



A Linguistic Approach to the Conjunction Fallacy

Ariel Cohen

Ben-Gurion University of the Negev

arikc@bgu.ac.il

Joint work with:

*Ralph Hertwig, Lavi Wolf, Almog Simchon,
Keren Flash, and Hagar Golan*

Rationality, Probability, and Pragmatics—ZAS Berlin, 2016

The Conjunction Fallacy

Tversky and Kahneman (1983):

Linda is thirty-one 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 antinuclear demonstrations

The Conjunction Fallacy

Rank the following statements according to their probabilities:

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

The Conjunction Fallacy

- The vast majority of respondents:
- $P(1c) > P(1b)$
- But note that $1c = 1a \ \& \ 1b$
- And it is impossible for a conjunction to be more probable than one of its conjuncts!
- Necessarily, $P(A \& B) \leq P(A)$
- This is known as the **Conjunction Fallacy**

The Conjunction Fallacy

Remarkably, it was found that the conjunction fallacy is also committed by:

1. PhD students in statistics
2. Subjects who had the rule $P(A \& B) \leq P(A)$ explicitly explained to them.

Gould's homunculus

Stephen Jay Gould (1988):

all groups of subjects, sophisticated students who ought to understand logic and probability as well as folks off the street corner, rank the last statement as more probable than the second. (I am particularly fond of this example because I know that the third statement is least probable, yet a little homunculus in my head continues to jump up and down, shouting at me—"but she can't just be a bank teller; read the description.")

Representativeness

Tversky and Kahneman (1983):

- People do not really follow the rules of probability, but rather a *representativeness* heuristic.
- Linda is more representative of (more similar to) a feminist bank teller than a bank teller.

Representativeness

- This view paints a rather grim picture of human rationality.
- How come even probability experts commit this fallacy?
- Tversky & Kahneman: even **they** don't follow the rules of probability.
- But is this plausible?

Saving rationality?

- Many researchers feel uneasy with this claim.

→ Misinterpretation theories:

Subjects misinterpret the task as intended, and actually reason rationally.

- Politzer and Noveck (1991): “the task achieves its effect because the pragmatic demands of the task encourage subjects to reinterpret the presented materials before carrying out the task's required extensional analysis.”

Saving rationality?

- But various types of misinterpretation theory are claimed to have been disconfirmed experimentally.
- It seems that there is no choice: people are fundamentally irrational (Moro 2009).
- But maybe there is a choice after all...

The Role of Linguistics

- A theory of **mis**interpretation must rely on the notion of interpretation.
- Interpretation is determined by a combination of **linguistic** rules: phonological, syntactic, semantic, and pragmatic.
- Hertwig (2000): “*The Linda problem... is laden with the ambiguity of natural language.*”

The Role of Linguistics

- T&K found Conjunction Fallacy also with non-linguistic stimuli:
- Consider a regular six-sided die with four green faces and two red faces. The die will be rolled 20 times.

1. **RGRRR**

2. **GRGRRR**

3. **GRRRRR**

The Role of Linguistics

- Note:
- **GRGRRR = G & RGRRR**
- **$P(\text{GRGRRR}) \leq P(\text{RGRRR})$**
- But subjects judged:
- **$P(\text{GRGRRR}) > P(\text{RGRRR})$**
- A fallacy!

The Role of Linguistics

- But: in the Linda experiment, subjects commit the CF even when the conjunction rule is explicitly explained.
 - Crucially, with T&K's die, explaining the rule to the subjects does effectively eliminate the CF.
- With linguistic stimuli, subjects' reasoning is not fallacious, so they see no need to correct their judgment

Linguistics: The Domain of quantification

- There is a linguistic question that is relevant here—quantifier domains.
- (2) Mary always takes John to the movies
(Rooth 1985)
 - Doesn't mean that every second, 24 hours a day, 365 days a year, Mary takes John to the movies...
- The domain of quantification is restricted
- How is the domain of adverbial quantification determined?

Linguistics: The Domain of quantification

- Answer: focus
- Roughly, new information
- Focused elements are typically stressed

Linguistics: The Domain of quantification

(2) a. Mary always takes JOHN to the movies.

=Whenever Mary takes someone to the movies, Mary takes John to the movies

=**always**_e

($\exists x$ Mary takes x to the movies at e
Mary takes John to the movies at e)

Linguistics: The Domain of quantification

- b. MARY always takes John to the movies.

= Whenever someone takes John to the movies, Mary takes John to the movies

= **always_e**

($\exists x$ x takes John to the movies at e
Mary takes John to the movies at e)

Linguistics: The Domain of quantification

- Hence:
- Non-focused material goes to the domain (Rooth 1985; Partee 1991)
- The entire sentence goes to the scope

Linguistics: The Domain of quantification

- Cohen (1999): quantificational adverbs express conditional probabilities (see also Åquist *et al.* 1980).
- Roughly:
- **always**(ψ, ϕ) is true iff $P(\phi \mid \psi) = 1$
- **never**(ψ, ϕ) is true iff $P(\phi \mid \psi) = 0$
- **sometimes**(ψ, ϕ) is true iff $P(\phi \mid \psi) > 0$
- **usually**(ψ, ϕ) is true iff $P(\phi \mid \psi) > 0.5$

...

Linguistics: The Domain of quantification

- Focus affects the domain of Q-adverbs
- Q-adverbs express conditional probabilities: their domain is mapped onto the reference class.

→ Focus affects the reference class of conditional probabilities

- Does this solution to the linguistic problem also help with the

Conjunction Fallacy?

Focus and the Conjunction Fallacy

Cohen (1996, 1997):

- Consider a variant of the Linda sentences:

(3) Linda is a feminist bank teller.

(4) Linda is a bank teller.

- Conjunction is expressed by intersective modification

→ A Conjunction Fallacy is expected:

- $P(3) > P(4)$

Focus and the Conjunction Fallacy

- But Politzer and Noveck (1991):
“performance on Linda-type problems is, at least partly, due to... the manner in which items are presented.”
- Now note the effect of focus:
(5) a. Linda is a FEMINIST bank teller.
b. Linda is a feminist BANK TELLER.
- Intuitively, (5a) still seems highly probable, but (5b) much less so

Focus and the Conjunction Fallacy

- If focus affects the reference class, people do not interpret (5a) as the **un**conditional probability:
 - (6) $P(\text{Linda is a feminist bank teller})$
 - but rather as conditional probability:
 - (7) $P(\text{Linda is a feminist bank teller} \mid \text{Linda is a bank teller})$
 $> P(\text{Linda is a bank teller})$
- The Conjunction Fallacy

Focus and the Conjunction Fallacy

- In contrast: with (5b), people consider a different conditional probability:

(8) $P(\text{Linda is a feminist bank teller} \mid \text{Linda is a feminist})$

- This probability is not high

→ No Conjunction Fallacy

- These judgments do not violate the conjunction rule, are not irrational, and are, in fact, eminently plausible.

"Already refuted"?

- Tversky and Kahneman's representativeness heuristic cannot account for focus effects.
- Yet, Fiedler (1988):
- “the hypothesis was considered, and refuted, by Tversky and Kahneman (1983, p. 302) that subjects interpret the task in terms of conditional probabilities, judging $P(B|A)$ instead of $P(A\&B)$ ”

"Already refuted"?

- But this is wrong
- Tversky & Kahneman examined whether “the respondents interpreted the conjunction (A and B) as... a conditional statement (A if B).”
- They used a truth value judgment task, and the results correspond to the conjunction truth table (see also Tentori and Crupi, 2012)

"Already refuted"?

- But: nowhere do we deny that *Linda is a feminist bank teller* is a conjunction!
- So it is hardly surprising that a truth value judgment task confirms this.
- The difference from T & K only comes into play with a probability judgment task:
- Conditional probability rather than unconditional probability

The Conjunction Fallacy without overt focus

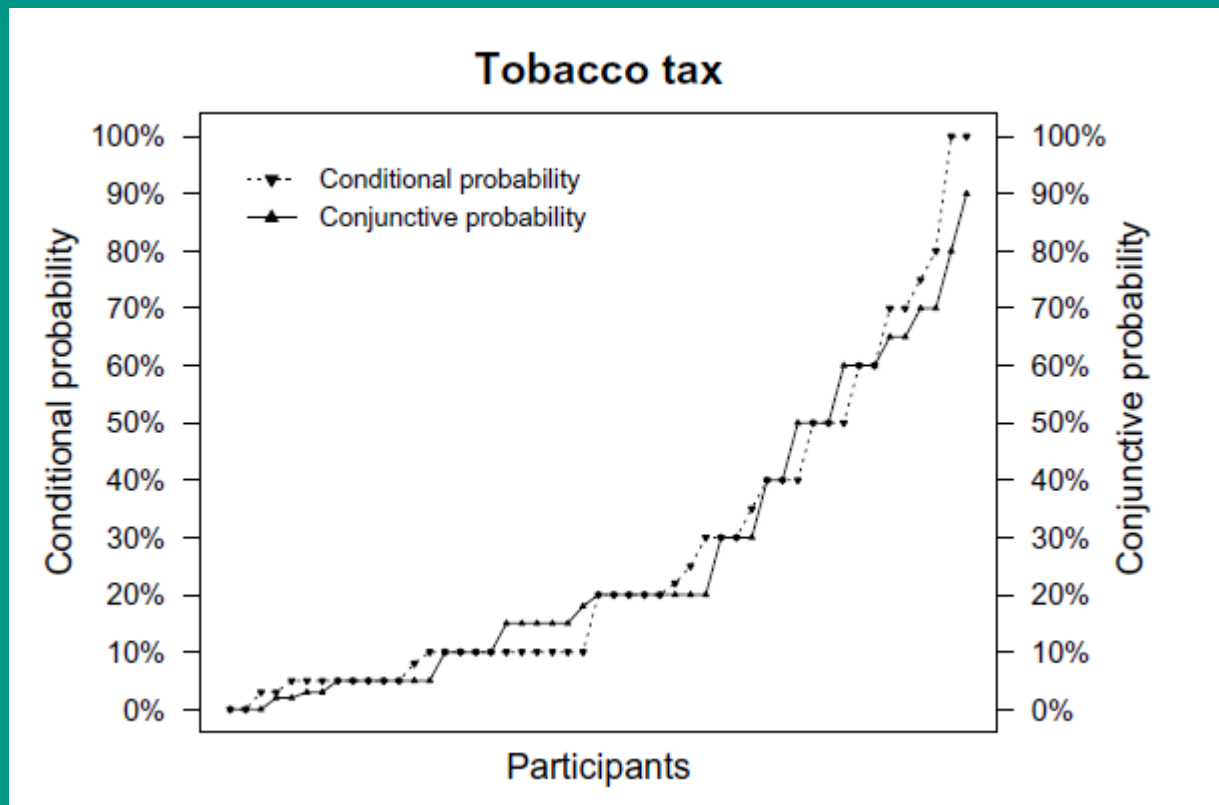
- But in the original experiment there is no overt indication of focus.
- Do subjects in this case also judge conditional, rather than unconditional, probabilities?

Evidence for Conditional Interpretation

- Hertwig *et al.* (2008) compare the judgments of two groups:
 1. Probability of a conjunction
 2. Conditional probability.

Evidence for Conditional Interpretation

Results—the judgments are virtually identical:



Back to Focus

- Without overt focus, subjects must choose the reference class themselves.
 - They choose where to mark focus:
 - (9) a. Linda is a FEMINIST bank teller.
(→ fallacy)
 - b. Linda is a feminist BANK TELLER
(→ no fallacy)
- Why do they assign focus as in (9a) rather than (9b)?

The Principle of Parsimony

- Crain *et al.* (1994): focus ambiguities.

(10) Dan only painted a chair.

(i) Narrow focus:

Dan only painted A CHAIR

= The only thing he painted was a chair

(ii) Wide focus:

Dan only PAINTED A CHAIR

= The only thing he did was paint a chair.

The Principle of Parsimony

- Crain *et al*: adult subjects prefer the narrow focus reading.
- Note that the narrow reading has more chances of being true:
- Suppose Dan painted a chair and also flew a kite.
 - (i) Narrow focus: true
 - (ii) Wide focus: false

The Principle of Parsimony

- Crain *et al.*: **The Principle of Parsimony**
- People prefer the reading that is true under most circumstances.
→ The most probable reading.

The Principle of Parsimony

- Marking focus on **feminist** results in an interpretation that is more probable than marking focus on **bank teller**.
- The Principle of Parsimony entails that this reading is preferred.
- **This** is the cause of the conjunction fallacy.

The Principle of Parsimony

- Crain *et al.*: children do **not** follow Parsimony—prefer wide focus.
- Prediction:
 - In Linda scenarios, children place focus over the entire conjunction
 - They judge the **un**conditional probability $P(\text{Linda is a feminist bank teller})$
 - No conjunction fallacy
- This is borne out (Davidson 1995)!

Interim Conclusions

- The conjunction fallacy is not really a fallacy: both children and adults behave rationally in their judgments.
- Adults follow the principle of parsimony:
 - They assign focus in a way that makes the interpretation more likely
 - They judge a conditional rather than unconditional probability
 - The result looks like a fallacy, but isn't

Experimental investigation

- This theory needs to be tested experimentally
- Our goals are to demonstrate that:
 1. Manipulation of focus as new information affects the fallacy
 2. Manipulation of focus in terms of prosody affects the fallacy
 3. Without overt focus, subjects assign focus in a way that leads to the fallacy

Experimental investigation

- This is work in progress;
- So far, we have established stage 1: manipulation of focus in terms of information structure.

Focus and order

- New information leads to focus
- The more recently an item is presented, the newer it is
 - Other things being equal, the more recently an item is presented, the more focused it is.
 - The probability will be conditional on the property that is presented **first**

Focus and order

- We predict an *order effect*:
- If the unexpected property (**bank teller**) is presented first, and the expected property (**feminist**) is presented last, the conditional probability will be:
 - $P(\text{feminist \& bank teller} \mid \text{bank teller})$
→ Much higher than $P(\text{bank teller})$

Focus and order

- If the expected property (**feminist**) is presented first, and the unexpected property (**bank teller**) is presented last, the conditional probability will be:
 - $P(\text{feminist \& bank teller} \mid \text{feminist})$
 - Roughly the same as $P(\text{bank teller})$

Focus and order

- Surprisingly: “There are no experimental data on order effects in conjunction fallacy experiments, when the judgments are performed in different orders.” (Franco 2009, 421)
- Charness *et al.* (2009) counterbalanced the order of presentation.
- Significantly: they got a much reduced fallacy rate: 58% vs. T&K’s 85%.

Experiment 1: Hebrew replication

- The language of our experiments was Hebrew
- We first replicated Tversky & Kahneman's results in Hebrew, to validate the Hebrew items of Linda and Bill.
- 27 native Hebrew Speakers
- Ages: 21-28(1st year undergrads)
- The experiment was done in class

Experiment 1: Hebrew replication

■ Results:

Scenario	Linda	Bill
Fallacy Rate	79% (23/27)	75% (22/27)

- A significant conjunction fallacy effect, comparable to T&K

Experiments 2-4: General design

- Subjects who were familiar with the Conjunction Fallacy, or who were non-native Hebrew speakers, were taken out of the analysis.
- Subjects estimated the probabilities of the unexpected property, the expected property, the conjunction, and a filler.

Experiments 2-4: General design

- As a measure of the effect of order on the fallacy, we analyzed the difference between the probability of the conjunction and the probability of the unexpected property (bank teller)
- $P(E\&UE) - P(UE)$

Experiment 2: within subject design

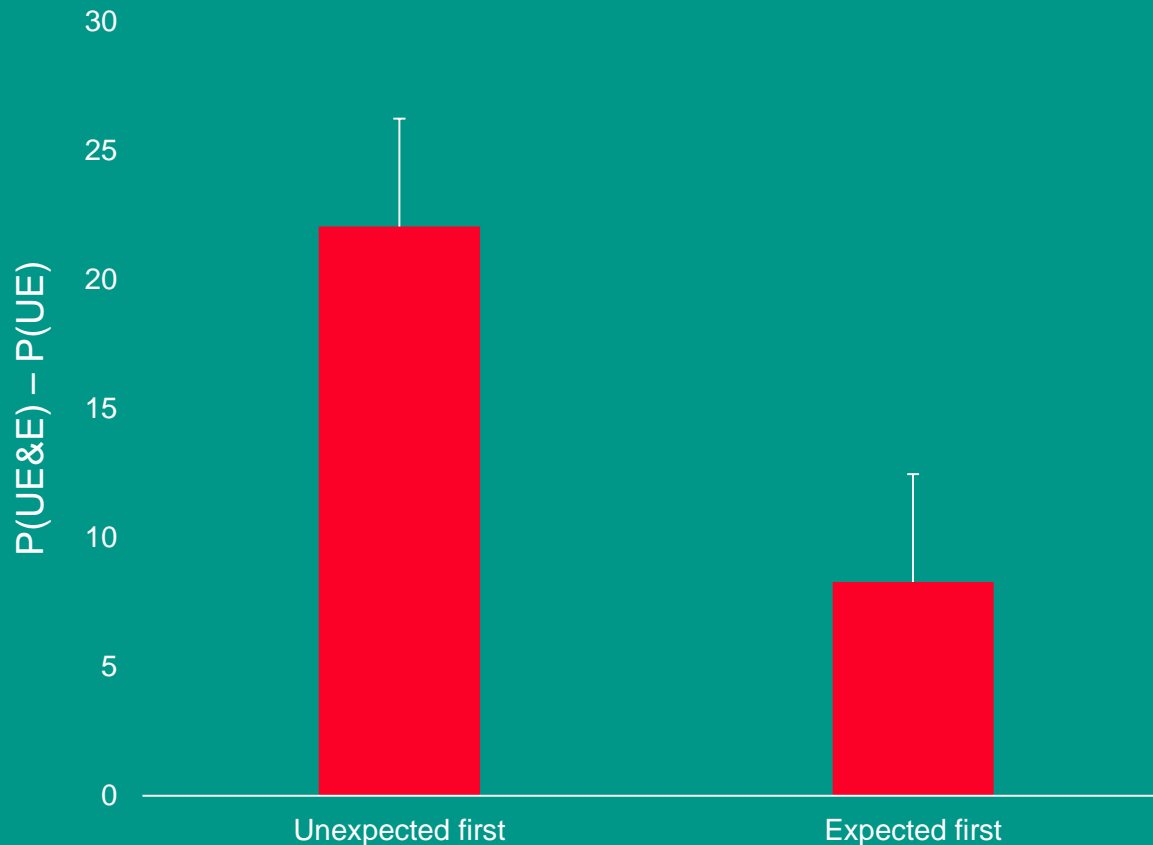
- Two conditions:
- Condition a: For Linda, the order was:
 1. Unexpected: *Linda is a bank teller*
 2. Conjunction: *Linda is a bank teller and is active in the feminist movement*
 3. Filler: *Linda is a teacher*
 4. Expected: *Linda is active in the feminist movement*
- For Bill, reversed order: *expected first*

Experiment 2: within subject design

- Condition b: switch the order of the items between Linda and Bill
- So:
 - Condition a: Linda **UE-E**, Bill **E-UE**
 - Condition b: Linda **E-UE**, Bill **UE-E**
- The experiment was conducted online using the Qualtrics platform

Experiment 2: within subject design

- Results: a significant effect



$p=.004$
 $N=34$

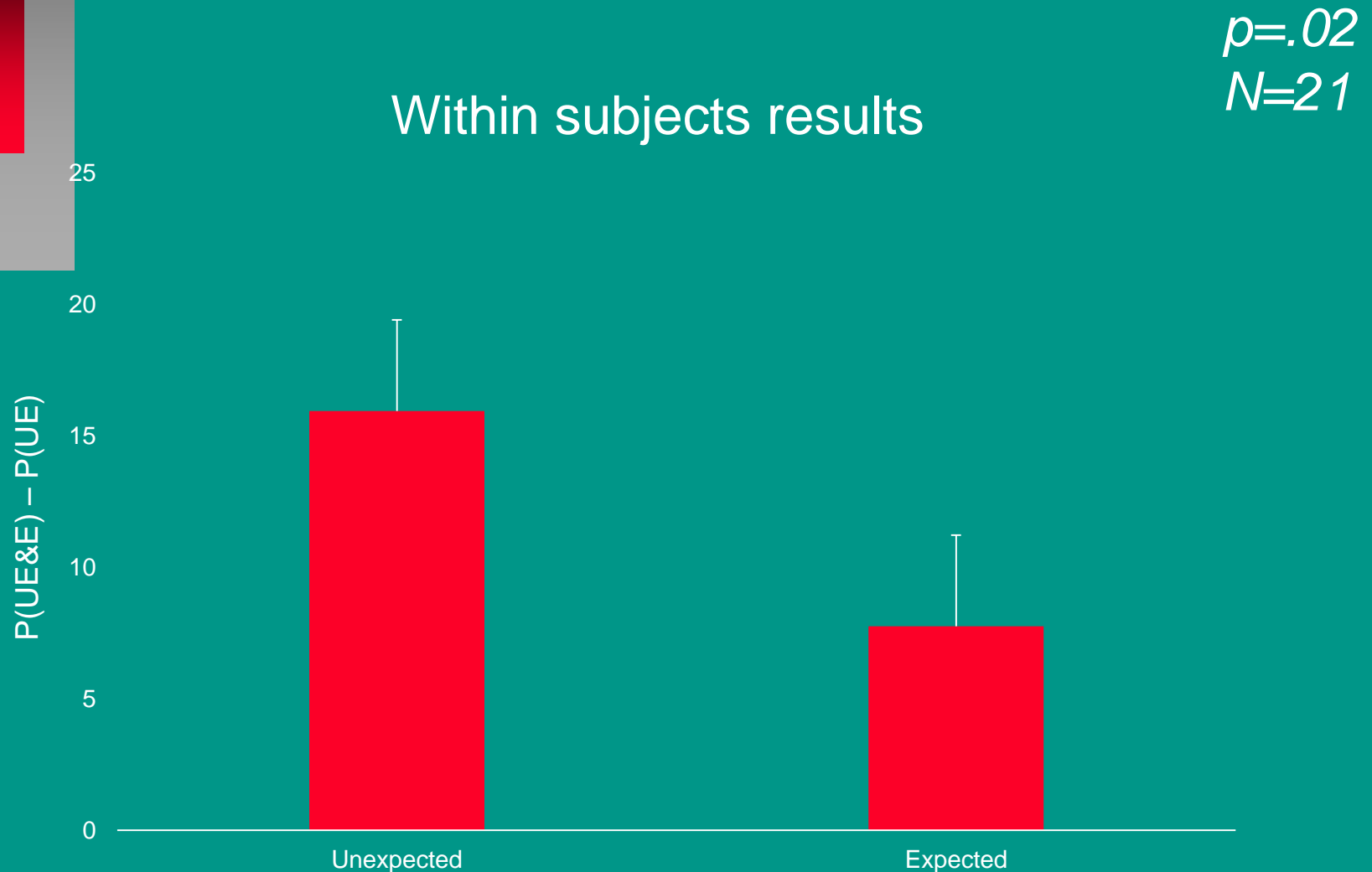
Experiment 3: within + between subjects

- We also checked for between-subjects effects
- To properly do this we needed to check all 4 conditions:
 - a. Linda: UE-E, Bill: E-UE
 - b. Linda: E-UE, Bill: UE-E
 - c. Linda: UE-E, Bill UE-E
 - d. Linda: E-UE, Bill E-UE
- The experiment was conducted in class

Experiment 3: within + between subjects

- We first verified within-subjects results:
- Conditions **a** and **b** only.
- (the ordering for Linda is the opposite of the ordering for Bill)

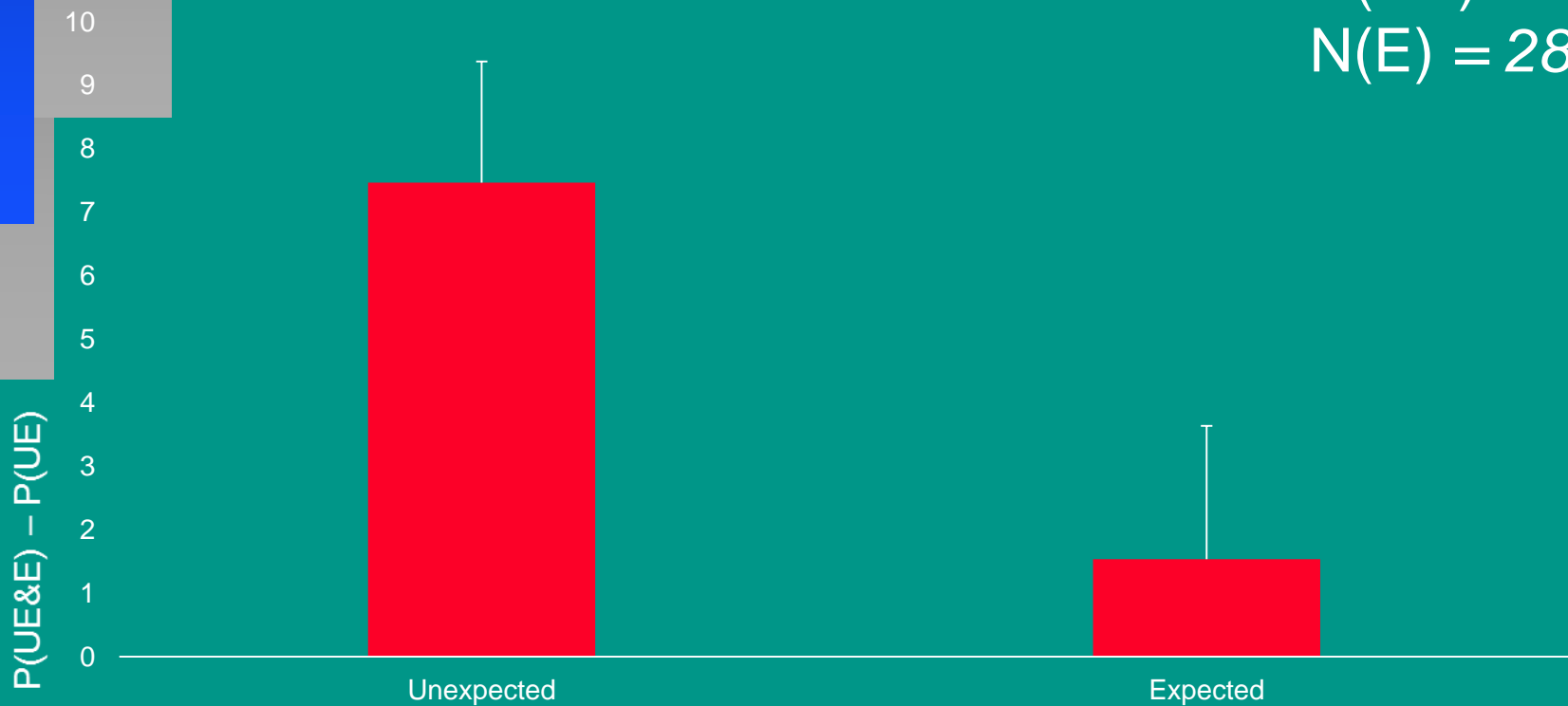
Experiment 3: within + between subjects



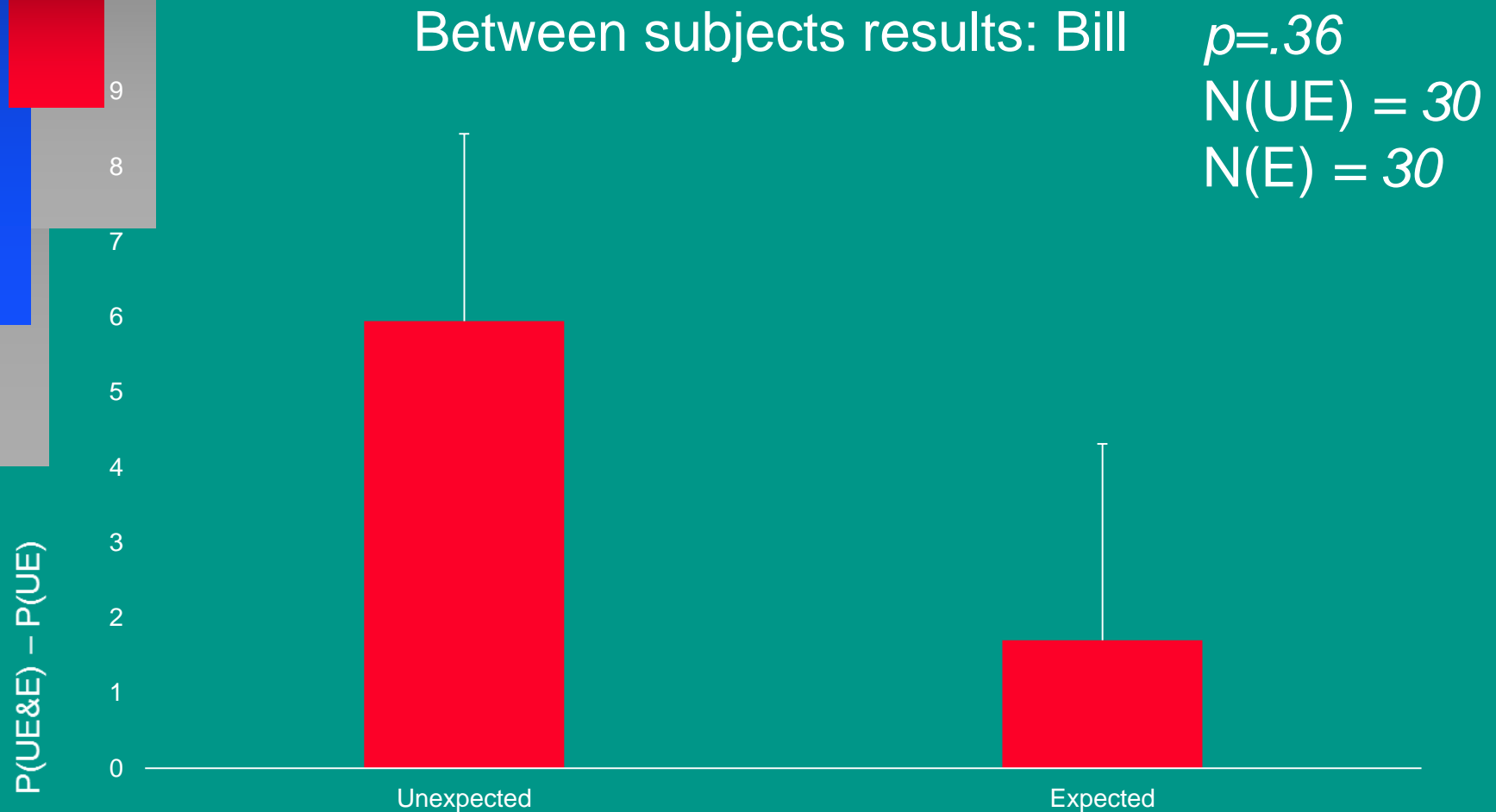
Experiment 3: within + between subjects

Between subjects results: Linda

$p = .03$
 $N(UE) = 32$
 $N(E) = 28$



Experiment 3: within + between subjects



Experiment 3: within + between subjects

- Within-subjects results are again highly significant
- Between-subjects results are significant for Linda
- Similar trend is shown for Bill, but the results are not statistically significant
- But there is another problem...

Is there an order effect?

- In the previous two experiments, the order of presentation was:
 1. Property (E or UE)
 2. Conjunction
 3. Filler
 4. Property (UE or E)
- Maybe this is not really an order effect, since subjects judge the conjunction before they see the second property?

Experiment 4: between subjects

- We now used a different order:

1. Property (E or UE)
2. Filler
3. Property (UE or E)
4. Conjunction

→ The conjunction is judged after both properties are judged

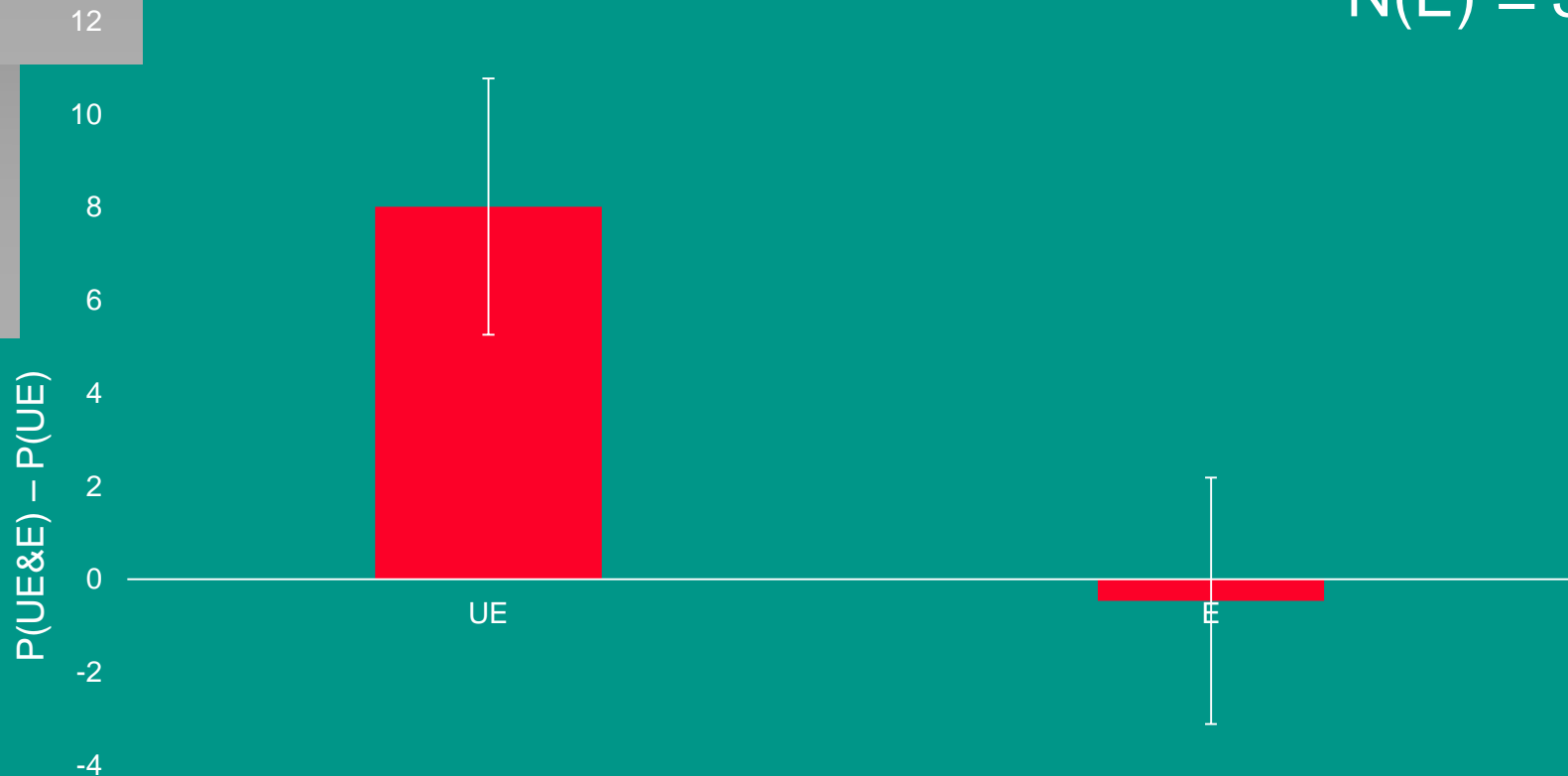
Experiment 4: between subjects

- The experiment was run online using the Qualtrics platform
- More subjects: 102
- Each subject saw only Linda or Bill

Experiment 4: between subjects

Results: between subjects
(Linda + Bill)

$p = .028$
 $N(\text{UE}) = 49$
 $N(\text{E}) = 53$



Non-native speakers

- Recall that in all the experiments, non-native speakers were not analyzed
- When we did look at their results, the data were quite noisy, and the effect was not demonstrated (though not enough data for a statistical analysis)
- This provides additional suggestive evidence that the conjunction fallacy is really a linguistic phenomenon.

Conclusion

- In four experiments, we have shown that manipulating focus, by changing the order of items, has a significant effect on the conjunction fallacy
- This effect has been shown across experimental settings (classroom and Qualtrics) and patterns of presentations, and is dependent on language fluency

Conclusion

- This effect is unexpected from the point of view of T&K's theory
 - But it follows directly from the theory according to which focus affects the reference class of probability judgments
- The Conjunction Fallacy is a linguistic phenomenon; it ought to be analyzed by linguistic means

Future work

- Investigate manipulation of focus by auditory means—intonation
- Provide direct evidence for assignment of focus according to Parsimony:
 - Have subjects read the items aloud, and provide an acoustic analysis
 - Eye tracking
 - ERP