The rationality of inferring causation from correlational language

Daniel Lassiter (dan.lassiter@ed.ac.uk)
Linguistics & English Language, University of Edinburgh

Michael Franke (michael.franke@uni-tuebingen.de)
Department of Linguistics, University of Tübingen

Abstract

Do people infer causal information from non-causal language? If they do, should they? To address these questions, we build on recent work showing that participants make asymmetric causal inferences from apparently symmetric correlational statements (e.g., "*A* is associated with *B*"). In Experiment 1, we dig further into the nature of such interpretive preferences—what we call "PACE effects"—in light of models and results from theoretical and experimental pragmatics and psycholinguistics, uncovering several linguistic factors that influence them. The existence of PACE effects does not yet show that associational language leads to causal implicatures strong enough to influence action choice in practical decision contexts. To show that it can, Experiment 2 offers new evidence from an experiment that explicitly compares the effects of causal vs. correlational claims on decision-making. Our results support a general picture in which causal inferences from correlation language are an intricate, but emimently rational, feature of natural language understanding.

Keywords: causal inference, correlation, natural language pragmatics, rationality

Introduction

There is a well-known mantra that *correlation does not imply causation*. Statisticians and scientists have long decried the human tendency to ignore this rule (Adams et al., 2017; Huff, 1954; Seifert et al., 2022). On one popular theory, deep-seated cognitive biases cause us to hallucinate causation where only statistical conclusions are appropriate (Ariely & Jones, 2008; Tversky & Kahneman, 1982).

From another perspective, this conclusion may seem too hasty. Causal learning is a core aspect of human cognition, crucial to reasoning, decision-making, categorization, and many other tasks (e.g. Danks, 2014; Rehder, 2017; Sloman, 2005). Given the importance of causal learning for so many other key areas of our cognition, it would be surprising if humans could not maintain the basic distinction between correlation and causation: a person unable to do this would have difficulty tying their shoes. Indeed, many theorists have argued that humans are able to combine statistical evidence with prior beliefs to draw sensible inferences about the structure and strength of causal relations (Gopnik et al., 2004; Griffiths & Tenenbaum, 2009).

The mere fact that evidence of correlation influences causal inference is not surprising or normatively problematic. By the common cause principle (Reichenbach, 1956), every non-accidental correlation is due to some sort of causal relationship. The existence of a correlation between *A* and *B* at least allows us to rule out a model in which these variables are causally unconnected. But this reasoning does not tell us whether *A* causes *B*, *B* causes *A*, some third variable causes both, or some more complex relationship obtains. From a Bayesian perspective, the point appears like this: once we have ruled out independence, any asymmetries in the posterior probabilities of the remaining causal models must have been present already in the prior.

The matter becomes more complex when we turn to the linguistic expression of correlation. If a speaker chooses to use an expression that *literally* conveys that that *A* and *B* are correlated, do we take this choice to implicate a causal relation? If so, is this inference rational?

Recent experimental work attempted to address the first question (Gershman & Ullman, 2023) and showed that participants who are presented with a simple correlational claim, with no background information to inform priors, do not treat the various possible causal models symmetrically. Concretely, the study conducted a forced-choice experiment in which participants were presented with a correlational claim such as "*A* is associated with *B*" and were asked to choose either "*A* causes *B*" or "*B* causes *A*". Since participants knew nothing about *A* and *B*, the reasoning above would lead us to expect purely random choice—50% for each option. Instead, they found that the majority of participants selected "*B* causes *A*". Strikingly, this pattern was reversed in a second experiment which used slightly different prompts, varying in details of their semantic content.

The experiments of Gershman & Ullman (2023) reveal an interesting new phenomenon that we call **PACE effects** (for *Preferences in Assignments of Cause and Effect*). In a series of studies partially replicating and extending Gershman & Ullman's results, we show that PACE effects are influenced by grammatical, semantic, and discourse-level features of language use. We also present a second set of experiments, embedded in a practical decision-making context, that show that the use of associational language can generate causal inferences strong enough to influence action choice in concrete decision contexts. These experiments indicate that causal inferences from non-causal language are not merely an artifact of the forced-choice format of Gershman & Ullman 2023 and our Exp.1: a speaker's choice to use non-causal language can lead to actionable changes in causal beliefs.

The results observed in previous work and in our experi-

ments can be understood in terms of rational pragmatic inference: a listener who has reason to think that the speaker is trying to convey causal information can use the form and meaning of the chosen sentence to infer what kind of causal information the speaker considers relevant. This explains why PACE effects are sensitive to linguistic factors, and how it can be rational to infer causation from correlational language.

Inferring implicit questions

Current pragmatic theory has makes considerable use of the classic hypothesis that every declarative sentence is interpreted as the answer to a question, which may be explicitly posed or pragmatically inferred (Carlson, 1983; Clark, 1979; Ginzburg, 1995; van Kuppevelt, 1995; Roberts, 2012). "Question under Discussion" or "QUD theory" has been successful in explaining aspects of linguistic form and interpretation (Beaver et al., 2017). Crucially, the choice of QUD is sensitive to overt linguistic form. For example, a clefted sentence like "It's COOKIES that I love" can be respond to "What do you love?", but not to "Who loves cookies?" or "What time is it?" (Clark & Haviland, 1977). As a result, when a statement does not respond to an overt question, the listener's search for an implicit QUD is tightly constrained.

QUD theory has proved useful in psycholinguistic research, for instance by underwriting the discourse relations that are crucial for maintaining coherence and motivating many discourse-level phenomena (Kehler & Rohde, 2017). Bayesian models of language understanding have also made crucial use of QUDs to account for ways that context, plausibility, and speaker goals interact to influence pragmatic interpretation (Kao et al., 2014; Hawkins et al., 2015; Sumers et al., 2023). In these models, an (implicit or explicit) QUD determines what is deemed **relevant**, and relevance in turn influences pragmatic inference in a variety of ways. For instance, Kao et al. (2014) construct an RSA model (Frank & Goodman, 2012) in which listeners attempt to jointly infer the QUD and the intended message, and the inferred QUD influences the listener's model of the speaker's production protocol. They show that integrating QUD inference into RSA can account for whether large numerical estimates ("It cost a thousand dollars") are interpreted literally (QUD "How much did it cost?") or as hyperbolic expressions of affect (QUD "How do you feel about it?").

As Clark (1979) discusses, we frequently use QUD inference to make conversation more efficient by responding not to a question that a speaker has explicitly posed, but to a related question that we think practically relevant. This is a special case of bridging (Clark, 1977). Suppose a customer asks "Can I get a bottle of Kraken rum for $8?" and the clerk responds "You can get a bottle of Black Seal rum for $8". The success of this conversational exchange depends on the speaker and listener's ability to coordinate on some fairly complex inferences. The clerk's failure to respond directly to the question, far from being uncooperative, functions to improve the efficiency of information exchange. Inferring that the customer wants to buy rum and has $8 to spend, the clerk skips over the intermediate answer and likely follow-up question, leaving the listener to work out what is missing:

- **Customer**: Can I get a bottle of Kraken for $8?
- Implicit: (**Clerk**: No, you can't.)
- Implicit: (**Customer**: What kind of rum can I get for $8?)
- **Clerk**: You can get a bottle of Black Seal for $8.

Inferring causal questions

We hypothesize that the PACE effects documented by Gershman & Ullman (2023) are related to QUD inference of the type that Clark (1979) and Kao et al. (2014) discuss. To illustrate, consider the piece of public messaging in Fig. 1, which is at the time of writing posted on elevator doors throughout the University of Edinburgh. The sign aims to promote

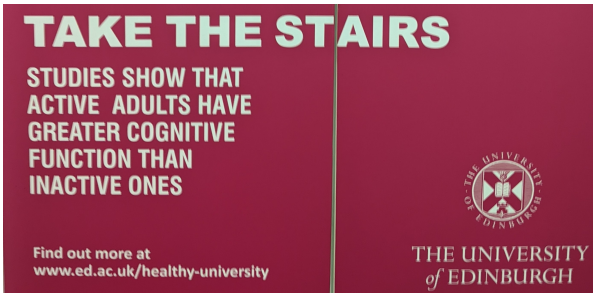


Figure 1: Correlational language used to convey existence and direction of a causal relationship.

a certain behavioral modification ("Take the stairs") by providing summary evidence of a correlation. Of course, the fact that "active adults have greater cognitive function" could be explained in many ways that would not support the causal efficacy of the recommendation to take the stairs. Nevertheless, the intended message—that physical activity **causes** improved cognitive function—comes across clearly.

How do we as interpreters of this signage effortlessly hone in on the intended causal relationship—inferring this message rather than some other plausible causal story? Our guiding hypothesis extends Clark's (1979) analysis of the liquor store conversation. The sign recommends taking the stairs. An astute listener would then respond "Why should I do that?". This QUD casts the next sentence as an explanation of why one should comply with the previous sentence's recommendation (cf. Hobbs, 1979; Asher & Lascarides, 2003; Kehler & Rohde, 2017). While the correlational information in this sentence is in principle compatible with models where **physical activity** $\Rightarrow$ **cognitive function**, or where **cognitive function** $\Rightarrow$ **physical activity**. However, only the former model would actually justify the author's recommendation to take the stairs. Assuming the writer is attempting to efficiently convey information that would allow them to accomplish the stated persuasive goal, a listener can conclude that the evidence adduced suggests a causal relationship that does in fact

support this goal. The writer's choice to convey this message indirectly can be seen as a device to improve efficiency. Because the intended causal claim can be recovered pragmatically, citing a correlation allows the author to compress both the causal claim and its evidential basis into a single statement.

Experiment 1: PACE effects

To test the hypothesis just described, we adapted Gershman & Ullman's (2023) experimental design in three studies manipulating the form and content of correlational statements.

Experiment 1a: Grammatical manipulation

Linguistic research has shown that every sentence is associated with a **topic** determining what the sentence is intuitively "about", with various syntactic and pragmatic effects (van Kuppevelt, 1995; Erteschik-Shir, 2007; Strawson, 1964). Crucially, the choice of topic places pragmatic constraints on the QUD, because the topic must be Given in the discourse and will typically be mentioned in the QUD (Clark & Haviland, 1977; Vallduví, 2016). These discourse-level features of interpretation are reflected in the form of sentences as well. We tend to place to place Given before New information, and in particular grammatical subjects are default topics in English (Clark & Haviland, 1977). For instance, (1) is more naturally associated with the QUD in (2a) than the one in (2b). This is because, absent further context, we tend to treat the subject phrase "The old judge" as the sentence topic in (1).

(1) The old judge took the bribe.

(2) a. What did the old judge do?
    b. Who took the bribe?

We hypothesized that the asymmetrical causal inferences noted Gershman & Ullman 2023 were due to the asymmetrical grammatical form of their stimuli. Absent further context, participants tend to read "*A* is associated with *B*" as being *about* the subject *A*, and so to infer a QUD that involves this item. For instance, if participants inferred a QUD along the lines of "Why does *A* happen?" or "What is the cause of *A*?", then a simple bridging inference would be enough to get from "*A* is associated with *B*" to "*A* is caused by *B*".

This reasoning suggests a simple manipulation: since the associational claim "*A* is associated with *B*" can equally be rendered as "*A* and *B* are associated", we modified Exp.1 of Gershman & Ullman 2023 to explicitly compare these semantically identical but grammatically different sentences.

Participants. We recruited 203 participants via Prolific (English native speaker, at least 10 previous studies, at least 90% approval rate). This and all other experiments reported in this paper were conducted on a website built using Magpie (Franke et al., 2021). Participants were paid 0.30 British pounds for an estimated 2 minutes of work (median of actual time taken was close to one minute for all three experiments).

Design, materials & procedure. This was a one-shot forced-choice experiment whose design followed Gershman & Ullman 2023 closely. We randomly presented participants with either the asymmetric "*A* is associated with *B*", or the symmetric "*A* and *B* are associated". The variables *A* and *B* were randomly filled with nonsense words used by Gershman & Ullman (2023) in their experiment 1: "Pneuben", 'Themaglin'", "Rebosen", "Denoden", "Agoriv","Flembers", and "Ceflar". Participants chose between "*A* causes *B*" and "*B* causes *A*", with the order of choices randomized. Once they had made a choice, they filled in a brief demographic questionnaire and the experiment ended.

Predictions. We expected that the asymmetric form would replicate the main result Gershman & Ullman 2023, Exp.1: the preferred interpretation would be that the second item causes the first. We predicted that participants would respond randomly in the symmetric condition, because the entire subject noun phrase "*A* and *B*" is topical. Sentence form thus does not provide clues suggesting that one item is Given, or that the sentence is "about" one rather than the other.

Results. The results in the asymmetric condition replicated Gershman & Ullman 2023: a majority (68%) chose the response option where the second-mentioned item caused the first ("*B* causes *A*"). This was significantly different from the chance level of .5 ($p < .001$ in a two-sided binomial test with $N = 94, k = 64$). By contrast, in the symmetric condition only 30% chose this response. This was also significantly different from chance ($p < 10^{-4}$ with $N = 109, k = 33$).

Discussion. The results are consistent with our hypothesis that PACE effects are conditioned by to discourse-level factors that relate to grammatical form. However, we did not anticipate the preference for "first-causes-second" in the symmetric condition. This result suggests that further factors involving sentence form that also play a role. In retrospect, this preference may be due to the Iconic Sequencing constraint on binomial ("*A* and *B*") constructions, motivated by Benor & Levy (2006). Such constructions are far more likely to put cause before effect than vice versa (e.g., "principal and interest" vs. the odd-sounding "interest and principal"). Iconic sequencing was in fact the "strongest and most frequently active" constraint in Benor & Levy's extensive corpus investigation (see p.252). It seems plausible that, lacking further clues from content or sentence form, participants' judgments were influenced in the symmetric condition by their knowledge of this general feature of binomial constructions; but further investigation would be needed to corroborate this suggestion.

Experiment 1b: Constrastive topic manipulation

To investigate our hypothesis that effects of linguistic form are mediated by discourse factors (e.g., QUD and topicality), we constructed a variant of the asymmetric condition in Exp.1 with an overt information-structural manipulation. Specifically, we tested the effects of a leading "As for ..." construction, which explicitly marks out the object of "for" as a contrastive topic (e.g. Büring, 2003).

Participants. We recruited 221 participants via Prolific, with the same conditions and payment as in Exp.1a.

Design, materials & procedure. The design was identical

to Exp.1a, except that we only tested the asymmetric frame, with the "as for" manipulation independent of grammatical role. Participants saw one of the following frames:

(3) a. "As for *A*, it is associated with *B*."
b. "As for *B*, *A* is associated with it."

Predictions. Based on our hunch that the relationship between grammatical form and PACE is mediated by discourse factors, we predicted that participants would not be sensitive to the subject/non-subject distinction with an overt contrastive topic. If so, the "as for" topic would be chosen as the cause at the same rate regardless of grammatical subject.

Results and discussion. Participants who saw (3a) chose "*B* causes *A*" at exactly the same rate (68%) as in the asymmetric condition of Exp.1a. Those who saw (3b) chose this response somewhat less often (54%). These results were marginally different ($\chi^2(1) = 3.5889$, $p = .058$). The study provided weak support for our hypothesis that information structure conditions PACE effects. It did not support our hypothesis that subjecthood is irrelevant once we control for information structure.

Experiment 1c: Adding category information

Inferring QUDs depends crucially on semantic and pragmatic plausibility. In Kao et al.'s (2014) RSA model of QUD inference, listeners assume that speakers are trying to be informative. This means that listeners should expect speakers to select and respond to QUDs whose answers are not obvious or already in common ground. We thus expected that providing category information to participants should influence their responses by influencing the expected informativity of different QUDs. Experiment 1c was intended to test this prediction.

Participants. We recruited 413 participants via Prolific, with the same conditions and payment as in Exps.1a and 1b.

Design, materials & procedure. The design was identical to Exps. 1a and 1b except for the stimuli. Participants saw one of the following at random:

(4) a. "The drug *A* is associated with *B*."
b. "*A* is associated with the drug *B*."
c. "The disease *A* is associated with *B*."
d. "*A* is associated with the disease *B*."

Predictions. On the QUD informativity hypothesis, when *B* is topical, a question about the causes of *B* should be dispreferred if the answer is obvious. In this case, a QUD about *B*'s effects should be preferred. We expected participants to reason as follows. The cause of taking a drug is obvious: the person who took it chose to do so. In contrast, the effects of a drug are often uncertain and of great interest. By contrast, both causes and effects of a disease are often uncertain and of interest. So, we expected participants to be more evenly split in PACE effects when they knew that one of the items was a disease. We expected that these effects would modulate the grammatical/information-structural effects discussed above.

Results. When participants were told that an item was a drug, they overwhelmingly chose that item as the cause: 88% for (4a), and 90% for (4b). When told that one item was a disease, the responses were more mixed and were strongly influenced by grammatical form. Participants who saw (4d) were evenly split in whether to treat the disease as cause or effect (48% "*A* causes *B*"). Those who saw (4c) chose "*A* causes *B*" 83% of the time. These responses differed significantly ($\chi^2(1) = 22.3$, $p < 10^{-5}$).

Discussion. The effects of linguistic form were strongly modulated when participants were able to map the content onto even skeletal causal knowledge. When participants were told that one item was a drug, the preference for informative QUDs overwhelmed the effects of grammatical position. The information that one item was a disease modulated, but did not eliminate, the effect of grammatical position. As compared to the asymmetric condition of Exp.1a, responses to both (4c) and (4c) were biased toward "*A* causes *B*", but much more so when the disease was the grammatical subject.

Exp.1 general discussion

The studies reported in Exp.1 explored several manipulations of linguistic form and content, in order to get a clearer picture of the source and nature of PACE effects. The hypotheses investigated were inspired by work in theoretical and experimental/computational pragmatics, particularly the idea that listeners enrich the interpretation of an out-of-context sentence by attempting to work out what question the speaker is trying to address—in other words, what the sentence is *for* in the conversation. We found that grammatical position and semantic content exert a large influence on PACE effects. We also uncovered possible evidence for a mediating role of an information-structural manipulation (constrastive topic) and of a statistical preference for cause before effect in binomial ("*A* and *B*") constructions. All of these factors deserve further attention. More broadly, the studies in Exp.1 suggest that PACE effects, far from being arbitrary, are highly sensitive to linguistic factors. With further investigation, there is good reason to believe that their precise nature will turn out to be attributable to the interaction of complex, but eminently rational, aspects of language understanding.

Experiment 2: Causal implicature

The previous experiments demonstrated systematic PACE effects for language expressing correlation, but PACE effects, as such, do not yet provide strong evidence for *practically relevant* causal implicatures. PACE effects show that certain constructions make one causal direction between two variables more likely than the other, but that does not necessarily imply that the degree of belief in causality as such is increased. In other words, PACE effects are in principle compatible with belief changes that are not *bona fide* examples of what we would want to address as causal implicatures, i.e., practically relevant increases in belief in a causal relationship.We therefore ran two experiments to test whether a statement like "*A* is associated with *B*", can actually convey

or suggest a causal connection to an extent that would affect practical decision making. The only difference between the studies is the exclusion or inclusion of additional information in the context description (see below).

Participants. We recruited 200 participants via Prolific (English Native, at least 10 previous studies, at least 90% approval rate). For technical reasons, N=198 data sets were received for Exp. 2a, and N=203 for Exp. 2b. Participants took an average of 2.69 minutes for Exp. 2a, and 2.69 minutes for Exp. 2b. They were paid 0.40 British Pounds.

Design, materials & procedure. The experiments are one-trial forced-choice designs with four between-subject conditions. Each subject supplied one data point for the critical trials. The four conditions differed only with respect to the sentence in bold in Fig. 2.

- **association:** "a high yield of xeliherb is associated with the presence of another plant called ralocrop."
- **intervention:** a high yield of xeliherb was obtained whenever another plant called ralocrop was cultivated as well.
- **commonCause:** a high yield of xeliherb is associated with the presence of another plant called ralocrop. But they also found that ralocrop only grows on particularly fertile grounds.
- **deniedCausation:** a high yield of xeliherb is associated with the presence of another plant called ralocrop. But there is no evidence that the cultivation of ralocrop causes a better yield of xeliherb.

Predictions. The condition of main interest is *association*. The other conditions function as reference categories. We expect different degrees of belief in the relevant causal relation (ralocrop causes a high yield of xeliherb) in different conditions. The auxiliary assumption is that participants with a greater degree of belief in the relevant causal relation should be more inclined to select the "both" option. We thus expect higher choice rates of the "both" option in the *association* condition than in the *deniedCausation* condition. The latter condition is intended to provide a lower bound on the strength of causal implicature. An upper bound on causal implicature strength is provided by the *intervention* condition, which has the highest rates of the "both" choice. Since Exp. 2b additionally makes more salient a potential cost, we expect that choice rates for the "both" option are generally lower.

Results. Fig. 3 shows the proportion of participants who chose "both". Impressionistically, participants respond with a choice option indicating an increased degree of belief in a causal relation more often for the *association* condition than for the *deniedCausation* condition. However, in a Bayesian logistic regression model this contrast was only credible for Exp. 2b. The posterior probability for contrast *deniedCausation* $<$ *association* was 0.96 with 95% credible interval of difference [-0.11; 2.08] for Exp. 2a, and 0.99 ([0.26; 1.98]) for Exp.2b. Moreover, the contrast *association* $<$ *intervention* is not credible in either experiment (0.76, [-0.82; 1.95] for Exp. 2a, and 0.52, [-0.92; 0.92] for Exp. 2b). The full results of the regression analyses are in Table 1.

comparison	Exp. 2a	Exp. 2b
denC $<$ comC	0.87, [-0.36 ; 1.46]	0.98, [0.13 ;1.68]
comC $<$ ass	0.78, [-0.67 ; 1.58]	0.68, [-0.64 ;1.08]
ass $<$ inter	0.77, [-0.89 ; 1.85]	0.52, [-0.90 ;0.93]
denC $<$ ass	0.96, [-0.11 ; 2.08]	0.99, [0.22 ;1.94]
denC $<$ inter	1.00, [0.33 ; 2.87]	0.99, [0.29 ;1.93]

Table 1: Full results of regression analysis.

Discussion. Exp.2 suggests that statements of association can influence practical decision making to a similar extent as descriptions of intervention effects. The data are in principle compatible with the idea that mere mentioning of "ralocrop" alone increases belief in a potential causal connection, and that a statement of association, while not actually *in*creasing beliefs in a causal connection, merely only *de*creases beliefs to a certain extent. Nevertheless, the studies show a behavioral effect that hints at different degrees of causal beliefs triggered by different linguistic expressions, some with a clearly non-causal literal meaning.

Conclusion

People do draw causal inferences from correlational language. Is this due to a cognitive error, or a rational feature of language understanding? In two sets of experiments we provided support for a picture in which it is rationally appropriate for listeners to draw rich, context-dependent causal inferences from correlational statements. We related the details of the context-dependence to details of the linguistic form and meaning of sentences, the discourse context in which they occur, and the practical decision context surrounding the discourse.

Building on a preliminary hypothesis drawn from discourse pragmatics and psycholinguistics, the studies in Exp.1 provided evidence that PACE effects (Preferences for Assigning Cause and Effect) are highly sensitive to the grammatical form and semantic content of sentences. We suggested that both factors can be related to discourse interpretation, in particular the effects of implicit QUDs and the contextual constraints on QUD inference. We also found more speculative evidence for a role of an explicit information-structural manipulation (contrastive topic) and of order effects in binomial ("*A* and *B*") constructions.

Exp.2 provided preliminary evidence that associational language can convey causal information that is strong enough to influence decision-making. This result may be explained by the fact that the scenario we constructed involves a communicative interaction embedded in a practical decision in a specific way: participants can reasonably expect that the science team would only provide correlational information if it is practically relevant to the decision at hand. If so, the causal

Context: You are leading a group of colonists to a far away planet called Xelifan-3. To survive on Xelifan-3, the colonists need constant supply of the seeds of a plant called xeliherb, which grows sparsely and only on Xelifan-3. You will need to cultivate xeliherb to ensure survival of the colonists. Your science team have explored the planet and found that ...

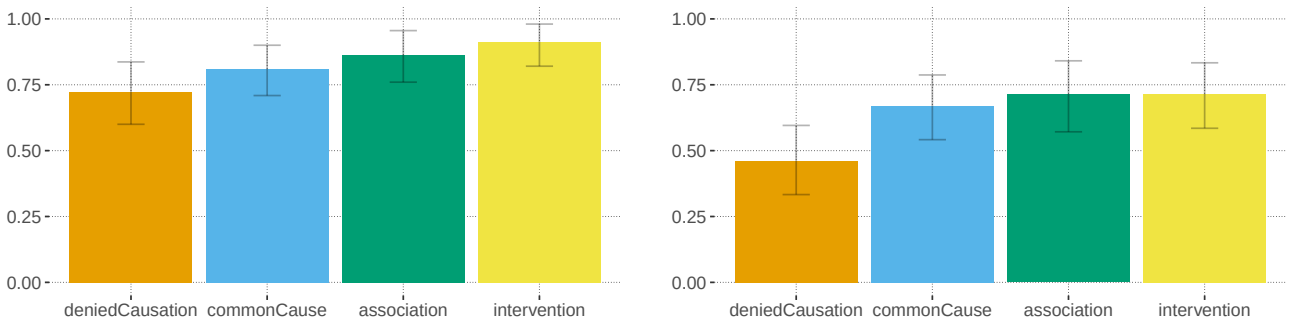
... a high yield of xeliherb is associated with the presence of another plant called ralocrop.

However, the cultivation of ralocrop is costly (water, energy resources).

Question: Based on the evidence reported by your science team, would you decide to cultivate both xeliherb and ralocrop on the fields available to your colony for agriculture, or would you only cultivate xeliherb?

BOTH XELIHERB AND RALOCROP | ONLY XELIHERB

Figure 2: Screenshot from main trial of Exp. 1b. For Exp. 1a the sentence "However, the cultivation ..." was ommitted.



(a) Exp. 2a

(b) Exp. 2b

Figure 3: Average "both" choices for different conditions in Experiments 2a and 2b. Error bars are bootstrapped confidence intervals for the mean.

inferences that participants made in Exp.2 are similar in nature to the intuitive causal interpretation of the public messaging in Fig.1. In both cases, only one direction of intended causal influence would be sufficient to rationalize a speaker's choice to provide this particular piece of correlational evidence.

While correlation does not imply causation, correlational *statements* can. There is virtually no previous linguistic work on the correlation vs. causation issue, and the research reported here generates more questions than answers. However, it does point the way toward a more detailed understanding of the linguistic and contextual factors that influence the rational inference of causation from correlational language.

References

Adams, R. C., Sumner, P., Vivian-Griffiths, S., Barrington, A., Williams, A., Boivin, J., ... Bott, L. (2017). How readers understand causal and correlational expressions used in news headlines. *Journal of experimental psychology: applied*, *23*(1), 1.

Ariely, D., & Jones, S. (2008). *Predictably irrational*. HarperCollins New York.

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge University Press.

Beaver, D. I., Roberts, C., Simons, M., & Tonhauser, J. (2017). Questions under discussion: Where information structure meets projective content. *Annual Review of Linguistics*, *3*, 265–284.

Benor, S. B., & Levy, R. (2006). The chicken or the egg? a probabilistic analysis of english binomials. *Language*, 233–278.

Büring, D. (2003). On D-trees, beans, and B-accents. *Linguistics and philosophy*, *26*(5), 511–545.

Carlson, L. (1983). *Dialogue games: An approach to discourse analysis*. Springer.

Clark, H. H. (1977). Bridging. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking.* Cambridge University Press.

Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive psychology*, *11*(4), 430–477.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension.* Ablex Publishing Company.

Danks, D. (2014). *Unifying the mind: Cognitive representations as graphical models*. MIT Press.

Erteschik-Shir, N. (2007). *Information structure: The syntax-discourse interface* (Vol. 3). OUP Oxford.

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

Franke, M., Ji, X., Ilieva, S., Rautenstrauch, J., & Klehr, M. (2021). *magpie: Minimal architecture for the generation of portable interactive experiments.* Retrieved from `https://magpie-experiments.org/`

Gershman, S. J., & Ullman, T. D. (2023). Causal implicatures from correlational statements. *Plos one*, *18*(5), e0286067.

Ginzburg, J. (1995). Resolving questions, I. *Linguistics and Philosophy*, *18*(5), 459–527.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and bayes nets. *Psychological Review*, *111*(1), 3.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological review*, *116*(4), 661.

Hawkins, R. X., Stuhlmüller, A., Degen, J., & Goodman, N. D. (2015). Why do you ask? good questions provoke informative answers. In *Cogsci 2015.*

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive science*, *3*(1), 67–90.

Huff, D. (1954). *How to lie with statistics*. Norton.

Kao, J. T., Wu, J. Y., Bergen, L., & Goodman, N. D. (2014). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences*, *111*(33), 12002–12007.

Kehler, A., & Rohde, H. (2017). Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes*, *54*(3), 219–238.

Rehder, B. (2017). Concepts as causal models: Categorization. In M. R. Waldmann (Ed.), *The oxford handbook of causal reasoning* (pp. 347–376). Oxford University Press New York, NY.

Reichenbach, H. (1956). *The direction of time* (Vol. 65). Univ of California Press.

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

Seifert, C. M., Harrington, M., Michal, A. L., & Shah, P. (2022). Causal theory error in college students' understanding of science studies. *Cognitive Research: Principles and Implications*, *7*(1), 4.

Sloman, S. A. (2005). *Causal models: How we think about the world and its alternatives*. OUP.

Strawson, P. F. (1964). Identifying reference and truth-values. *Theoria*, *30*(2), 96–118.

Sumers, T. R., Ho, M. K., Griffiths, T. L., & Hawkins, R. D. (2023). Reconciling truthfulness and relevance as epistemic and decision-theoretic utility. *Psychological Review*.

Tversky, A., & Kahneman, D. (1982). Causal schemas in judgments under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117–128). Cambridge University Press.

Vallduví, E. (2016). Information structure. In M. Aloni & P. Dekker (Eds.), *Cambridge handbook of formal semantics.* Cambridge University Press.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, *31*(01), 109–147.