Proceedings of SALT 35: 64–83, 2025

Which alternatives matter? The role of the Question Under Discussion and speaker knowledge in conditionals[*]

Ebru Evcen
UC San Diego

Alan Bale
Concordia University

David Barner
UC San Diego

Abstract Conditional perfection (CP) is often analyzed as a quantity implicature that arises when "*if p*, *q*" is strengthened by excluding relevant alternatives to *p*. If this analysis is correct, CP should not arise in contexts that do not license exhaustification, such as when the Question Under Discussion (QUD) does not make alternative antecedents salient, or when the speaker cannot be assumed to know them. To date, experimental studies on CP have revealed limited or inconclusive evidence in support of these typical signatures of quantity implicatures. In two experiments, we demonstrate robust support for both signatures: Experiment 1 found that CP arose only when the conditional answered an antecedent-focused QUD; consequent-focused and neutral QUDs did not lead to perfection. Experiment 2 showed that CP was further constrained by the speaker's epistemic access: listeners computed a perfected reading only when the speaker appeared to know the full set of relevant antecedents. Our findings provide robust evidence that both QUD and the speaker's knowledge state influence the interpretation of conditional statements and that CP is indeed a quantity implicature.

Keywords: conditional perfection, quantity implicature, question under discussion, epistemic reasoning, scalar alternatives

1 Introduction

Conditional perfection (CP; Geis & Zwicky 1971) is a phenomenon where conditionals of the form "*if p*, *q*" are understood exhaustively by listeners to mean "*if and only if p*, *q*," thereby treating them as biconditionals[1]. For instance, although the literal meaning of "If you mow the lawn, you will receive \$5" requires only that mowing the lawn is a sufficient condition for earning five dollars, listeners often infer the stronger meaning that mowing the lawn is the only way to obtain five dollars. Formally, CP corresponds to an inference pattern of the form $p \Box\!\!\rightarrow q \Rightarrow r \Box\!\!\rightarrow \neg q$ for any contextually salient alternative antecedent r, where $\rightarrow$ serves as theory-neutral notation for modal implication. For concreteness, we adopt a Stalnaker-style semantics in which $p \Box\!\!\rightarrow q$ is true at a world w just in case the closest p-world to w is a q (Stalnaker 1968). Under bivalence, $\neg(r \Box\!\!\rightarrow q)$ entails $r \Box\!\!\rightarrow \neg q$. An analogous pattern can also be derived in a Lewis-style semantics (Lewis 1973), assuming standard auxiliary conditions such as homogeneity across the relevant set of closest r-worlds (also see the discussion in von Fintel 2001).[2]

---

[*] For discussion of this topic, we thank members of the Language and Development Lab at UC San Diego, members of S-Babble at UC San Diego, audiences at SALT 35, anonymous SALT reviewers, and Katherine Jamison and Grace Carlson for help with the video stimuli.

[1] We use exhaustification in a broad, theory-neutral sense to mean strengthening. Our use of the term does not presuppose a grammatical implementation (e.g., via an Exh operator), but is compatible with both grammatical and Gricean accounts.

CP is commonly analyzed as arising from a presumption of exhaustivity as a quantity implicature (Cornulier 1983; Geis & Zwicky 1971; Horn 2000; Levinson 2000; van der Auwera 1997; von Fintel 2001), often grounded in Gricean theories of pragmatic reasoning (e.g., Horn 1972; Grice 1975; Gazdar 1979; Geurts 2010), or similarly inspired grammatical algorithms (e.g., Chierchia 2004; Fox 2007; Chierchia, Fox & Spector 2012). These accounts, while differing in their details, converge on the idea that CP arises when the condition stated in the antecedent p is understood to exhaust the set of conditions sufficient for the consequent q. On this view, a listener infers that the antecedent p exhausts all sufficient conditions for q, strengthening the meaning of the conditional. Once the listener identifies other antecedents that could plausibly bring about q, as in cases where several tasks are known ways of earning five dollars, the strengthened inference is suspended. If CP is a quantity implicature of this sort, then factors known to block scalar implicatures should likewise limit conditional perfection. The present study focuses on two such factors that regulate the salience and relevance of alternatives. First, the Question Under Discussion determines which alternatives are contextually relevant and therefore available for exhaustification. Second, the epistemic state of the speaker constrains whether the listener may assume that the speaker is in a position to rule out these alternatives.

Critical to the present paper, however, previous accounts argue that CP should not arise if listeners identify alternatives to p that could also lead to q. For instance, a statement like, "If Mary exercises, she will lose weight," is less likely to be interpreted pragmatically (e.g., as "If and only if Mary exercises, will she lose weight") because alternative causes, such as going on a diet, are easily imagined (Cummins 1991, 1995; Janveau-Brennan & Markovits 1999; De Neys & Everaerts 2008; Blochowiak, Castelain, Rodriguez-Villagra & Musolino 2022). Relatedly, because perfection is thought to involve exhaustification of the antecedent, it should arise when the Question Under Discussion (QUD; Roberts 1996, 2012) relates to antecedents that will bring about the consequent, but not when the QUD pertains to the antecedent itself. For example, if asked, "What will happen if I mow the lawn?", the statement, "If you mow the lawn, you will receive $5," no longer implies that mowing the lawn is the only way to receive $5, since the listener considers alternatives to the consequent, rather than to the antecedent, and therefore does not exhaustify the antecedent.

---

[2] Our empirical contribution does not hinge on any particular choice between these frameworks; the results reported here are compatible with either analysis.

The idea that conditional inferences are contingent upon access to alternatives is also found in accounts of other forms of implicature, but with important differences. For example, on neo-Gricean accounts of implicature, given the statement in (1), "There are cookies in some of the boxes," the listener infers that the speaker believes that the stronger statement in (2) is false, and therefore that some, but not all of the boxes contain cookies. In particular, the listener reasons that if the speaker had evidence that the stronger statement in (2) were true, then they should have uttered that statement instead (e.g., Maxim of Quantity). Given this, they infer that the speaker does not believe the stronger statement to be true. Crucially, this inference relies on the "competence assumption", the belief that the speaker has sufficient knowledge to judge whether the alternative statement is false (Sauerland 2004; Geurts 2010).

(1) There are cookies in some of the boxes.

(2) There are cookies in all of the boxes.

What's critical to observe is that, in such cases, a strong scalar implicature arises when a statement is enriched via the negation of stronger alternative statements. Given this fact, a strong implicature is not expected to arise if the listener believes that the speaker is ignorant about the status of stronger statements (violating the competence assumption). For example, if the listener watches the speaker look into only one of three possible cookie boxes and then hears the statement in (1), they should conclude that the speaker is ignorant regarding the stronger alternative contained in (2), and therefore fail to negate it to compute an implicature. Similarly, the inference should not arise if the contextually defined QUD activates different alternative utterances, instead. For example, previous studies find that if the speaker is asked a question like, "Are there cookies in any of the boxes?", a reply containing the utterance in (1) may be less likely to generate the implicature that the statement containing "all" is false (Degen 2013; Degen & Tanenhaus 2015; Zondervan, Meroni & Gualmini 2008; Ronai & Xiang 2021; Kursat & Degen 2020). This is because the QUD relates not to the quantity of boxes that contain cookies in this context, but instead to the kinds of things contained in some of the boxes.

How do conditionals differ from putative forms of scalar implicatures? First, although most accounts agree that perfection arises via exhaustification of the antecedent, they disagree about what this entails, and whether it involves the negation of particular alternatives (akin to some/all inferences), or instead relies on a different form of inference (see van der Auwera 1997 for a comprehensive overview of proposed scales and related discussion). While it is tempting to assume that perfection, like other quantity implicatures, is computed by negating stronger alternatives, such an approach is difficult to maintain because it is often impossible to define the set of relevant alternatives that should be negated. For example, van der Auwera 1997 suggests that more informative alternatives to a statement like "if p then q" might be generated by substituting *p* with all other possible antecedents—for example, "if a then q," "if b then q," "if c then q," and so on. If the speaker thought it were possible to receive $5 by trimming a hedge, they would have mentioned it. However, this approach faces a challenge: while Horn scales rely on a closed set of well-defined alternatives known to both speaker and listener, this proposal requires generating and negating an unbounded set of possible alternatives, a process that is widely considered psychologically implausible (Horn 2000; von Fintel 2001). Given this, Horn (2000) offers an alternative account, proposing that the stronger scalemate to "if p then q" is the unconditional "q, no matter what." By choosing a conditional rather than an unconditional form, the speaker signals that the consequent *q* depends on some condition *p*. While this account supports the inference that *q* is not unconditional or "for free," it does not rule out the possibility of other antecedents and therefore does not directly derive perfection (von Fintel 2001).

Such considerations have led others to argue that, unlike some/all implicatures, conditional perfection relies on a more global exhaustivity assumption (Cornulier 1983; Groenendijk & Stokhof 1984; von Fintel 2001). Specifically, on this account, listeners infer that the absence of any mention of alternative conditions by the speaker implies that no such alternatives exist (Cornulier 1983). This inference is drawn from the expectation that the speaker, if aware of any other sufficient conditions, would have mentioned them (similar to quantity implicatures). This sidesteps the challenge of explicitly defining scales, but leaves a critical question unanswered: what might drive the assumption that the speaker should mention all possible sufficient conditions? To address this, von Fintel 2001 provides an account of CP that integrates the theory of questions proposed in Groenendijk & Stokhof 1984 into a pragmatic framework. The central idea is that exhaustivity arises naturally in question-answer pairings (see also Franke 2009; Nadathur 2013). For instance, in response to the question, "Who is in the garden?", the answer "Mary" is typically interpreted as exhaustive, implying that Mary is the only person in the garden. This interpretation relies on two key assumptions: (1) the speaker follows the Gricean principles of being cooperative and informative, and (2) the speaker is competent regarding the question they are answering (van Rooij 2004; van Rooij & Schulz 2004). By analogy, when a conditional is understood as answering an explicit or implicit question about what is sufficient for the outcome, listeners often treat it exhaustively, which yields the perfected interpretation.

A further consequence of this distinction is that alternatives limit strengthened readings for conditional perfection in a different way than they do for classic scalar implicatures. In scalar implicatures, accessing a stronger alternative (e.g., all) is part of deriving the strengthened reading: the listener identifies the stronger alternative and negates it. For conditional perfection, however, the relevant alternatives are other salient antecedent conditions r such that $r \rightarrow q$. Whether these alternatives are negated depends on the QUD. When the QUD requires an exhaustive list of sufficient conditions, for example, "Under what conditions will q occur?", the listener expects the speaker to mention all antecedents that could license q. In such contexts, hearing only p in "If p, then q" may lead the listener to infer that no other antecedent r licenses q, yielding a perfected interpretation. This is analogous to interpreting "Mary" as the exhaustive answer to "Who is in the garden?". In contrast, a QUD such as "What happens if p?" requires an exhaustive list of consequents, not antecedents. Because this QUD does not treat $r \rightarrow q$ as a possible answer, these antecedent alternatives remain unnegated, and conditional perfection does not arise.

On most accounts, the speaker's epistemic state plays a parallel role, since exhaustification should target only alternatives whose truth value is known to the speaker. Under a Gricean approach, for instance, deriving a quantity implicature from an utterance expressing p requires first identifying a relevant alternative q (Matsumoto 1995), typically determined by the QUD. Once q is established as relevant, the listener reasons, by the Maxim of Quantity, that if the speaker knew q to be true, they would have said so explicitly. Because the speaker did not, the listener infers that the speaker does not know that q is true. Crucially, this inference reflects ignorance rather than negation: the speaker may simply lack knowledge about q's truth value. To conclude that q is false, the listener must assume that the speaker is epistemically competent with respect to q – that they know whether q or its negation holds. Once this assumption is in place, ignorance can be reinterpreted as evidence for falsity, yielding the strengthened inference. Although this discussion adopts a Gricean perspective, the same reasoning can be incorporated into grammatical theories of implicature (Fox 2007; Chierchia et al. 2012). In such accounts, an exhaustification operator can be defined so that it negates only those alternatives that both answer the QUD and are known to the speaker, ensuring that exhaustification depends on epistemic access as well as contextual relevance.[3]

---

[3] Formally, if KS(x) denotes that the speaker knows whether x, then assuming $KS(q) \vee KS(\neg q)$ and $\neg KS(q)$ entails $KS(\neg q)$, which in turn entails $\neg q$. In grammatical terms, this can be captured by defining the set of alternatives ALT(p) so that $ALT(p) \subseteq ANS_{\text{QUD}} \wedge ALT(p) \subseteq \{q : KS(q) \vee KS(\neg q)\}$, ensuring that the exhaustification operator negates only those alternatives that are both relevant and epistemically accessible.

Interestingly, the role of speaker knowledge in conditional perfection has received very little attention. This is surprising, given robust evidence from other types of implicatures that highlight the critical role of speaker knowledge (e.g., Bergen & Grodner 2012; Breheny, Ferguson & Katsos 2013; Goodman & Stuhlmüller 2013; Hochstein, Bale & Barner 2018; Bale, Noguchi, Rolland & Barner 2025). Empirical work on conditional perfection has focused on testing the effects of the QUD on perfection, with some finding strong evidence that QUD impacts perfection (e.g., Farr 2019, and others finding little to no evidence of such effects (e.g., Cariani & Rips 2023; Grusdt, Liu & Franke 2023). Meanwhile, the role of the speaker's epistemic state in the interpretation of conditionals has yet to be tested empirically.

In one of the few studies to explore the role of the QUD, Farr (2019) investigated whether manipulating the QUD changes how conditionals are interpreted. Participants read short scenarios in which one character, Kerstin, asked another, Sahra, about seafood prices. The situation allowed for two ways to earn 2.50 euros, by selling either an eel or a pike. Depending on the condition, Kerstin asked either a consequent-focused question ("What happens if I sell an eel?") or an antecedent-focused question ("When do I get 2.50 euros?").[4] Sahra's response was always the same: "If you sell an eel, you get 2.50 euros." Participants were more likely to judge the response as complete or sufficient when it followed the antecedent-focused question, consistent with an effect of the QUD. However, as noted in Cariani & Rips 2023, these results may not reflect true conditional perfection. Instead, participants' dissatisfaction with the consequent-focused question may have stemmed from missing information. Knowing that two ways to earn money existed, listeners might have expected Sahra to mention both. Thus, the observed pattern could reflect incomplete disclosure rather than a perfected reading.

To address this, Cariani & Rips (2023) conducted seven follow-up experiments using short vignettes. In one example, participants read that "John took a test covering Chapters 4 to 6 that has not yet been graded." They were then asked one of two questions: either an antecedent-focused one ("What are all the ways John could do well on the test?") or a consequent-focused one ("What could happen if John understood Chapter 5?"). In both conditions, a third character, Mary, responded with the statement "If John understood Chapter 5, then John did well on the test." Participants then judged whether this response implied that John had in fact understood

---

[4] Different labels for these QUDs appear in the literature. Farr (2019) distinguishes between "what-if-p?" questions, which focus on variation in the consequent (referred to here as consequent-focused), and "when-q?" questions, which focus on variation in the antecedent (referred to here as antecedent-focused). Cariani & Rips (2023) introduce the terms ANT? (antecedent-directed QUDs) and CONS? (consequent-directed QUDs), depending on whether variation is constrained to the antecedent or consequent, respectively. Grusdt et al. (2023) use "if-p?" for consequent-focused QUDs and "will-q?" for antecedent-focused QUDs, also including a "neutral" category.

Chapter 5, given that he did well on the test. Across experiments, responses did not differ reliably across the two QUD conditions. Interestingly, when Cariani & Rips (2023) added the explicit operator only (e.g., "That is the only way John could have done well on the test"), participants were more likely to derive a perfected interpretation in the consequent-focused condition than in the antecedent-focused one. The authors concluded that demanding exhaustivity is not, by itself, enough to guarantee perfection. They also noted that consequent-focused questions in their experiment might have been harder to answer exhaustively, since multiple possible explanations for success (e.g., guessing, cheating, or luck) were always plausible.

Grusdt et al. (2023) also examined the impact of QUD on conditional perfection, but in more controlled settings designed to minimize world knowledge. Participants viewed animated scenes involving simple physical events, such as an upper block striking a ball and causing a lower block to fall. One character posed either a consequent-focused question ("What happens if the upper block falls?") or an antecedent-focused question ("Will the lower block fall?"), to which another character, Bob, replied, "If the upper block falls, the lower block will fall." Participants then chose between two possible scenes: an exhaustive one, where only the upper block's fall caused the lower block to drop, and a non-exhaustive one, where other factors could also produce the effect. A small QUD effect was found, participants showed a slight preference for exhaustive interpretations when the question focused on the antecedent, but this effect was inconsistent and disappeared when low-level visual features, such as the stability of the blocks, varied across trials. The authors concluded that the influence of QUD on conditional interpretation was not robust and might be due to the unintended features of the stimuli.

In summary, most previous accounts treat conditional perfection as a form of exhaustivity inference, wherein listeners assume that: (1) the antecedent of a conditional provides an exhaustive answer to a contextually-specified question under discussion, e.g., "What are the conditions for the consequent?", and (2) a knowledgeable speaker should list all possible antecedents they believe to be true, implying that unmentioned alternatives are not believed to be true. By extension, conditional perfection is thought to be canceled if the speaker is perceived as ignorant about alternative antecedents or if the QUD is interpreted as focusing on the consequent rather than the antecedent. In the present study, we report data from two novel experiments that investigate the role of alternative utterances in conditional perfection using the same experimental paradigm. One in which we test for the influence of different QUDs by providing simple contexts that assess the computation of implicatures with straightforward yes-no questions, which do not require metalinguistic judgments, and another where we use the same methods to test for the influence of the speaker's epistemic state.

2 Experiment 1

In Experiment 1, using a task modeled after Hochstein et al. 2018, we tested how participants interpret conditionals across three different types of QUDs that focused either on the antecedent or the consequent. Participants watched videos in which a character, Mary, pressed three buttons (red, blue, and orange), which each produced an animal sound audible only to her (because she was wearing headphones; see Figure 1). In each video, another character asked one of three questions: a consequent-focused question (e.g., “What will happen if I press the blue button?”), an antecedent-focused question (e.g., “Which of these buttons will play a dog sound?”), or a neutral question (e.g., “What will happen if I press the buttons?”). Mary always responded with the same conditional statement: “If you press the blue button, it will play a dog barking.” The type of question determined the relevant set of alternatives: consequent-focused questions put the focus on different sounds (e.g., “it will play a cat meowing”, “it will play a lion roaring”), while antecedent-focused questions targeted different buttons (e.g., “if you press the red button”, “if you press the orange button”). We predicted that if perfection is sensitive to QUD, then exhaustive interpretations should arise under the antecedent-focused questions because the alternatives to the antecedent must be considered. In particular, since the speaker explicitly mentioned the blue button, participants might infer that only the blue button plays the dog sound, as other buttons would have been mentioned otherwise. In contrast, for the consequent-focused or neutral questions, where the alternative utterances concerned the consequent, this exhaustive inference should not arise if perfection is sensitive to the QUD. All methods and analyses were preregistered prior to data collection. The preregistration is available at https://osf.io/9tcmw. All data and analysis code for this are available at https://osf.io/vq72b.

2.1 Participants

The study was conducted using the PCIbex platform (Zehr & Schwarz 2018). We recruited 104 native speakers of English through Prolific, limiting participation to users located in the United States. Following the preregistered criteria, data from six participants were excluded: one due to technical issues and five for failing attention checks, leaving a final sample of 98 participants. All participants were compensated at a rate equivalent to California’s minimum hourly wage at the time ($15.50/hr).

2.2 Materials and Procedure

Participants were randomly assigned to one of three between-subjects conditions corresponding to the type QUDs: antecedent-focused, consequent-focused, or neutral. Each trial contained two short videos. In the first, a character named Mary pressed three colored buttons (red, blue, and orange). Each button produced a distinct animal sound that only Mary could hear through her headphones. After this phase, participants were asked a comprehension question (e.g., Does Mary know what sound the orange button makes?) to ensure they understood that Mary had learned the sound produced by each button (see Figure 1).

Experiment 1

Mary puts her headphones on.

Mary presses all three buttons and listens.

Mary responds, "If you press the blue button, it will play a dog barking"

Another speaker asked one of three questions before Mary's response:		
Consequent-focused	Antecedent-focused	Neutral
"What will happen if I press the blue button?"	"Which of these buttons will play a dog sound?"	"What will happen if I press the buttons?"
Critical question: Do you think the orange button plays a dog sound? (Yes / No / Can't tell)		

Figure 1: Still images from video trials in Experiment 1, depicting Mary putting on headphones, listening to the buttons, and responding with a conditional statement

In the second video, Mary spoke with another character who posed one of the three QUD-type questions: a consequent-focused question (e.g., What will happen if I press the blue button?), an antecedent-focused one (e.g., Which of these buttons will play a dog sound?), or a neutral one (e.g., What will happen if I press the buttons?). Mary's reply was the same for all three types of questions: "If you press the blue button, it'll play a dog barking." After viewing this exchange, participants were asked to respond to the critical question, e.g., "Do you think the orange button will play a dog sound?"

Before the experiment began, participants read instructions that explained the task and the structure of the videos. Participants were told that Mary did not know in advance which sound each button made and that she would offer clues based on what she heard. Because participants could not hear the sounds themselves, their task was to infer what sound each unpressed button might make using Mary's statements. The instructions emphasized that buttons could, in principle, share the same sound or differ from one another, and that Mary was cooperative rather than deceptive. After each video, the participants answered the questions choosing "Yes," "No" or "Can't tell," with the final option reserved for cases where Mary's statements did not provide enough information to draw a conclusion. The critical conditional statement given by Mary was consistent across trials, but the manipulation of the QUD allowed us to investigate whether participants interpreted the statement literally, resulting in more “Can’t tell” responses, or pragmatically, leading to more “No” responses. “Yes” responses were technically possible but not predicted for critical items. Each participant completed seven trials, with button–sound pairings and trial order pseudorandomized. Three of these were attention checks. In these control trials, Mary said, “If you press any of the buttons, it will play a dog barking.” and we expected participants to respond “Yes” when asked about another button in these control trials.

2.3 Results and Discussion

Participants responded as expected on attention-check trials (98.63% correct), confirming they understood the task. In line with preregistered exclusions, we removed 1.62% of trials that failed the comprehension check about speaker knowledge, and another 3.53% of data points (14 responses) where participants selected “Yes” on critical trials, unexpected but possible responses.

To investigate how QUD type influenced interpretation, we fit a generalized logistic mixed-effects model predicting “No” (vs. “Can’t tell”) responses. The model included QUD Type as a fixed effect and random intercepts for both participants and items. Adding random slopes for QUD Type caused overfitting and was therefore omitted. The model revealed a significant effect of QUD Type ($\chi^2(2) = 57.32$, $p < .001$). Pairwise comparisons with Bonferroni correction (using emmeans in R) showed that “No” responses were much more frequent for antecedent-focused questions (M = 0.65, SE = 0.10) than for consequent-focused (M = 0.22, SE = 0.10; β = –18.08, SE = 2.46, z =–7.34, $p < .001$) or neutral questions (M = 0.29, SE = 0.11; β = –17.30, SE = 2.37, z =–7.31, $p < .001$). No reliable difference emerged between consequent-focused and neutral conditions ($p > .05$). Overall, antecedent-focused questions prompted more pragmatic interpretations, while the other two QUD types yielded more literal readings.

Experiment 1 showed a clear effect of the QUD on conditional perfection. When the conditional served as an answer to an antecedent-focused question, participants frequently produced perfected readings. In contrast, perfection was rare when the QUD targeted the consequent, and similarly low in the neutral condition. This is expected if listeners align exhaustification with the QUD: a neutral question like “What will happen if I press the buttons?” makes both antecedent- and consequent-based alternatives available (e.g., “If you press the orange button, it will play a dog sound,” “If you press the red button, it will play a pig sound,” or “If you press the orange button, it will play a cat sound.”). Strengthening against all of them would incorrectly imply that the speaker learned nothing from pressing the other

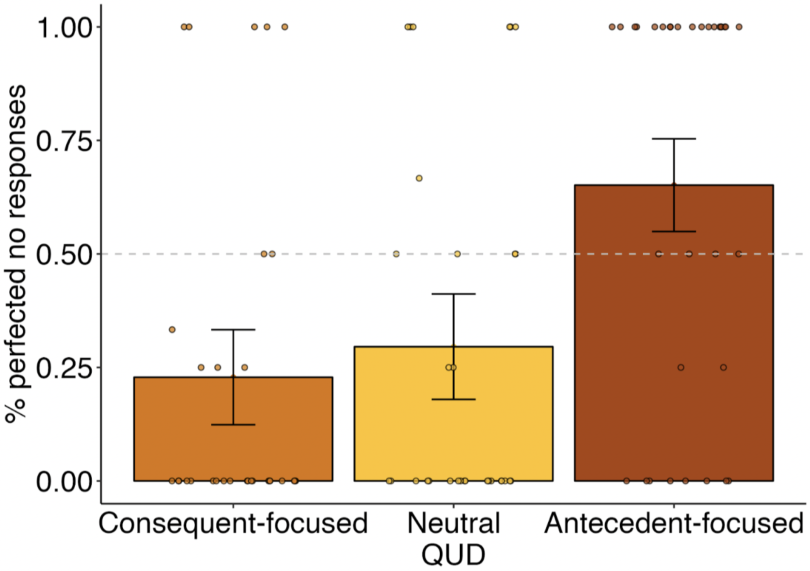


Figure 2: Percentage of perfected "no" responses by QUD Type in Experiment 1. Each dot represents a participant. Error bars represent the standard error of the mean, computed across participants

buttons, contradicting the scenario. Prior work shows that such contradictions block strengthening (Ginzburg 1995a,b; Beck & Rullmann 1999; van Rooij 2004). Overall, the results support the view that the QUD constraints which alternatives are considered viable during interpretation.

3 Experiment 2

In Experiment 2, using the same paradigm in Experiment 1, we tested how the speaker's epistemic state modulates perfection for antecedent-focused questions (e.g., "Which of these buttons will play a dog sound?"). Participants viewed trials where the speaker, Mary, pressed and listened to either two buttons (partial knowledge trials) or all three buttons (full knowledge trials) before making a conditional statement about the sounds the buttons produced, such as, "If you press the blue button, it will play a dog barking," similar to Experiment 1 (see Figure 3). We reasoned that if listeners' inferences depend on assuming that the speaker has complete information, then perfection should drop when Mary lacks knowledge of all possibilities. The reasoning follows prior work on scalar implicatures: strengthening an utterance like *some* to *not all* requires that the listener take the speaker to be informed about stronger alternatives but to have chosen not to assert them (e.g., Bergen & Grodner

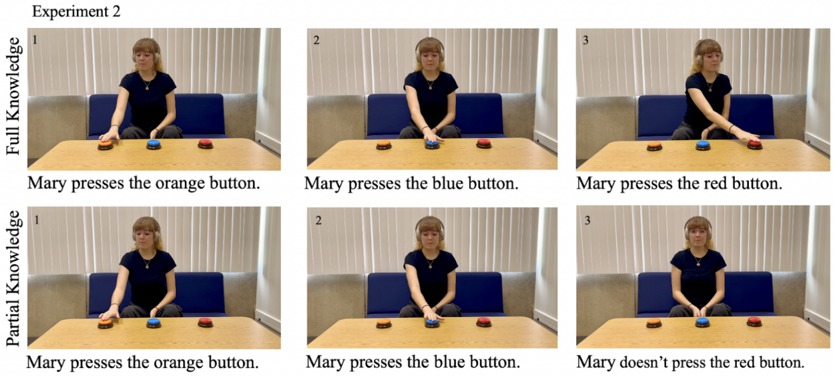
2012; Breheny et al. 2013; Hochstein et al. 2018; Goodman & Stuhlmüller 2013). All methods and analyses were preregistered prior to data collection. The preregistration is available at https://osf.io/e7n2f. All data and analysis code for this are available at https://osf.io/vq72b.

3.1 Participants

Seventy-five native English speakers were recruited through Prolific, restricted to users located in the United States. Following the preregistered criteria, three participants were excluded, one for failing attention checks and two for incorrect responses on knowledge-check questions, with 72 participants in the final sample. Each participant was compensated at the California minimum wage rate at the time $(15.50/hour)$.

3.2 Materials and Procedure

The structure of the task closely followed that of Experiment 1, except that we now manipulated Mary's epistemic access as a within-subject variable while holding the QUD constant as antecedent-focused ("Which of these buttons will play a dog sound?").



Mary presses the orange button. Mary presses the blue button. Mary presses the red button.

Mary presses the orange button. Mary presses the blue button. Mary doesn't press the red button.

Figure 3: Still images from video trials in Experiment 2, showing Mary pressing all three buttons in the Full Knowledge condition and only two buttons in the Partial Knowledge condition.

Each trial contained two short videos. In the first, Mary either pressed all three buttons (full-knowledge trials) or only two (partial-knowledge trials). Participants then answered a comprehension question confirming whether Mary knew what sound

the remaining button produced. In the second video, Mary was asked the antecedent-focused question and replied, “If you press the blue button, it will play a dog barking.” Participants then judged whether the orange button, which Mary either had or had not tested, made the same sound. A literal reading of the conditional should yield a “Can’t tell” response; a perfected reading should yield a “No” response.

Each participant completed nine trials in total: three in each knowledge condition and three attention-check trials. Attention checks repeated the format from Experiment 1, using statements such as “If you press any of the buttons, it will play a dog barking”, triggering a “Yes” response when asked about the other buttons. To balance potential order effects, two pseudorandom trial orders were created so that half the participants began with full-knowledge trials and the other half with partial-knowledge ones.

3.3 Results and Discussion

Participants responded to the attention check questions as expected (98.14% correct), confirming that participants understood the task. Trials in which participants misidentified Mary’s knowledge state were removed (3.24% of all data), as were the few unexpected “Yes” responses in critical trials (12 data points; 2.77%).

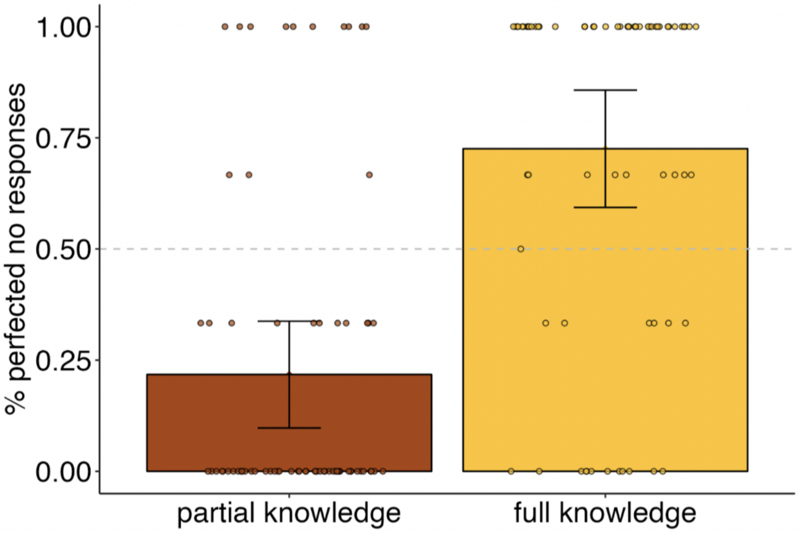


Figure 4: Percentage of perfected “No” responses by Knowledge State in Experiment 2. Each dot represents a participant. Error bars represent the standard error of the mean, computed across participants.

To investigate whether participants perfected conditionals less often when the speaker lacked complete information, we ran a generalized logistics mixed-effects model that predicted the probability of “No” (vs. “Can’t tell”) responses from the fixed effect of Knowledge State (Full vs. Partial), with random intercepts for both participants and items. The model showed a robust effect of Knowledge State ($\chi^2(1) = 26.01$, $p < .001$). Participants were far more likely to respond “No” when Mary had full knowledge (M = 0.72, SE = .13) than when she had partial knowledge (M = 0.21, SE = .12; β = –11.07, SE = 2.86, z =-3.86, $p < .001$). These findings suggest that participants were more likely to adopt a pragmatic, perfected interpretation when the speaker demonstrated full knowledge of the alternatives (see Figure 4).

4 General Discussion

In the present study, we asked how listeners interpret conditional statements when the discourse context varies the relevance of alternative propositions. Across two experiments, we tested two factors that, on many accounts, set the space of alternatives for exhaustification: the QUD, which determines which propositions count as possible answers, and the speaker’s epistemic state about those alternatives. In Experiment 1, we found that perfection depended on how the QUD structured the space of alternatives. When the conditional functioned as an answer to an antecedent-focused question, one whose possible answers vary the antecedent while holding the consequent constant, listeners inferred that unmentioned antecedents would not yield the same outcome, and thus perfected the conditionals. In contrast, when the conditional answered a consequent-focused or neutral question, the relevant alternatives no longer varied the antecedent, and correspondingly, conditional perfection was less likely. In Experiment 2, when the speaker was presented as having full epistemic access to all antecedents, perfection was robust: listeners inferred that unmentioned antecedents did not lead to the consequent. When the speaker was ignorant, those inferences were less likely.

Taken together, these findings provide converging evidence that conditional perfection is sensitive to the contextually defined availability of alternative utterances. This pattern is expected under analyses that treat CP as a type of quantity-based strengthening, where “if p then q” is interpreted relative to a set of alternative antecedents such as “if s then q” or “if r then q” (Cornulier 1983; van der Auwera 1997; Horn 1972; von Fintel 2001). On such accounts, if the relevant alternatives are determined by the QUD, and if exhaustification targets only alternatives whose truth value can be assumed to be known to the speaker, then strengthened readings should arise only when the discourse licenses both assumptions. The present results match these predictions. At the same time, they do not require that CP be derived in exactly the same way as classic scalar implicatures. As noted in the Introduction, the alternatives relevant for perfection are not given by lexical scales but by the discourse structure itself, often through the QUD. This distinguishes CP from cases where alternatives are linguistically encoded and forms a fixed scale, and raises the possibility that perfection reflects a broader mechanism tied to question–answer structure rather than to scales per se.

One question left open by this study is why some previous reports fail to find effects of the QUD. Some studies of conditional perfection rely on tasks that ask participants to make explicit judgments about the truth or felicity of an utterance. For instance, Cariani & Rips (2023) presented participants with conditionals such as “if p then q” under different QUD manipulations, and then asked whether “if p then q” together with “q” implies “p”. Responding to these questions requires participants to maintain the QUD, the conditional statement, and a hypothetical set of truth-conditional possibilities in working memory. These metalinguistic tasks may therefore impose substantial cognitive demands that go beyond the kinds of inferences people make in everyday conversation. In contrast, the button-choice method used here, and in other work on implicature (e.g., Hochstein et al. 2018; Kampa & Papafragou 2020) prompts listeners to draw an inference about the state of the world by pragmatically strengthening the speaker’s utterance — much like they would in everyday conversation — rather than asking them to evaluate the truth, felicity, or acceptability of the utterance itself. While our study cannot directly resolve this question, future work should examine the role that different experimental designs play in eliciting perfection.

In our experiments, we focused on simple causal scenarios such as “If you press the blue button, it’ll play a dog sound,” where we deliberately controlled the context. Extensive work in the reasoning literature has shown that people strengthen conditionals differently depending on the surrounding context (e.g., Newstead 1997; Evans & Twyman-Musgrove 1998; Ohm & Thompson 2004; Evans, Neilens, Handley & Over 2008). For instance, inducements, such as promises and threats like “If you wash the car, I’ll let you borrow it tonight,” are far more likely to be interpreted biconditionally than advice conditionals, such as tips and warnings like “If you take the highway, you’ll get there faster” (Fillenbaum 1976; Evans & Twyman-Musgrove 1998). On the surface, one could argue that both kinds of sentences answer the same antecedent-focused question, something like “What do I need to do for q to happen?”, yet listeners treat them quite differently. One explanation might be that even questions that are exhaustive on the surface may not always require an exhaustive answer (e.g., von Fintel 2001; Nadathur 2013). Conditionals in advice contexts offer one useful option, much like naming a nearby shop when someone asks where to buy Italian newspapers rather than listing every possible store. Inducements, by contrast, are often understood as specifying the full set of conditions relevant to bringing about the outcome. Under this expectation, listeners are more likely to strengthen the conditional. How these expectations interact with antecedent-focused QUDs remains an open question for future work.

5 Conclusion

In conclusion, this work makes two main contributions. Empirically, our findings provide robust evidence that both QUD and the speaker's knowledge state influence the interpretation of conditional statements. While previous studies have reported mixed effects of QUD on conditional perfection, our results demonstrate that exhaustive answers under consequent-focused QUDs reliably yield perfection. Furthermore, we show that the speaker's epistemic status plays an additional role: the set of alternative utterances considered by the listener depends not only on the QUD but also on whether the speaker is perceived as fully knowledgeable. Theoretically, it supports a quantity-based account of conditional perfection, on which listeners treat a speaker's response as exhaustive, assuming that all relevant sufficient conditions have been mentioned. Crucially, a failure to find sensitivity to QUD and speaker knowledge would have counted strongly against treating CP as an implicature. Instead, both experiments showed robust effects of these factors, aligning CP with well-attested signatures of quantity-based strengthening. At the same time, the patterns we observe are also compatible with grammatical approaches in which exhaustivity is introduced by covert operators (e.g., Herburger 2015; Chierchia 2004; Fox 2007), since those accounts can likewise restrict strengthening to contextually relevant and epistemically accessible alternatives. Thus, while our findings support an implicature-based analysis, they do not preclude related grammatical implementations, and the broader landscape of perfection phenomena may well accommodate both.

References

van der Auwera, Johan. 1997. Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics* 27(3). 261–274. doi:10.1016/S0378-2166(96)00058-6.

Bale, Alan, Hiromu Noguchi, Maxime Rolland & David Barner. 2025. Competence by default: Do listeners assume that speakers are knowledgeable when computing conversational inferences? *Journal of Semantics* doi:10.1093/jos/ffae016.

Beck, Sigrid & Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7(3). 249–298. doi:10.1023/A:1008373224343.

Bergen, Leon & Daniel J. Grodner. 2012. Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38(5). 1450–1460. doi:10.1037/a0027850.

Blochowiak, Joanna, Thomas Castelain, Omar A. Rodriguez-Villagra & Julien Musolino. 2022. If and only if people were logical! the effect of pragmatic enrichment on reasoning with abstract and realistic materials. *Journal of Pragmatics* 197. 137–158. doi:10.1016/j.pragma.2022.05.013.

Breheny, Richard, Heather J. Ferguson & Napoleon Katsos. 2013. Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition* 126(3). 423–440. doi:10.1016/j.cognition.2012.11.012.

Cariani, Fabrizio & Lance J. Rips. 2023. Experimenting with (conditional) perfection: Tests of the exhaustivity theory. In Stefan Kaufmann, David Over & Gaurav Sharma (eds.), *Conditionals: Logic, Semantics, Psychology*, 235–274. Springer.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and Beyond*, vol. 3 The Cartography of Syntactic Structures, 39–103. Oxford University Press. doi:10.1093/oso/9780195171976.003.0003.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 3, 2297–2332. De Gruyter.

Cornulier, Benoît. 1983. 'if'and the presumption of exhaustivity. *Journal of Pragmatics* 7(3). 247–249. doi:10.1016/0378-2166(83)90012-7.

Cummins, Denise D. 1991. Children's interpretations of conditional statements. *Journal of Experimental Child Psychology* 52(3). 431–448. doi:10.1016/0022-0965(91)90074-R.

Cummins, Denise D. 1995. Naive theories and causal deduction. *Memory & Cognition* 23(5). 646–658. doi:10.3758/BF03197265.

De Neys, Wim & David Everaerts. 2008. Developmental trends in everyday conditional reasoning: The retrieval and inhibition interplay. *Journal of Experimental Child Psychology* 100(4). 252–263. doi:10.1016/j.jecp.2008.01.007.

Degen, Judith. 2013. *Alternatives in pragmatic reasoning*. University of Rochester.

Degen, Judith & Michael K. Tanenhaus. 2015. Processing scalar implicature: A constraint-based approach. *Cognitive Science* 39(4). 667–710. doi:10.1111/cogs.12171.

Evans, Jonathan St BT, Helen Neilens, Simon J Handley & David E Over. 2008. When can we say 'if'? *Cognition* 108(1). 100–116. doi:10.1016/j.cognition.2008.02.001.

Evans, Jonathan St BT & Joanna Twyman-Musgrove. 1998. Conditional reasoning with inducements and advice. *Cognition* 69(1). B11–B16. doi:10.1016/S0010-0277(98)00062-6.

Farr, Margaret C. 2019. Focus influences the presence of conditional perfection: Experimental evidence. In *Sinn und Bedeutung (SuB) 15*, 225–240. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/374.

Fillenbaum, Samuel. 1976. Inducements: On the phrasing and logic of conditional promises, threats, and warnings. *Psychological Research* 38(3). 231–250. doi:10.1007/BF00309774.

von Fintel, Kai. 2001. Conditional strengthening. *Unpublished manuscript* .

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. London: Palgrave Macmillan. doi:10.1057/9780230210752_4.

Franke, Michael. 2009. *Signal to act: Game theory in pragmatics*. University of Amsterdam.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.

Geis, Michael L & Arnold M Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2(4). 561–566.

Geurts, Bart. 2010. *Quantity Implicatures*. Cambridge: Cambridge University Press.

Ginzburg, Jonathan. 1995a. Resolving questions, i. *Linguistics and Philosophy* 18(5). 459–527. doi:10.1007/BF00985365.

Ginzburg, Jonathan. 1995b. Resolving questions, ii. *Linguistics and Philosophy* 18(6). 567–609. doi:10.1007/BF00983299.

Goodman, Noah D. & Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science* 5. 173–184. doi:10.1111/tops.12007.

Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Speech Acts*, 41–58. Academic Press. doi:10.1163/9789004368811_003.

Groenendijk, Jeroen & Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. In Fred Landman & Frank Veltman (eds.), *Varieties of Formal Semantics: Amsterdam Colloquium*, 143–170. Foris.

Grusdt, Benjamin, Menghan Liu & Michael Franke. 2023. Testing the influence of quds on the occurrence of conditional perfection. *Experiments in Linguistic Meaning (ELM) 2* 104–116. doi:10.3765/elm.2.5413.

Herburger, Elena. 2015. Conditional perfection: the truth and the whole truth. In *Semantics and Linguistic Theory (SALT) 25*, 615–635.

Hochstein, Laura, Alan Bale & David Barner. 2018. Scalar implicature in absence of epistemic reasoning? the case of autism spectrum disorders. *Language Learning and Development* 14(3). 224–240. doi:10.1080/15475441.2017.1343670.

Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in*

English: University of California, Los Angeles Doctoral dissertation.

Horn, Laurence R. 2000. From if to iff: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32(3). 289–326. doi:10.1016/S0378-2166(99)00053-3.

Janveau-Brennan, Geneviève & Henry Markovits. 1999. The development of reasoning with causal conditionals. *Developmental Psychology* 35(4). 904–911. doi:10.1037/0012-1649.35.4.904.

Kampa, Athina & Anna Papafragou. 2020. Four-year-olds incorporate speaker knowledge into pragmatic inferences. *Developmental Science* 23(3). e12920. doi:10.1111/desc.12920.

Kursat, Leyla & Judith Degen. 2020. Probability and processing speed of scalar inferences is context-dependent. In *Annual Conference of the Cognitive Science Society (CogSci) 42*, 1236–1242.

Levinson, Stephen C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Lewis, David. 1973. *Counterfactuals*. Oxford: Basil Blackwell.

Matsumoto, Yo. 1995. The conversational condition on horn scales. *Linguistics and Philosophy* 21–60.

Nadathur, Prerna. 2013. If...(and only if): conditional perfection and completeness.

Newstead, Stephen E. 1997. Conditional reasoning with realistic material. *Thinking & Reasoning* 3(1). 49–76. doi:10.1080/135467897394428.

Ohm, Eyvind & Valerie A Thompson. 2004. Everyday reasoning with inducements and advice. *Thinking & reasoning* 10(3). 241–272. doi:10.1080/13546780342000061.

Roberts, Craige. 1996. Information structure in discourse: Toward a unified theory of formal pragmatics. *Ohio State University Working Papers in Linguistics* 49. 91–136.

Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 1–69. doi:10.3765/sp.5.6.

Ronai, Eszter & Ming Xiang. 2021. Pragmatic inferences are qud-sensitive: An experimental study. *Journal of Linguistics* 57(4). 841–870. doi:10.1017/S0022226720000389.

van Rooij, Robert. 2004. Utility of mention-some questions. *Research on Language and Computation* 2(3). 401–416. doi:10.1007/s11168-004-1975-0.

van Rooij, Robert & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13. 491–519. doi:10.1007/s10849-004-2118-6.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and philosophy* 27(3). 367–391. doi:10.1023/B:LING.0000023378.71748.db.

Stalnaker, Robert C. 1968. A theory of conditionals. In *Ifs: Conditionals, Belief,* *Decision, Chance and Time*, 41–55. Dordrecht: Springer.

Zehr, Jeremy & Florian Schwarz. 2018. Penncontroller for internet based experiments (ibex).

Zondervan, Arjen, Luisa Meroni & Andrea Gualmini. 2008. Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. In *Semantics and Linguistic Theory (SALT) 18*, 765–777.

Ebru Evcen
Department of Linguistics
University of California, San Diego
9500 Gilman Drive #0109
La Jolla, CA 92093-0108
eevcen@ucsd.edu

David Barner
Department of Psychology
University of California, San Diego
9500 Gilman Drive #0109
La Jolla, CA 92093-0109
dbarner@ucsd.edu

Alan Bale
Linguistics Program
Concordia University
1455 de Maisonneuve W.,
Montréal, Québec Canada H3G 1M8
alan.bale@concordia.ca