives music lessons' is not very likely (low posterior), but is confirmed by e. So if high posteriors drive the conjunction fallacy, then subjects should be more likely to commit the conjunction fallacy by choosing $h1 \wedge h3$ over $h1$. Meanwhile, if it is confirmation that drives the conjunction fallacy, subjects should be more likely to commit the conjunction fallacy by choosing $h1 \wedge h2$ over $h1$. In experiments with this structure, preceded by norming studies with a different group of participants to estimate the relevant posterior and confirmation quantities, Tentori et al. find evidence that subjects tend to choose $h1 \wedge h2$ over $h1 \wedge h3$ when they commit the conjunction fallacy. They conclude that it is indeed confirmation, not posteriors, that drives the conjunction fallacy.

But there is a confound. In examples like **Violinist** above, there are *two* salient differences between $h2$ and $h3$. One is the difference that Tentori et al. focus on: 'O. gives music lessons' is confirmed but improbable, while 'O. owns an umbrella' is not confirmed but is probable. But there is a second difference: given the vignette e, 'O. gives music lessons' is *conversationally relevant*—it contributes to the topic or question under discussion (see Roberts 2012 and §5 below)—while 'O. owns an umbrella' is not conversationally relevant. To see this intuitively, note that if someone told you that O. has a degree in violin performance, and went on to tell you that he gives music lessons, that would feel like a coherent conversational move; while if they went on to tell you that he owns an umbrella, it would seem odd, and you would wonder why they are telling you this.[1]

On some theories, this type of conversational relevance is a central determinant of the conjunction fallacy. Indeed, Tversky and Kahneman (1983) suggested one possible explanation of the conjunction fallacy was that subjects ranked responses by *informativity* rather than *probability*. The relevant notion of informativity, however, plausibly depends on what question is at stake: relative to the question 'Is Linda a feminist?', 'Linda is a feminist bankteller' is very informative, while 'Linda is a bankteller' is not—even though it is more probable. Although Tversky and Kahneman do not pursue this explanation, more recent work by Levi (2004), Dorst and Mandelkern (2021), and Sablé-Meyer and Mascarenhas (2021) develop theories of the conjunction fallacy where relevance plays a central role, as we will discuss.

In short, Tentori et al.'s (2013) results on their own are consistent with two interpretations: either that confirmation is *the only driver* of the conjunction fallacy, or that posteriors can also be drivers, depending on whether they are *revelant* or not. In other words, relevance may be a central *determinant* of the conjunction fallacy. Of course, very often confirmation and relevance go together: if you're told it looks cloudy, that both increases the probability that it will rain *and* makes the question of *whether it will rain* conversationally relevant. But other times relevance and confirmation come apart: if you're told it looks cloudy, that confirms *it will rain and the speaker's birthday is in July*,[2] but does *not* make the question of whether *it will rain and the speaker's birthday is in July* conversationally relevant.

We have two goals in this paper. The first, local goal is to pull these two drivers apart in order to critically assess Tentori et al.'s conclusions. The second, broader goal is to begin to empirically investigate whether conversational relevance—again, in the specific form of a question under discussion—is an important contributor to the conjunction fallacy, as recent theories predict. We pursue these goals by modifying Tentori et al.'s experiments, adding a vignette which makes both $h2$ and $h3$ conversationally relevant but does not change the probabilities or degrees of confirmation—so that $h2$ is still confirmed but improbable while $h3$ is still probable but not confirmed, but *both* $h2$ *and* $h3$ *are relevant*. For instance, in the **Violinist** case, we compare Tentori et al.'s version, where the context is simply 'O. has a degree in violin performance', to a version where the context instead is the following:

> Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected

[1] A variety of studies have looked at the interaction of Gricean (1989) pragmatic reasoning with the conjunction fallacy (e.g. Adler 1984; Agnoli and Krantz 1989; Dulany and Hilton 1991; Gigerenzer 1991). However, these have mostly focused on *conversational implicatures*, which have been shown not to explain all cases of the conjunction fallacy; see Moro 2009 for an overview. However, we know of no study that, like ours, directly examines the role of *conversational relevance*, the part of Gricean pragmatics that has been developed into (e.g.) the *question under discussion* model of Roberts (2012).

[2] If *rain* and *July* are independent both unconditionally and given e—as, surely, they are—then $P(rain|e) > P(rain)$ iff $P(rain \wedge July|e) > P(rain \wedge July)$.

person, Dan, and starts asking Dan questions. She finds out that Dan has **a degree in violin performance**.

This context—unlike the original—makes the question of whether the person of interest has an umbrella *relevant* without intuitively changing its *probability*. We will directly manipulate relevance by comparing rates of the conjunction fallacy in minimal pairs like this—e.g. the **Violinist** case with and without this context—to detect (1) how much Tentori et al.'s findings support the claim that confirmation drives the conjunction fallacy, rather than that either confirmation or relevance does; and (2) whether relevance itself, apart from confirmation, can be a determinant of the conjunction fallacy.

2 Norming study

Mirroring Tentori et al.'s (2013) excellent design, we first conducted a norming study to ensure that our subjects' perceived probabilities and levels of confirmation of the hypotheses are comparable to those in the original experiment, and that adding contexts does not significantly affect these quantities.[3,4]

2.1 Participants and procedure

We used the English translations of the materials in Tentori et al.'s Experiment 2 to keep the design and materials as close as possible to the original study. Our experiment includes one additional factor, namely the presence or absence of relevant context.

The probability task aims to compare the participants' perceived values of $Pr(h2 \mid e \wedge h1)$ and $Pr(h3 \mid e \wedge h1)$. The task provides the participants with a context conveying e and $h1$, and then asks how probable the hypothesis $h2$ (or $h3$) is given the context. The question was asked in a frequency format (e.g., 'How many of them do you think are $\left\{ \begin{smallmatrix} h2 \\ h3 \end{smallmatrix} \right\}$?') just as in Tentori et al.'s original task—a strategy that encourages more reliable probability judgments (Gigerenzer, 1991). The confirmation task compares the participants' perceived confirmatory values of $h2$ and $h3$. We first presented $h1$ as background information and $h2$ (or $h3$) as a hypothesis, and then e as a new piece of information. We asked to what degree the evidence strengthens or weakens the given hypothesis (see Mastropasqua et al. 2010 for a justification of this task as a measure of confirmation).

Participants saw contexts that made $h3$ relevant to the question under discussion. The context for Tentori et al.'s **Violinist** scenario is provided in (1). We used the same contexts for both the probability task and the confirmation task.

(1) An illustration of a confirmation task with a relevant context:

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that Dan is **an expert mountaineer**.

Consider the **following hypothesis (which could be true or false)**:

Dan **gives music lessons**.

[3] All data and analysis for our three experiments can be found in our OSF respository by following THIS LINK.

[4] All studies reported in this manuscript received ethics approval by the *Comité d'évaluation de l'éthique de l'INSERM*, under research protocol *Le langage et les capacités cognitives connexes*.

Table 1: Our norming study

Task	Context	Hypothesis	Scenario			
			Violinist	Swiss man	Italian student	Swedish woman
Prob	Yes	*h*2	.32	.53	.12	.15
		*h*3	.80	.83	.14	.36
	No	*h*2	.30	.54	.08	.03
		*h*3	.66	.71	.07	.33
Conf	Yes	*h*2	+4.8	+3.7	-1.8	+0.4
		*h*3	-0.5	-0.1	-0.5	-3.2
	No	*h*2	+7.2	+4.2	-0.4	+0.9
		*h*3	-0.3	-0.2	+1.9	-1.3

> Now you are given a new piece of information concerning Dan:
>
> Dan has **a degree in violin performance**.
>
> How does the new piece of information that Dan has **a degree in violin performance** affect the hypothesis that Dan **gives music lessons**?

For each scenario, the participants were asked about either *h*2 or *h*3. We pseudo-randomized the assignment of the hypotheses, following Tentori et al.'s design.

We recruited 241 participants on *Prolific* who are British and native speakers of English. Among them, we excluded 90 participants who did not pass two attention checks (see appendix A.3). We also excluded one participant who entered ill-formatted answers to the target questions (e.g., wrote a value less than 0 or greater than 100 for the frequency task). The median time each participant spent on the experiment was 3 minutes and 49 seconds and the average reward was £7.86/hr.

2.2 Results and discussion

Table 1 shows the means of participants' responses in the probability task and the confirmation task. We analyzed the results separately for each task as the outcomes diverge depending on the task. For both tasks, all mixed effects models either failed to converge or resulted in a singular fit so we used a simple linear regression model with the *lm* function in *R*, fitting the participants' responses with the following predictors: (i) HYPOTHESIS the subjects were presented with, with 2 levels (*h*2 vs. *h*3, where *h*2 is expected by design to rank high on confirmation and low on probability, conversely for *h*3; between-subjects), (ii) CONTEXT with 2 levels (yes-context vs. no-context; between-subjects), and (iii) the interaction of HYPOTHESIS with CONTEXT. The main effect of HYPOTHESIS was significant for the probability task ($p < .001$) and the confirmation task ($p < .001$). But a putative main effect of CONTEXT did not reach significance for either the probability task ($p = .361$) or the confirmation task ($p = .060$). The interaction term was also not significant for the probability task ($p = .481$) or the confirmation task ($p = .939$).

Comparing Tentori et al.'s results summarized in Table 4 to ours in the '*no-context*' condition, all scenarios except **Italian student** show a similar trend: the probability of *h*2 is judged lower than the

Table 2: Output of the model of the probability task in our norming study

Coefficient	Estimate	*SE*	*t*	*p*-value
intercept	0.23	0.04	6.36	$< .001$ ***
yes-context	0.05	0.05	0.91	.361
*h*3	0.21	0.05	3.96	$< .001$ ***
*h*3:yes-context	0.05	0.07	0.71	.481

Table 3: Output of the model of the confirmation task in our norming study

Coefficient	Estimate	*SE*	*t*	*p*-value
intercept	0.64	0.02	26.11	$< .001$ ***
yes-context	-0.07	0.03	-1.89	.060 .
*h*3	-0.15	0.03	-4.18	$< .001$ ***
*h*3:yes-context	0.004	0.05	0.08	.939

probability of *h*3 but the confirmatory value of *h*2 is judged higher than the confirmatory value of *h*3. The result is in line with Tentori et al.'s original norming study and thus provides concrete grounds for our conjunction fallacy experiments. Regarding the failure of replication in the **Italian student** scenario, we speculate that British (our participants) and Italian (Tentori et al.'s) participants might have different ideas of what Italian undergraduate students are likely to do.

The presence of relevant context did not significantly influence the perceived probabilities or levels of confirmation for both *h*2 and *h*3. While it raised participants' judgment about *h*3 slightly more than judgment about *h*2, the difference was very small. Overall, the results do not allow us to completely disregard the potential worry that the introduction of context adds a confound to the forthcoming conjunction fallacy experiments, namely that context directly manipulates the perceived probabilities and levels of confirmation of *h*2 and *h*3. However, the results suggest that the direct influence of context to perceived probabilities and levels of confirmation does not fully explain people's behavior. Moreover, since our norming study replicated the crucial intended contrast regarding how *h*2 and *h*3 fare on probability and confirmation, we see no reason to think that any potential impact of context in participants' probability and confirmation judgments interferes with our design. Consequently, we conducted two experiments using these normed materials to investigate whether

Table 4: Tentori et al.'s norming study

Task	Hypothesis	Scenario			
		Violinist	Swiss man	Italian student	Swedish woman
Prob	*h*2	.35	.68	.16	.19
	*h*3	.67	.83	.12	.25
Conf	*h*2	+5.6	+4.7	+3.9	+2.6
	*h*3	-0.1	-0.6	-0.4	-4.1

relevance is a determinant of the conjunction fallacy by testing the effect of adding a context which makes $h3$ relevant without changing its probability.

3 Experiment 1

3.1 Participants and procedure

Experiment 1 uses the English translations of the items in Tentori et al.'s conjunction fallacy study, with a few adjustments. The central change is that we added the between-subjects CONTEXT factor. By contrasting the responses in the context condition with the responses in the no-context condition, we can detect any effect of conversational relevance.

The second change we made is that rather than asking participants to choose the most probable from the three options $h1$, $h1 \wedge h2$, and $h1 \wedge h3$, we contrasted two hypotheses at a time, i.e., $h1$ vs. $h1 \wedge h2$ or $h1$ vs. $h1 \wedge h3$. This addresses an issue overlooked in the original Tentori et al. study: asking which hypothesis is the most probable potentially conceals a lot of conjunction errors; $h1 \wedge h2$ could have been participants' top choice, but they could have made an $h1 \wedge h3$ conjunction error at the same time. Our updated design detects such hidden errors. Overall, we adopted a 2×2 design, crossing HYPOTHESIS with CONTEXT.

We recruited 150 native speakers of English from the United Kingdom via *Prolific*. Among them, 66% were female and the mean age was 37. Half of the participants were provided with contexts that made $h3$ relevant (i.e. the CONTEXT condition), and the other 75 were not (the NO-CONTEXT condition). For pseudo-randomization, each of the two groups was further divided into 3 subgroups, where the subgroups differed in the order in which the scenarios were presented. The median time each participant spent on the experiment was 3 minutes and the average reward was £15.00/hr.

3.2 Results and discussion

Table 5 summarizes participants' responses. We analyzed the data using a generalized linear mixed effects model with the *glmer* function in *R*. We coded the outcome variable CONJ_ERROR which was valued 'yes' if a participant judged the presented conjunction more probable than $h1$, and 'no' otherwise. The model contained (i) HYPOTHESIS with 2 levels ($h1 \wedge h2$ vs. $h1 \wedge h3$; within-subjects) (ii) CONTEXT with 2 levels (yes-context vs. no-context; between-subjects), (iii) the interaction between HYPOTHESIS and CONTEXT, and (iv) by-participant random intercepts. The interaction term HYPOTHESIS : CONTEXT was positive and significant ($p = .001$). A model comparison between the full model and a simpler model without the interaction term using the likelihood ratio test revealed that the former outperforms the latter ($p = .002$). Tables 6 and 7 summarize the fitted model and the model comparison results, respectively.

When contexts that make $h3$ relevant were not provided, we found a trend reminiscent of Tentori et al. (2013)'s report: the rate of conjunction errors was notably higher when $h1$ was contrasted with $h1 \wedge h2$ (43%) than when it was contrasted with $h1 \wedge h3$ (25%). However, crucially, when such contexts were provided, this contrast disappeared, and the rate of conjunction errors in both conditions were comparable (35% vs. 32%). The significant interaction between HYPOTHESIS and CONTEXT confirms our hypothesis that the presence of context making $h3$ relevant increased the rate of conjunction errors for $h1 \wedge h3$ more than for $h1 \wedge h2$.

Table 5: Our conjunction fallacy result (Experiment 1). The percentage points in parentheses indicate the proportion of responses within each HYPOTHESIS × CONTEXT condition.

HYPOTHESIS	CONJ_ERROR	CONTEXT	
		No	Yes
$h1 \wedge h2$	No	172 (57%)	195 (65%)
	Yes	128 (43%)	105 (35%)
$h1 \wedge h3$	No	226 (75%)	205 (68%)
	Yes	74 (25%)	95 (32%)

Table 6: Output of the model of Experiment 1 looking for the interaction effect between HYPOTHESIS ($h1 \wedge h2$ vs. $h1 \wedge h3$) and CONTEXT (yes-context vs. no-context)

Coefficient	Estimate	*SE*	*z*	*p*-value
intercept	-0.45	0.23	-1.99	.047 *
yes-context	-0.49	0.32	-1.51	.130
$h1 \wedge h3$	-1.14	0.21	-5.43	$<$ .001 ***
$h1 \wedge h3$:yes-context	0.93	0.29	3.20	.001 **

Table 7: Likelihood ratio test (Experiment 1) comparing model with interaction term (model 1) and model without interaction term (model 2)

Model	*df*	LogLik	Chisq	*p*-value
model 1	5	-671.34		
model 2	4	-676.33	9.99	.002 **

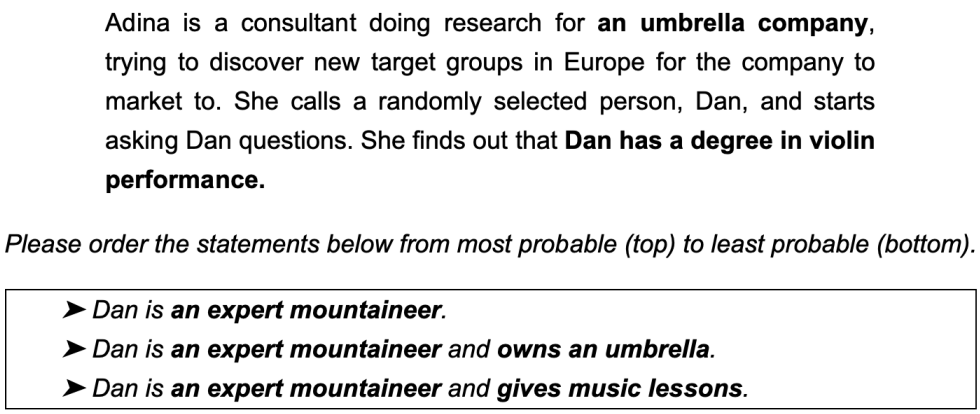


Figure 1: An example of the contextualized ranking test from Experiment 2

We should emphasize again that our methodology was slightly different from Tentori et al.'s, for the reasons we explained above, making a direct comparison with their results tricky. Crucially, we only presented two hypotheses at a time. Yet, it is possible to compare the results within our experiment, as we directly manipulated the categorical variable CONTEXT. We found that simply adding a context which makes $h3$ relevant increased rates of conjunction errors for $h1 \wedge h3$, in particular, leading to comparable rates of errors for $h1 \wedge h2$ and $h1 \wedge h3$ (recall that $h2$ is always contextually relevant thanks to the vignette e). This provides experimental evidence for the claim that *conversational relevance*, independently of confirmation, is one determinant of the conjunction fallacy.[5]

4 Experiment 2

In Experiment 2, we present all three competing hypotheses ($h1, h1 \wedge h2$, and $h1 \wedge h3$) in every question, just as in Tentori et al.'s original experiment. Recall that Experiment 1 only presented two hypotheses at a time and asked to select the most probable, contrasting $h1$ with either $h1 \wedge h2$ or $h1 \wedge h3$. As noted earlier, this allowed us to detect hidden conjunction errors in cases where $h1 \wedge h2$ and $h1 \wedge h3$ are both deemed more likely than $h1$. While this served the central purpose of Experiment 1, it resulted in a departure from Tentori et al.'s original experimental design. Experiment 2 improves on this by presenting participants with a ranking task with all three hypotheses. This way each participant saw all three hypotheses at once, as in Tentori et al.'s study, while the ranking element allowed us to detect conjunction errors with either of the options of interest.

As illustrated in Figure 1, we asked participants to *order* the three competing hypotheses from most to least probable. By looking at the order between $h1 \wedge h2$ or $h1 \wedge h3$ on the one hand and $h1$ on the other, we can check whether people committed either of the potential conjunction errors each experimental scenario allowed.

[5] An anonymous reviewer observes that the proportion of participants who did not commit the fallacy is rather high in our experiment (57% for $h1 \wedge h2$ and 75% for $h1 \wedge h3$ in the no-context condition) compared to what was reported in Tentori et al.'s Experiment 2 (28%). This is an intriguing fact, and we can only speculate that this should be due to differences in the two populations, in particular the very different mean ages (*Prolific* population vs. Italian undergraduates), as well as the level of experience taking psychological experiments.

Vera is a consultant doing research for **a newspaper company**, trying to discover new target groups for the company to market to. She calls a randomly selected person, Mary, and starts asking Mary questions. She finds out that **Mary lives in Liverpool.**

Please order the statements below from most probable (top) to least probable (bottom).

➤ *Mary is a fan of Manchester United.*
➤ *Mary lives in England.*
➤ *Mary is married.*

Figure 2: Second practice trial of Experiment 2 (yes-context condition)

4.1 Participants and procedure

At the beginning of the experiment, we provided the participants with two practice trials to familiarize them with the task. In the first practice trial, we asked participants to order the letters 'a', 'b', and 'c' in alphabetical order (the initial order was randomized). In the second practice trial, we provided natural language sentences and asked the participants to order them based on how probable they are. As depicted in Figure 2, the second trial was somewhat suggestive of what the main task would look like, but crucially the sentences included no conjunctions. The conjunction fallacy tasks followed the practice trials. We reused all four scenarios from Experiment 1, and just like in Experiment 1, half of the participants were presented with contexts that made hypothesis $h3$ relevant. The order of statements was randomized for all trials to suppress order effects.

We recruited 599 participants via *Prolific*. We decided to be more conservative with the sample size and increased it because we found relatively low rates of conjunction errors in Experiment 1. Moreover, Experiment 2 uses a new methodology involving ranking potential conjunction fallacy triggers.[6] Among the participants, 58% were female and their mean age was 37. We used the first practice trial as a control and excluded 21 participants who did not properly order the letters. 277 participants remained in the 'no-context' condition (9 excluded) and 301 remained in the 'yes-context' condition (12 excluded). For pseudo-randomization, each of the two groups was further divided into 3 subgroups, where the subgroups differed in the order in which the scenarios were presented. The median time each participant spent on the experiment was 3 minutes and 19 seconds and the average reward was £9.05/hr.

4.2 Results and discussion

We coded the results in the following way. From each trial by each participant we created two observations, one for how they treated $h1 \wedge h2$ and another for how they treated $h1 \wedge h3$. Our responses

[6] In a pilot study, we recruited 255 participants and analyzed the responses of 242 after excluding those who failed to properly order the letters in the first practice trial. We got a marginally significant interaction term ($p = .046$). We ran a power analysis using the SIMR package, and the simulation results suggested that we need roughly 600 participants to achieve the power of .8 at .01 significance level (each run gave a slightly different result due to the lack of analytic solution for mixed effects models).

Table 8: Our conjunction fallacy result (Experiment 2). The percentage points in parentheses indicate the proportion of responses within each HYPOTHESIS × CONTEXT condition.

HYPOTHESIS	CONJ_ERROR	CONTEXT	
		No	Yes
$h1 \wedge h2$	No	832 (75%)	844 (70%)
	Yes	276 (25%)	360 (30%)
$h1 \wedge h3$	No	906 (82%)	860 (71%)
	Yes	202 (18%)	344 (29%)

Table 9: Output of the model of Experiment 2 looking for the interaction effect between HYPOTHESIS ($h1 \wedge h2$ vs. $h1 \wedge h3$) and CONTEXT (yes-context vs. no-context)

Coefficient	Estimate	*SE*	*z*	*p*-value
intercept	-2.05	0.20	-10.13	$< .001$ ***
yes-context	0.29	0.28	1.01	0.311
$h1 \wedge h3$	-0.62	0.13	-4.73	$< .001$ ***
$h1 \wedge h3$:yes-context	0.51	0.18	2.85	.004 **

column CONJ_ERROR was filled in as 'yes' if participants ranked $h1 \wedge hn$ as more probable than $h1$, for n the hypothesis in the observation in question.

Table 8 summarizes participants' responses. We analyzed the data using a generalized linear mixed-effects model with the *glmer* function in *R*, fitting participants' responses into the largest converging model which includes (i) HYPOTHESIS with 2 levels ($h1 \wedge h2$ vs. $h1 \wedge h3$; within-subjects), (ii) CONTEXT with 2 levels (yes-context vs. no-context; between-subjects), (iii) the interaction between HYPOTHESIS and CONTEXT, and (iv) by-participant random intercepts and random slopes for CONTEXT. The estimate of the interaction term HYPOTHESIS : CONTEXT was positive and significant ($p = .004$). A model comparison between the full model and a simpler one lacking the interaction term using the likelihood ratio test revealed that the former outperforms the latter ($p = .004$). Tables 9 and 10 summarize the fitted model and the model comparison results, respectively.

Again, we found a trend in the direction we would expect if conversational relevance is one driver of the conjunction fallacy. In the absence of relevant context, we observed a higher rate of $h1 \wedge h2$ conjunction errors (25%) than $h1 \wedge h3$ conjunction errors (18%). However, presenting participants with a context making $h3$ relevant boosted the rate of $h1 \wedge h3$ conjunction errors to 29%, which is

Table 10: Likelihood ratio test (Experiment 2) comparing model with interaction term (model 1) and model without interaction term (model 2)

Model	*df*	LogLik	Chisq	*p*-value
model 1	7	-2058.8		
model 2	6	-2062.2	8.16	0.004 **

comparable to the 30% rate of $h1 \wedge h2$ conjunction errors. Given how we encoded the dependent variable CONJ_ERROR, the significant interaction between HYPOTHESIS and CONTEXT indicates that the presence of relevant context had a greater effect on participants' judgments regarding the ranking between $h1 \wedge h3$ and $h1$ than the ranking between $h1 \wedge h2$ and $h1$.[7]

5 General discussion

Tentori et al. (2013) found that the conjunction fallacy was more likely to occur for certain hypotheses that were confirmed by the vignette but not very probable ($h1 \wedge h2$) than others that had a high posterior probability but were not confirmed ($h1 \wedge h3$). They interpreted this as showing that it is confirmation, not high posteriors, that drives the conjunction fallacy.

Our first, negative goal was to point out a confound in their study which casts doubt on their conclusions: in the cases they used, $h1 \wedge h2$ was *both* confirmed *and* conversationally relevant, while $h1 \wedge h3$ had a high posterior but was conversationally *ir*relevant. This leaves open the possibility that the difference they found between confirmation and posteriors may have been due at least partly to the difference in relevance. We tested this hypothesis by constructing conditions which held the facts about confirmation and posteriors fixed, but ensured that both hypotheses were conversationally relevant. We found that making both hypotheses relevant by adding a context increased the rate at which participants committed the conjunction fallacy for the high-posterior hypothesis ($h1 \wedge h3$). Indeed, it made the rates of conjunction fallacy entirely comparable in the high-confirmation ($h1 \wedge h2$) and high-posterior ($h1 \wedge h3$) conditions (35% vs. 32% in Experiment 1, and 30% vs. 29% in Experiment 2).

These results cast doubt on Tentori et al.'s conclusions that confirmation "prevails as a determinant of the conjunction fallacy" over posterior probability (2013, p. 250), since they did not control for the role of conversational relevance. Of course, this is consistent with confirmation in fact playing an important role in determining the conjunction fallacy—but we cannot take Tentori et al.'s experiments to have demonstrated that confirmation is a more important determiner of the conjunction fallacy than posteriors.

Our second, positive goal was to begin to empirically explore the role of conversational relevance in the conjunction fallacy. On this front, the results of our two experiments are clear: conversational relevance is one determinant of the conjunction fallacy. Clear support for this claim comes from our finding that changing $h3$ from conversationally irrelevant to conversationally relevant—without changing its degree of confirmation or probability—substantially increased rates of the conjunction fallacy involving $h3$. This is the central positive contribution of our paper. In the remainder of this discussion section, we explore the ramifications of this finding for theories of the conjunction fallacy.

The idea that conversational relevance may matter for the conjunction fallacy has not to our knowledge been directly tested. However, this idea *has* been mooted in theoretical discussions of the conjunction fallacy. Indeed, the roots of this idea can be already be found in Tversky and Kahneman (1983), who suggested that one explanation of the fallacy could be that subjects aim to be *informative* in their answers, and hence choose the more informative conjunction over the less informative

[7] The overall rate of conjunction errors was notably even lower than what we had observed in Experiment 1, see footnote 5. As before, we can only speculate as to the causes of these low error rates, since our specific ranking paradigm is to the best of our knowledge novel. We suspect that allowing the participants to manually change the order of statements would have made them more attentive than simply choosing one option, but the methodological question here is interesting and deserves further study.

conjunct. Crucially, informativity is closely related to relevance: whether something is informative depends on what we are trying to figure out. If we are trying to find out whether Linda is a feminist, then 'Linda is a feminist' is informative, while if we are trying to find out whether Linda is a bankteller, then 'Linda is a feminist' is not very informative at all.

Tversky and Kahneman (1983) themselves quickly dismissed the informativity approach—the problem, they said, is that there is no explanation of *why* subjects would tap into informativity when they were asked about probability. However, the idea was raised again by Levi (2004), and has recently been given extensive formal exposition and defense in different ways by Sablé-Meyer and Mascarenhas (2021), Guerrini et al. (2022), and Dorst and Mandelkern (2021).

The idea behind all these approaches is that informativity guides subjects' responses to conjunction fallacy stimuli because informativity guides conversation, and, indeed, cognition in general.[8] We cannot expand on our stronger cognition-wide agenda in any detail here, so we will proceed in this section focusing exclusively on the special case of properly communicative situations, for simplicity and concreteness.

The basic situation that agents find themselves in when communicating with each other is this: some question is (implicitly or explicitly) posed, and cooperative agents make assertions or guesses that are informative about the answer to that question (see Roberts 1989 for a standard exposition of this view in linguistic pragmatics).

There are different ways to leverage these uncontroversial facts about question-answer dynamics in conversation as accounts of the conjunction fallacy. Sablé-Meyer and Mascarenhas (2021), building on the general framework of Koralus and Mascarenhas (2013, 2018) on erotetic, or question-based, processes in higher cognitive faculties, propose that participants are trying to answer a question with two possible answers, namely the ones given explicitly as options, resulting in the answer set $\{\text{bankteller}, \text{bankteller} \wedge \text{feminist}\}$. They look for an answer to this question elsewhere in the environment, and find that the only plausible locus for it is in the description of Linda that occurs in the vignette preceding the question. Interpreted as a relevant answer to the question, the description of Linda can only be a hint in the direction of the conjunction, rather than the simpler option. Sablé-Meyer and Mascarenhas (2021) frame the account in these confirmation-theoretic terms, which—given our current findings—may be too limited. But Guerrini et al. (2022) demonstrated that confirmation-theoretic behavior in question-answer dynamics can be derived by standard extensions of the Rational Speech Act model to accommodate questions under discussion (Frank and Goodman, 2012). This suggests that generalizing an account like this beyond a confirmation-theoretic framework, to account for our present findings, may be possible.

A different recent account, from Dorst and Mandelkern (2021); Mandelkern and Dorst (2022), is part of a broader theory of belief and assertion. The proposal is that cognition in general, and *conversational dynamics* in particular, are guided by a tradeoff between informativity and probability (expected accuracy). Rather than appealing to confirmation, Dorst and Mandelkern spell out their notion of informativity in terms of *what proportion of answers to the relevant question* a given answer rules out. For instance, if the question is 'Who will win the race?' and A, B, and C are the candidates, then 'A' is a more informative answer than 'A or B', since it rules out more potential answers to the question. Thus 'A' may be a *better guess* than 'A or B' about the question 'Who will win?',

[8] In this regard, we are very sympathetic to the fundamental tenets of relevance theory (Sperber and Wilson, 1986), as we strongly suspect that the role of relevance in integrating information to answer a question isn't predicated on proper communicative contexts. An actual comparison between the two approaches we consider here and relevance theory is however impossible, since relevance theory is in fact a framework and not a concrete theory existing at the same level of mathematical rigor as the two theories we discuss here.

even though it is less probable. Since this notion of informativity depends on what question is under discussion and is *not* based on confirmation, it is a potentially helpful notion for making sense of our current results.

On this view, questions are partitions—i.e. mutually exclusive and collectively exhaustive descriptions of the way the world could be. Since the vignette and possible answers make salient two binary questions—*Is Linda a feminist?* and *Is Linda a bank teller?*—the question under discussion will at least include the possible answers to these questions: {*feminist teller*, *feminist non-teller*, *non-feminist teller*, *non-feminist non-teller*}.[9] In such a context, 'feminist bank teller' is more informative than 'bank teller', since the former rules out three of the possible complete answers while the latter only rules out two. Hence in trying to answer the question under discussion, subjects may well choose the more informative but less likely option *feminist teller* over the less informative option *teller*, in conformity with the general cognitive and communicative practice of balancing informativity and accuracy.

To see precisely what this means for the materials in Tentori et al. and our experiments, recall the umbrella scenario. In the original version we critiqued here, subjects are told only that O. has a degree in violin performance. Hence whether or not O. has an umbrella is not contextually relevant: it doesn't answer any relevant question under discussion. By contrast, questions about *music* are made relevant by this set-up; subsequent claims about O.'s musical career will be felt to be more relevant than claims about whether O. has an umbrella. Consequently, subjects will judge 'O. is an expert mountaineer and gives music lessons' to be more informative (relative to this context) than 'O. is an expert mountaineer'. By contrast, 'O. is an expert mountaineer and owns an umbrella' will *not* be more informative than 'O. is an expert mountaineer' in this context, since 'O. owns an umbrella' does not address the context's question under discussion. Thus this theory predicts that subjects will not generally be inclined to rate 'O. is an expert mountaineer and owns an umbrella' as more likely than 'O. is an expert mountaineer'.

Crucially, when we change the context to include *both* information about umbrellas *and* information about music—as in our versions of the experiment—it becomes relevant whether O. owns an umbrella in addition to whether he gives music lessons. In this context, *both* 'O. is an expert mountaineer and gives music lessons' *and* 'O. is an expert mountaineer and owns an umbrella' will be more informative than 'O. is an expert mountaineer', since both answer relevant questions, and so this theory predicts that subjects will commit the conjunction fallacy with both (provided their posteriors on 'O. gives music lessons' and 'O. owns an umbrella' are high enough)—matching our findings.[10]

There is much to be said in favor and against different incarnations of the broad idea that *relevance matters*. On the one hand, Sablé-Meyer and Mascarenhas's account yields a general theory

[9]The move away from what superficially looks like the target question "bankteller or feminist bankteller?" to some conjunction of "feminist?" and "bankteller?" is by no means unprecedented. In the quantum-probability theory approach to the conjunction fallacy, the very same move is invariably made, only there it is supposed that a choice must be made between *first* addressing "feminist?" *and then* "bankteller?" or the other way around (Pothos and Busemeyer, 2022; Pothos et al., 2017).

[10]As a reviewer points out, the conjunction fallacy can occur without framing or conversational context (Costello, 2009a; Wedell and Moro, 2008), casting doubt on the ability of conversational-relevance-based accounts to explain the breadth of the data. Indeed, we agree that the conjunction fallacy is a broad phenomenon, and it's unlikely that any single theory will explain all its instances. As Dorst and Mandelkern 2021, §§2 and 4.3 argue, if the question under discussion model is correct then answering questions will always involve responding to a (perhaps implicit) question under discussion, even if it is not explicitly set by an earlier part of the conversation. Finally, it is worth noting that within the present dialectic, the confirmation-theoretic account *also* requires some contextually-salient bit of information (like the vignette) to predict any instances of the conjunction fallacy, so this worry applies to both types of accounts.

of *representativeness* reasoning as documented by Kahneman and Tversky (1973) and Tversky and Kahneman (1983), as well as a host of ostensibly *deductive* reasoning fallacies from the literature, providing an ambitiously unified approach to a broad range of puzzling reasoning patterns. However, its framing in terms of confirmation means it needs some supplementation to apply to the present findings. Here we want to emphasize instead the points of convergence between these recent theories, again restricting ourselves to properly communicative special cases for simplicity and concreteness: the idea that apparent fallacies can arise out of the fundamentally rational cooperative enterprise of sharing and coordinating on information. Since that enterprise is guided by *questions*, it's to be expected that the conjunction fallacy should also be sensitive to what question is made salient.[11]

6 Conclusion

Our experiments prompt two conclusions. First, Tentori et al.'s (2013) interpretation of their experiments as showing that confirmation, rather than posteriors, is the prime driver of the conjunction fallacy is too quick, for those results may in fact depend crucially on conversational relevance.

Second, and more broadly: conversational relevance matters to the conjunction fallacy. Whether a conjunct is relevant in a given context is one of the factors that determines whether subjects will make the conjunction fallacy with that conjunct. This is an important new empirical observation, which supports a theoretical account on which relevance is one of the factors that leads subjects to rank a conjunction as more likely than one of its conjuncts. We surveyed what to our knowledge are the only two mathematically rigorous accounts of this kind in the literature.

Our main goal here was not to argue for a particular account, but rather to put forward this novel empirical finding as a constraint on any positive account, and to briefly examine the only two families of accounts in the literature that directly and rigorously involve relevance with respect to questions. We think that the fact that relevance matters to the conjunction fallacy supports more generally an approach to psychological effects that pays careful attention to the linguistic context in which the judgments in question are elicited, with the goal of making sense of some of these judgments with tools from linguistics. From this point of view, our findings here contribute to a broader research program which aims to bring the insights of linguistics and philosophy of language to bear on the study of human reasoning.

References

Adler, J. E. (1984). Abstraction is uncooperative. *Journal for the Theory of Social Behaviour*.

Agnoli, F. and Krantz, D. (1989). Suppressing natural heuristic by formal instruction: The case of the conjunction fallacy. *Cognitive Psychology*, 21(515-550).

Chung, W. and Mascarenhas, S. (2023). Modality, expected utility, and hypothesis testing. *Synthese*, 202(11).

Costello, F. J. (2009a). Fallacies in probability judgments for conjunctions and disjunctions of everyday events. *Journal of Behavioral Decision Making*, 22:235–251.

Costello, F. J. (2009b). How Probability Theory Explains the Conjunction Fallacy. *Journal of Behavioral Decision Making*, 22:213–234.

[11] A different, intriguing possibility, suggested to us by an anonymous reviewer for this journal, would be to incorporate relevance into an averaging model of the conjunction fallacy.

Crupi, V., Fitelson, B., and Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Thinking & Reasoning*, 14(2):182–199.

Dorst, K. and Mandelkern, M. (2021). Good guesses. *Philosophy and Phenomenological Research*.

Dulany, D. E. and Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition*, 9(1):85–110.

Fantino, E., Kulik, J., Stolarz-fantino, S., and Wright, W. (1997). The conjunction fallacy: A test of averaging hypotheses. *Psychonomic Bulletin & Review*, 4(1):96–101.

Frank, M. C. and Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084):998.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology*, 2(1):83–115.

Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press, Cambridge.

Guerrini, J., Sablé-Meyer, M., and Mascarenhas, S. (2022). An explanation of representativeness: Contrastive confirmation-theoretical reasoning motivated by question-answer dynamics. Talk given at the 44th Annual Meeting of the Cognitive Science Society.

Juslin, P., Nilsson, H., and Winman, A. (2009). Probability Theory, Not the Very Guide of Life. *Psychological Review*, 116(4):856–874.

Kahneman, D. and Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4):237.

Koralus, P. and Mascarenhas, S. (2013). The erotetic theory of reasoning: bridges between formal semantics and the psychology of deductive inference. *Philosophical Perspectives*, 27:312–365.

Koralus, P. and Mascarenhas, S. (2018). Illusory inferences in a question-based theory of reasoning. In *Pragmatics, Truth and Underspecification*, pages 300–322. Brill.

Levi, I. (2004). Jaakko Hintikka. *Synthese*, 140(1):37–41.

Mandelkern, M. and Dorst, K. (2022). Assertion is weak. *Philosophers' Imprint*, 22(19):1–20.

Mangiarulo, M., Pighin, S., Polonio, L., and Tentori, K. (2021). The Effect of Evidential Impact on Perceptual Probabilistic Judgments. *Cognitive Science*, 45(1).

Mastropasqua, T., Crupi, V., and Tentori, K. (2010). Broadening the study of inductive reasoning: Confirmation judgments with uncertain evidence. *Memory & cognition*, 38(7):941–950.

Moro, R. (2009). On the nature of the conjunction fallacy. *Synthese*, 171(1):1–24.

Nilsson, H., Winman, A., Juslin, P., and Hansson, G. (2009). Linda is not a bearded lady: Configural weighting and adding as the cause of extension errors. *Journal of Experimental Psychology: General*, 138(4):517.

Pothos, E. M. and Busemeyer, J. R. (2022). Quantum cognition. *Annual Review of Psychology*, 73:749–778.

Pothos, E. M., Busemeyer, J. R., Shiffrin, R. M., and Yearsley, J. M. (2017). The rational status of quantum cognition. *Journal of Experimental Psychology: General*, 147(7):968–987.

Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12:683–721.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69.

Sablé-Meyer, M. and Mascarenhas, S. (2021). Indirect illusory inferences from disjunction: a new bridge between deductive inference and representativeness. *Review of Philosophy and Psychology*, 12(2).

Sperber, D. and Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University

Press Cambridge, MA.

Tenenbaum, J. B. and Griffiths, T. L. (2001). The rational basis of representativeness. In Moore, J. D. and Stenning, K., editors, *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, pages 1036–1042.

Tentori, K. and Crupi, V. (2012). How the conjunction fallacy is tied to probabilistic confirmation: Some remarks on Schupbach (2009). *Synthese*, 184(1):3–12.

Tentori, K., Crupi, V., and Russo, S. (2013). On the determinants of the conjunction fallacy: Probability versus inductive confirmation. *Journal of Experimental Psychology: General*, 142(1):235.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4):293–315.

Wedell, D. H. and Moro, R. (2008). Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type. *Cognition*, 107(1):105–136.

A Norming study materials

A.1 Probability task

Italian student

No-context

Consider 100 people who are **Italian undergraduate students** and have **red hair**.

How many of them do you think studied abroad in Barcelona in 2017?

Consider 100 people who are **Italian undergraduate students** and have **red hair**.

How many of them do you think spent their summer holidays in America in 2017?

Yes-context

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Italian undergraduate students** and have **red hair**.

How many of them do you think studied abroad in Barcelona in 2017?

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Italian undergraduate students** and have **red hair**.

How many of them do you think spent their summer holidays in America in 2017?

Swedish woman

No-context

Consider 100 people who are **Swedish women** and **study in Italy**.

How many of them do you think work as a model?

Consider 100 people who are **Swedish women** and **study in Italy**.

How many of them do you think have brown hair?

Yes-context

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Swedish women** and **study in Italy**.

How many of them do you think work as a model?

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Swedish women** and **study in Italy**.

How many of them do you think have brown hair?

Swiss man

No-context

Consider 100 people who are **Swiss men** and **like making Italian desserts**.

How many of them do you think can ski?

Consider 100 people who are **Swiss men** and **like making Italian desserts**.

How many of them do you think have a driving license?

Yes-context

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Swiss men** and **like making Italian desserts**.

How many of them do you think can ski?

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who are **Swiss men** and **like making Italian desserts**.

How many of them do you think have a driving license?

Violinist

No-context

Consider 100 people who have **a degree in violin performance** and are **expert mountaineers**.

How many of them do you think give music lessons?

Consider 100 people who have **a degree in violin performance** and are **expert mountaineers**.

How many of them do you think own an umbrella?

Yes-context

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who have **a degree in violin performance** and are **expert mountaineers**.

How many of them do you think give music lessons?

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She is interviewing a focus group of 100 people who have **a degree in violin performance** and are **expert mountaineers**.

How many of them do you think own an umbrella?

A.2 Confirmation tasks

Italian student

No-context

Carlo has **red hair**.

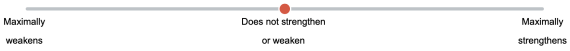
Consider the following **hypothesis (which could be true or false)** concerning Carlo:

Carlo **studied abroad in Barcelona in 2017**.

Now you are given a new piece of information concerning Carlo:

Carlo is **an Italian undergraduate student**.

How does the new piece of information that Carlo is **an Italian undergraduate student** affect the hypothesis that Carlo **studied abroad in Barcelona in 2017**?



Carlo has **red hair**.

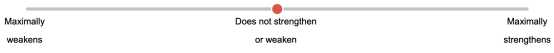
Consider the following **hypothesis (which could be true or false)** concerning Carlo:

Carlo **spent his summer holidays in America in 2017**.

Now you are given a new piece of information concerning Carlo:

Carlo is **an Italian undergraduate student**.

How does the new piece of information that Carlo is **an Italian undergraduate student** affect the hypothesis that Carlo **spent his summer holidays in America in 2017**?



Yes-context

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Carlo, and starts asking Carlo questions. She finds out that Carlo has **red hair**.

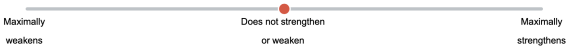
Consider the following **hypothesis (which could be true or false)** concerning Carlo:

Carlo **studied abroad in Barcelona in 2017**.

Now you are given a new piece of information concerning Carlo:

Carlo is **an Italian undergraduate student**.

How does the new piece of information that Carlo is **an Italian undergraduate student** affect the hypothesis that Carlo **studied abroad in Barcelona in 2017**?



Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Carlo, and starts asking Carlo questions. She finds out that Carlo has **red hair**.

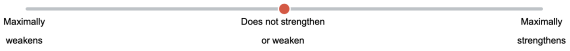
Consider the following **hypothesis (which could be true or false)** concerning Carlo:

Carlo **spent his summer holidays in America in 2017**.

Now you are given a new piece of information concerning Carlo:

Carlo is **an Italian undergraduate student**.

How does the new piece of information that Carlo is **an Italian undergraduate student** affect the hypothesis that Carlo **spent his summer holidays in America in 2017**?



Swedish woman

No-context

Alice **studies in Italy**.

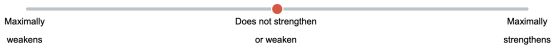
Consider the following **hypothesis (which could be true or false)** concerning Alice:

Alice **works as a model**.

Now you are given a new piece of information concerning Alice:

Alice is **a Swedish woman**.

How does the new piece of information that Alice is **a Swedish woman** affect the hypothesis that Alice **works as a model**?



Alice **studies in Italy**.

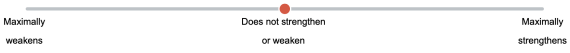
Consider the following **hypothesis (which could be true or false)** concerning Alice:

Alice has **brown hair**.

Now you are given a new piece of information concerning Alice:

Alice is **a Swedish woman**.

How does the new piece of information that Alice is **a Swedish woman** affect the hypothesis that Alice has **brown hair**?



Yes-context

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Alice, and starts asking Alice questions. She finds out that Alice **studies in Italy**.

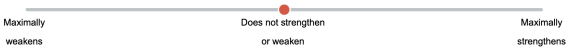
Consider the following **hypothesis (which could be true or false)** concerning Alice:

Alice **works as a model**.

Now you are given a new piece of information concerning Alice:

Alice is **a Swedish woman**.

How does the new piece of information that Alice is **a Swedish woman** affect the hypothesis that Alice **works as a model**?



Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Alice, and starts asking Alice questions. She finds out that Alice **studies in Italy**.

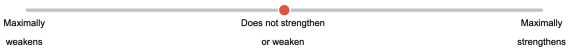
Consider the following **hypothesis (which could be true or false)** concerning Alice:

Alice has **brown hair**.

Now you are given a new piece of information concerning Alice:

Alice is **a Swedish woman**.

How does the new piece of information that Alice is **a Swedish woman** affect the hypothesis that Alice has **brown hair**?



Swiss man

No-context

Noah **likes making Italian desserts**.

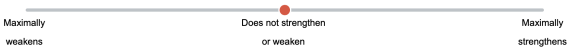
Consider the following **hypothesis (which could be true or false)** concerning Noah:

Noah **can ski**.

Now you are given a new piece of information concerning Noah:

Noah is **a Swiss man**.

How does the new piece of information that Noah is **a Swiss man** affect the hypothesis that Noah **can ski**?



Noah **likes making Italian desserts**.

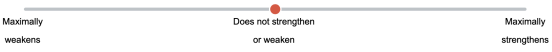
Consider the following **hypothesis (which could be true or false)** concerning Noah:

Noah **has a driving license**.

Now you are given a new piece of information concerning Noah:

Noah is **a Swiss man**.

How does the new piece of information that Noah is **a Swiss man** affect the hypothesis that Noah **has a driving license**?



Yes-context

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Noah, and starts asking Noah questions. She finds out that Noah **likes making Italian desserts**.

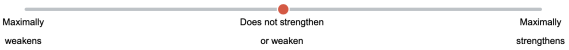
Consider the following **hypothesis (which could be true or false)** concerning Noah:

Noah **can ski**.

Now you are given a new piece of information concerning Noah:

Noah is **a Swiss man**.

How does the new piece of information that Noah is **a Swiss man** affect the hypothesis that Noah **can ski**?



Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Noah, and starts asking Noah questions. She finds out that Noah **likes making Italian desserts**.

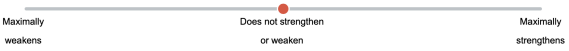
Consider the following **hypothesis (which could be true or false)** concerning Noah:

Noah **has a driving license**.

Now you are given a new piece of information concerning Noah:

Noah is **a Swiss man**.

How does the new piece of information that Noah is **a Swiss man** affect the hypothesis that Noah **has a driving license**?



Violinist

No-context

Dan is **an expert mountaineer**.

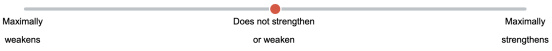
Consider the following **hypothesis (which could be true or false)** concerning Dan:

Dan **gives music lessons**.

Now you are given a new piece of information concerning Dan:

Dan has **a degree in violin performance**.

How does the new piece of information that Dan has **a degree in violin performance** affect the hypothesis that Dan **gives music lessons**?



Dan is **an expert mountaineer**.

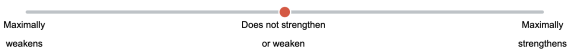
Consider the following **hypothesis (which could be true or false)** concerning Dan:

Dan **owns an umbrella**.

Now you are given a new piece of information concerning Dan:

Dan has **a degree in violin performance**.

How does the new piece of information that Dan has **a degree in violin performance** affect the hypothesis that Dan **owns an umbrella**?



Yes-context

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that Dan is **an expert mountaineer**.

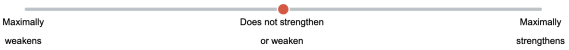
Consider the following **hypothesis (which could be true or false)** concerning Dan:

Dan **gives music lessons**.

Now you are given a new piece of information concerning Dan:

Dan has **a degree in violin performance**.

How does the new piece of information that Dan has **a degree in violin performance** affect the hypothesis that Dan **gives music lessons**?



Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that Dan is **an expert mountaineer**.

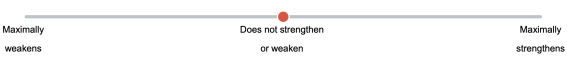
Consider the following **hypothesis (which could be true or false)** concerning Dan:

Dan **owns an umbrella**.

Now you are given a new piece of information concerning Dan:

Dan has **a degree in violin performance**.

How does the new piece of information that Dan has **a degree in violin performance** affect the hypothesis that Dan **owns an umbrella**?



A.3 Attention checks

In order to facilitate our research, we are interested in knowing certain facts about you. Specifically, we are interested in whether you actually take this time to read the directions; if not, then the data we collected based on your responses will be invalid. So, in order to demonstrate that you have read the instructions, please ignore the question on this page (i.e., don't answer it), and simply write "I have read the instructions" in the box labeled "Any comments or questions" on the next page. Thank you very much.

Have you attended university?

- No I have not
- Yes, but I didn't graduate
- Yes, and I graduated

Continue

Any comments or questions?

Continue

B Experiment 1 materials

Italian student

No-context

Consider the following text:

Carlo is an Italian undergraduate student.

*Which is more probable?**

- Carlo has **red hair** and **studied abroad in Barcelona in 2017**.
- Carlo has **red hair**.

Consider the following text:

Carlo is an Italian undergraduate student.

*Which is more probable?**

- Carlo has **red hair**.
- Carlo has **red hair** and **spent his summer holidays in America in 2017**.

Yes-context

Consider the following text:

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Carlo, and starts asking Carlo questions. She finds out that **Carlo is an Italian undergraduate student.**

*Which is more probable?**

- Carlo has **red hair** and **studied abroad in Barcelona in 2017**.
- Carlo has **red hair**.

Consider the following text:

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Carlo, and starts asking Carlo questions. She finds out that **Carlo is an Italian undergraduate student.**

*Which is more probable?**

- Carlo has **red hair** and **spent his summer holidays in America in 2017**.
- Carlo has **red hair**.

Swedish woman

No-context

Consider the following text:

Alice is a Swedish woman.

*Which is more probable?**

- Alice **studies in Italy** and **works as a model**.
- Alice **studies in Italy**.

Consider the following text:

Alice is a Swedish woman.

*Which is more probable?**

- Alice **studies in Italy** and has **brown hair**.
- Alice **studies in Italy**.

Yes-context

Consider the following text:

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Alice, and starts asking Alice questions. She finds out that **Alice is a Swedish woman.**

*Which is more probable?**

- Alice **studies in Italy**.
- Alice **studies in Italy** and **works as a model**.

Consider the following text:

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Alice, and starts asking Alice questions. She finds out that **Alice is a Swedish woman.**

*Which is more probable?**

- Alice **studies in Italy**.
- Alice **studies in Italy** and has **brown hair**.

Swiss man

No-context

Consider the following text:

Noah is a Swiss man.

*Which is more probable?**

- Noah **likes making Italian desserts**.
- Noah **likes making Italian desserts** and **can ski**.

Consider the following text:

Noah is a Swiss man.

*Which is more probable?**

- Noah **likes making Italian desserts** and **has a driving license**.
- Noah **likes making Italian desserts**.

Yes-context

Consider the following text:

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Noah, and starts asking Noah questions. She finds out that **Noah is a Swiss man.**

*Which is more probable?**

- Noah **likes making Italian desserts** and **can ski**.
- Noah **likes making Italian desserts**.

Consider the following text:

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Noah, and starts asking Noah questions. She finds out that **Noah is a Swiss man.**

*Which is more probable?**

- Noah **likes making Italian desserts**.
- Noah **likes making Italian desserts** and **has a driving license**.

Violinist

No-context

Consider the following text:

Dan has a degree in violin performance.

*Which is more probable?**

- Dan is **an expert mountaineer**.
- Dan is **an expert mountaineer** and **gives music lessons**.

Consider the following text:

Dan has a degree in violin performance.

*Which is more probable?**

- Dan is **an expert mountaineer**.
- Dan is **an expert mountaineer** and **owns an umbrella**.

Yes-context

Consider the following text:

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that **Dan has a degree in violin performance.**

*Which is more probable?**

- Dan is **an expert mountaineer** and **gives music lessons**.
- Dan is **an expert mountaineer**.

Consider the following text:

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that **Dan has a degree in violin performance.**

*Which is more probable?**

- Dan is **an expert mountaineer**.
- Dan is **an expert mountaineer** and **owns an umbrella**.

C Experiment 2 materials

Italian student

No-context

Carlo is an Italian undergraduate student.

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Carlo has **red hair**.*
- ➤ *Carlo has **red hair** and **studied abroad in Barcelona in 2017**.*
- ➤ *Carlo has **red hair** and **spent his summer holidays in America in 2017**.*

Yes-context

Maria is a consultant doing research for **an American travel agency**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Carlo, and starts asking Carlo questions. She finds out that **Carlo is an Italian undergraduate student.**

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Carlo has **red hair** and **spent his summer holidays in America in 2017**.*
- ➤ *Carlo has **red hair**.*
- ➤ *Carlo has **red hair** and **studied abroad in Barcelona in 2017**.*

Swedish woman

No-context

Alice is a Swedish woman.

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Alice **studies in Italy** and has **brown hair**.*
- ➤ *Alice **studies in Italy**.*
- ➤ *Alice **studies in Italy** and **works as a model**.*

Yes-context

Becky is a consultant doing research for **a shampoo company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Alice, and starts asking Alice questions. She finds out that **Alice is a Swedish woman.**

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Alice **studies in Italy** and has **brown hair**.*
- ➤ *Alice **studies in Italy**.*
- ➤ *Alice **studies in Italy** and **works as a model**.*

Swiss man

No-context

Noah is a Swiss man.

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Noah **likes making Italian desserts** and **can ski**.*
- ➤ *Noah **likes making Italian desserts**.*
- ➤ *Noah **likes making Italian desserts** and **has a driving license**.*

Yes-context

Natalie is a consultant doing research for **a car company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Noah, and starts asking Noah questions. She finds out that **Noah is a Swiss man.**

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Noah **likes making Italian desserts**.*
- ➤ *Noah **likes making Italian desserts** and **can ski**.*
- ➤ *Noah **likes making Italian desserts** and **has a driving license**.*

Violinist

No-context

Dan has a degree in violin performance.

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Dan is **an expert mountaineer**.*
- ➤ *Dan is **an expert mountaineer** and **owns an umbrella**.*
- ➤ *Dan is **an expert mountaineer** and **gives music lessons**.*

Yes-context

Adina is a consultant doing research for **an umbrella company**, trying to discover new target groups in Europe for the company to market to. She calls a randomly selected person, Dan, and starts asking Dan questions. She finds out that **Dan has a degree in violin performance.**

Please order the statements below from most probable (top) to least probable (bottom).

- ➤ *Dan is **an expert mountaineer** and **gives music lessons**.*
- ➤ *Dan is **an expert mountaineer**.*
- ➤ *Dan is **an expert mountaineer** and **owns an umbrella**.*