

Experimenting with (Conditional) Perfection: Tests of the Exhaustivity Theory¹

Fabrizio Cariani & Lance J. Rips

1. Introduction

Conditional claims like (1) can, in the right context, convey additional information—the sort of information carried by (2) or (3):

- (1) If she turns in the final paper, she will pass the class.
- (2) Only if she turns in the final paper, will she pass the class.
- (3) If she does not turn in the final paper, she will not pass the class.

More generally, in the appropriate contexts, from an utterance of *If A, B* hearers can infer the biconditional *A if and only if B*. Following Geis and Zwicky (1971), we will call this inference *conditional perfection* (sometimes simply *perfection*, since it is the only kind of perfection we will be concerned with).

There is strong reason to think that perfection is not a purely logical inference—that is, it's not to be explained exclusively as a semantic entailment of conditional sentences.² After all, theories of conditionals universally agree that *If A, B* does not entail *If B, A* (and that it does not entail *A if and only if B*). Moreover, there are many conditional statements and contexts that do not trigger perfection.

- (4) If this cactus grows native to Idaho, then it's not an *Astrophytum* (Lilje 1972).

¹ The complete list of vignettes for the experiments reported in this paper is available at

<https://github.com/fcariani/perfection>

We thank Jonathan Cohen, Paul Égré, Michael Franke, Joshua Knobe, Paolo Santorio, Katrin Schulz, Eva Wittenberg, and Yimei Xiang for comments on this project. We also thank audiences at the 2016 Pacific APA, a conference on causal reasoning at the University of Amsterdam, the UIC Cognitive group, Northwestern Phling, Sembabble at UCSD, and the Gothenburg Logic Seminar. Kai von Fintel delivered comments on the paper at the 2016 APA that significantly shaped its later incarnations. For their assistance with the experiments, we thank John Glines, Gabrielle McCarthy, Jane Ko, Sabrina Kohanzad, Michael Yi, Kathy Tian, Camdyn Parrott, Joshua Kennedy, Alyssa Pierce, and Nathaniel Yoon. The final version is also indebted to three anonymous reviewers from a prior journal submission.

² We will later consider a view on which conditional perfection does arise as a semantic entailment, but only when additional covert material is present. In other words, even at that point, we remain committed to the view that it is not a semantic entailment of *the conditional*.

Experimenting with (Conditional) Perfection / 2

- (5) If you look at this Canaletto painting, you'll get a good idea of what the Canal Grande looks like (Herburger 2015).

It is natural, then, to explore the idea that perfection arises, when it does, from some kind of pragmatic reasoning. Perhaps this sort of pragmatic reasoning is systematic enough to be derivable from general pragmatic principles (that is without leveraging very specific assumptions about context).

In this paper, we present a series of experiments designed to test one of the most promising pragmatic accounts of perfection. This is von Fintel's (2001) idea that whether perfection arises depends on what kinds of questions are, implicitly or explicitly, driving the inquiry of the participants to the conversation (we present the essentials of this account in the next section). Because our findings are mostly negative, we emphasize at the outset that von Fintel's proposal is avowedly speculative. We are interested in testing it experimentally because it is a plausible working hypothesis. Even if it were not the whole story, it is worthwhile exploring which elements of it are supported by experimental scrutiny and which aren't.

Before proceeding, note that, while perfection is interesting in its own right as a problem for pragmatics, it has a wider significance. For instance, we believe that an account of conditional perfection might illuminate why participants in reasoning experiments are sometimes willing to endorse the fallacious inferences of *Affirming the Consequent* and *Denying the Antecedent*—the invalid cousins of *Modus Ponens* and *Modus Tollens*.³ Indeed, the link between perfection and these argument forms is at the center of our experiments.

2. From implicatures to exhaustivity.

Conditional perfection, we said, is not a purely logical inference. More generally, it seems wrong to maintain that it is the sort of inference that arises solely on the basis of the linguistic material that is explicitly made available by a conditional sentence. Geis and Zwicky (1971) illustrated the phenomenon of perfection with:

- (6) If you mow the lawn, I will give you five dollars.

No doubt, they had in mind a context in which (6) conveys the information that mowing the lawn is the only way for the hearer to get the five dollars. But it's easy to come up with contexts in which this is not true. For instance, imagine (6) being uttered immediately after:

³ The connection between conditional perfection and the fallacies is already made in Geis & Zwicky (1971, p. 562). For reviews of the experimental findings concerning the conditional fallacies, see Evans and Over (2004, chap. 3) and Oaksford and Chater (2007, chap. 5). For a proposal about how to account for the findings, given a story about perfection, see Cariani and Rips (2017).

Experimenting with (Conditional) Perfection / 3

(7) If you tidy your room, I will give you five dollars.

In this case, the surrounding discourse seems to block perfection. In general, linguistic context can determine whether the inference is licensed or blocked.

Another important aspect of conditional perfection is that there are multiple paths to generate a biconditional interpretation from a conditional (Van Canegem-Ardijns & Van Belle 2008). In particular, for a given conditional *If A, B*, we may note at least four distinct paths to perfection, depending on whether we strengthen with:

- converse: *if B, A*
- obverse: *if ~A, ~B*
- exhaustified conditional: *Only If A, B*
- exhaustified contrapositive: *Only If ~B, ~A*

As Van Canegem-Adrijns and Van Belle (2006) note, these paths are not equivalent (see also Bonnefon & Polizer, 2010; Franke 2009, p. 235). For instance, they claim that (8) invites the inference to (9) but not to (10).

- (8) If you empty a bucket of oil on the street, the street will get slippery.
- (9) If you do not empty a bucket of oil on the street, the street won't get slippery.
- (10) Only if you empty a bucket of oil on the street, will the street get slippery.

The focus in our experiments is on perfection inferences that go via converse and obverse.

So, what features of context license perfection? There is a long history of controversial pragmatic analyses.⁴ Fast-forwarding to one of the later stages of this debate, Horn (2000) proposes a view that (with interpretational help from von Stechow 2001) we understand as follows: conditional perfection is licensed when *If A, B* is in Gricean competition with *B, no matter what*. By uttering *If A, B*, a cooperative speaker conveys her inability to make the stronger statement that *B* holds unconditionally (that is, to assert: *B, no matter what*). That does license a hearer to infer something that goes beyond *If A, B*. But, as von Stechow (2001) points out, that license is well short of perfection. Even if one's assertive utterance of *If A, B* signals reluctance to assert *B, no matter what*, it does not rule out another antecedent *D*, not equivalent to *A*, such that *If D, B*. So it does not imply that *A* is the only such antecedent (i.e., that *A* is necessary as well as sufficient for *B*).

The exact upshot of Horn's proposal depends on the background theory of conditionals. Let us adopt one theory as a starting point. Suppose that conditionals are universal quantifiers over a contextually set domain of possible worlds.

⁴ In addition to Geis and Zwicky (1971), see van der Auwera (1997), Horn (2000), and von Stechow (2001).

Experimenting with (Conditional) Perfection / 4

Strict: *If A, B* is true in context *C* at world *w* iff for all $v \in C(w)$, either *A* is false at *v* or *B* is true at *v*

Here, $C(\cdot)$ denotes a function from worlds to sets of worlds. Intuitively, this is interpreted as the set of worlds that are relevant to the evaluation of the conditional, given that the world of evaluation is *w*.

Given this semantics, the most likely explanation for one's not being in a position to assert *B, no matter what* would be that one believes that *B* might be false in some of the relevant worlds. But if the additional information conveyed by *If A, B* is just that *B* might be false in some salient possibility, we are far from perfection. Perfection would require not just that *B* fail at some of the relevant worlds at which *A* fails; it requires the truth of the converse conditional (i.e., *If B, A*), and so it requires that every relevant *B*-world is an *A*-world.⁵ Taking stock: Gricean competition between *If A, B* and *B, no matter what* seems to get at a real phenomenon. But that phenomenon is not perfection.

Why, then, is it sometimes possible to get the full perfection inference? Von Fintel (2001) sketches a different sort of pragmatic account, following an insight in Cornulier (1983). Commenting on the example,

(11) One can take this seat if one is disabled or one is older than 70,

Cornulier remarks:

For we can suppose, very roughly, that in [*One can take this seat if one is disabled or if one is older than 70*] the word *if* keeps its merely sufficient condition meaning, and that the utterance situation suggests that if other sufficient conditions (allowing one to sit there) did exist, they would have been mentioned, so that the only mentioned property (to be disabled or to be older than 70) is the only property which gives one the right to sit there (presumption of exhaustivity). (Cornulier, 1983: 248)

Incidentally, Cornulier's example is especially interesting because it is a conditional permission, which might actually trigger a particularly strong form of perfection inference. (See Section 3 for discussion; however, our experiments did not target conditional permissions).

Von Fintel's (2001) central move is to connect Cornulier's talk about exhaustivity with off-the-shelf work on exhaustivity in response to questions (specifically Groenendijk & Stokhof, 1991; for a survey of work on exhaustivity see van Rooij & Schulz 2003; for another development of Cornulier's insight, see Franke 2009). The key assumption is that we generally, though not always, interpret simple answers to questions as exhaustive. If the doctor asks, "What did you

⁵ Or at least, minimally, that both *A* and *B* fail at the actual world. The Horn-inspired proposal fails this more minimal standard as well.

Experimenting with (Conditional) Perfection / 5

drink last night?” and Lucy replies “Two glasses of wine,” we take Lucy to provide an exhaustive list of what she drank.

This idea can be extended to conditionals. Consider a question Q such that (i) $If A, B$ is a possible answer to Q and (ii) when $if A, B$ is provided as an answer to Q , the conditional is naturally given an exhaustive interpretation. There are many different questions conditionals might be used to answer. It will be useful to isolate two categories of questions: the first category consists of questions about how the consequent might come about. We call them *consequent-directed* (abbreviated [CONS?]). Here are some examples of [CONS?] questions:

- What are all the ways in which B ?
- How might B happen?
- Is B true?

The second category consists of questions about what follows from the antecedent.

- What follows from A ?
- What happens if A ?

To make this concrete, consider an example like (6). This might be an answer to [CONS?] questions like: “What are all the ways in which I might get five dollars?”, “Will I get five dollars?”; it might also be an answer to [ANT?] questions like “What happens if I mow the lawn?”. Moreover, we should not suppose that these are the only questions: [CONS?] and [ANT?] questions are not exhaustive categories.

Von Stechow's claim is that [CONS?] questions, but not [ANT?] ones, might help yield something like perfection. His argument starts, like ours, with the assumption that the truth-conditions for $If A, B$ are captured by **Strict**.

Here, then is one possible way of deriving perfection given the assumptions we have on the table. Suppose that a [CONS?] question Q_{cons} was asked and that it was answered by $If A, B$. Then:

Step 1: provided that $If A, B$ is understood as an exhaustive answer to Q_{cons} , the speaker is not in a position to assert $If D, B$ for any D that competes with A .

Step 2: provided that the speaker is informed about the truth-values of these conditionals, she must believe all conditionals of the form $If D, B$ are false.

Step 3: if all conditionals of the form $If D, B$ are false for every antecedent that competes with A , then *if not A, not B* must be true.

Experimenting with (Conditional) Perfection / 6

To justify the reasoning from Step 2 to Step 3 we need another assumption. One approach might stipulate that the space of competitors to A is particularly rich:

Competitors: the antecedents that compete with A are all those antecedents D that do not entail A .

This entails that there are *lots* of alternatives. In particular, it entails that for each *not* A world w , there is an alternative conditional (*If* S_w , B) where S_w is a sentence that is only true in w .⁶

Another way of justifying this step, a more plausible one in our view, is to assume (i) that conditionals satisfy the principle of Conditional Excluded Middle (CEM) so that from $\sim(\textit{If } D, B)$ one might infer *If* D , $\sim B$ and (ii) that every relevant possibility is included in some competitor or other. We think that this second approach fits best with the rest of von Stechow's theoretical framework—as von Stechow's (1997, section 7) argues that **Strict** truth conditions for *If* can be made compatible with CEM by hypothesizing that conditionals come with a homogeneity presupposition.⁷

Either way, von Stechow's account predicts that conditional perfection should arise precisely when the other assumptions in the above reasoning are satisfied. We take **Strict** to be a valuable working hypothesis, so the key assumptions to focus on, as we move to the experimental part of our paper are: (i) that there is some principled way in which the relevant competitors to A are generated; (ii) that *If* A , B is understood as an exhaustive answer to [CONS?] questions; and (iii) that the speaker is relevantly informed.

3. Designing experimental tests of perfection inferences.

Before going through the battery of experiments we ran, it's useful to describe informally how we set about experimenting with perfection. There are many decisions to be made in setting up such an experiment, and small variants might result in significant changes.

⁶ Concluding the reasoning: asserting *If* A , B would implicate $\sim(\textit{If } S_w, B)$, which entails that there is a world that verifies S_w & $\sim B$. But since only w verifies S_w , then w must verify $\sim B$. Since w was an arbitrary world that does not verify A , it follows that every $\sim A$ -world is a $\sim B$ world.

⁷ Of course, the canonical way of validating CEM is to adopt Stalnaker's semantics for conditionals (Stalnaker 1968, 1981): *If* A , B is true at w iff B is true at the selected A -world. Yet another option, one that we have some stake in, starts with the argument in Cariani and Santorio (2018) that *will* is a 'selectional' modal (that its contribution is to select a world out of a modal base). They note that combining a restrictor semantics in the style of Kratzer (1991, 2012) with selectional modals gives something roughly like Stalnaker's semantics for conditionals of the form *If* A , *will* B —specifically, a semantics that validates CEM. One might extend this insight to a broader class of conditional sentences by postulating that conditionals can sometimes restrict covert selectional modals (Cariani, 2021).

Experimenting with (Conditional) Perfection / 7

One guiding idea that motivated us is that if perfection arises, we should see unusually high endorsement rates for the conditional fallacies of *Affirming the Consequent* (AC) and *Denying the Antecedent* (DA) (see Geis & Zwicky, 1971).

(AC) *B, If A, B. Therefore: A*

(DA) *not A, If A, B. Therefore: not B.*

The idea here is that if, in context, *If A, B* conveys *If B, A* then, barring interference, the endorsement rate for AC should approach the endorsement rate for Modus Ponens. Similarly, if *If A, B* conveys *If not A, not B*, then DA should also approach Modus Ponens.

We have chosen to test endorsement rates for conditional inferences partly because we could rely on a wealth of established and very robust data concerning people's endorsement of these patterns. Figure 1 summarizes endorsement rates for MP, MT, AC, and DA from earlier experiments. Incidentally, it reveals one of the important discoveries in the psychology of reasoning: The endorsement rate of Modus Ponens is higher than the endorsement rate of Modus Tollens (even for bare, non-modalized conditionals).⁸

We could have chosen a different approach had we decided to test whether people endorse the inference from *If A, B* to *If not A, not B*. As we just noted, our choice of two-premise arguments (e.g., *If A, B* and *Not A to Not B*) allowed us to use the large set of earlier results, summarized in Figure 1, as a baseline for our manipulations. One could argue that the one-premise approach is a more direct test of whether people perfect conditionals. However, we know of no evidence to suggest that people find the one-premise arguments easier to think about than their two-premise counterparts. For example, the rate of endorsement for the inference from *If A, B* to *If not A, not B* (for "contingent inferences," such as "If the mushroom is red, it is edible," in Fillenbaum, 1975) is within the range of endorsements for DA arguments in the studies summarized in Figure 1. While our data are primarily of the two-premise kind, we recognize both kinds of data as significant to an account of conditional perfection. We report evidence of the one-premise type in Experiment 7.

⁸ See the references mentioned in Note 3 for reviews of this evidence. The data in Figure 1 come from Evans, Newstead, and Byrne's (1993) review of earlier studies of conditional inference (their Table 2.4). As Evans et al. state, the participants were adults (typically college students). They were "normally given the premises and conclusions and asked whether it follows, or else given a list of conclusions including the normal one and 'nothing follows' to choose from. All these studies involve either so called 'abstract' problem material or ones which are concrete but arbitrary, so that prior beliefs and pragmatic associations are not likely to be cued." (pp. 35-36). Recall that our approach is to test the endorsement rates of AC and DA. When perfection is triggered, these inferences aren't fallacies, but actually valid applications of modus ponens or modus tollens. Thus, when we think about, say, AC, there are two possibilities that might lead to different endorsement rates:

- i. people infer from *If A, B* to *If B, A* and then apply MP to *If B, A* and *B*.
- ii. people infer from *If A, B* to *If not A, not B* and then apply MT to *If not A, not B* and *B*.

This difference might matter to our project, and we should be mindful of it in interpreting our results.

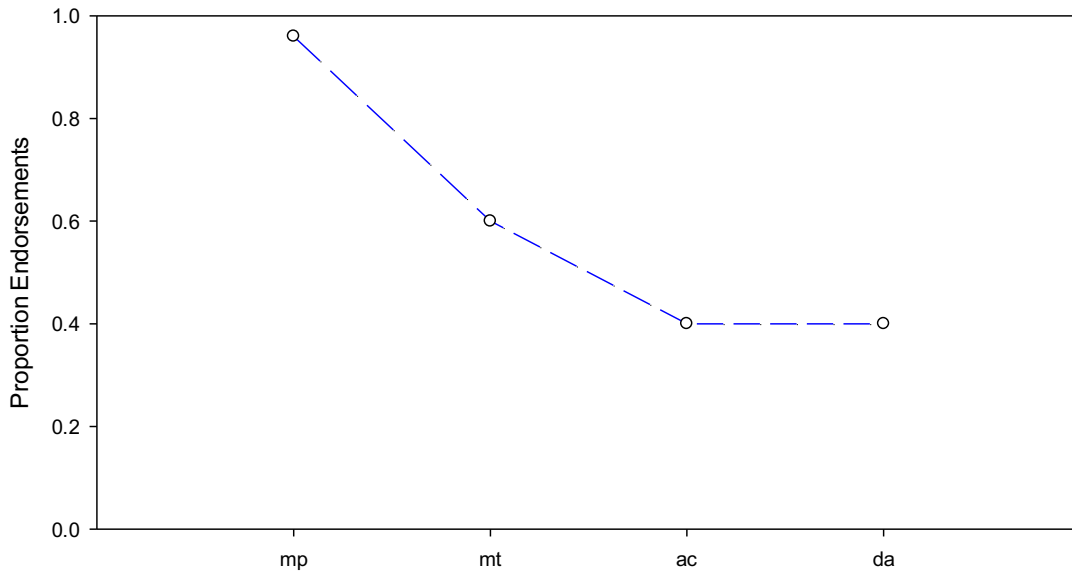


Figure 1. Standard endorsement rates for the four inference patterns modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da) across previous experiments (data from Evans, Newstead, & Byrne, 1993, Table 2.4 and are weighted averages from seven earlier experiments).

We note that psychological research provides many examples in which college-aged participants endorse AC and DA at rates greater than those that appear in Figure 1. The Figure 1 experiments used conditionals for which people do not have strong beliefs in the necessity of the antecedent for the consequent. But if prior beliefs *do* suggest (causal) necessity, participants find AC and DA more congenial (e.g., Cummins, Lubart, Alksnis, & Rist, 1991; Marcus & Rips, 1979; Staudenmayer, 1975; Thompson, 1994). For example, a conditional like *If the butter is heated, then it melts* accords with people's belief that heating is both necessary and sufficient for the butter melting. It's difficult to think of other ways to melt butter aside from heating it. So an AC or DA argument with such a conditional appears correct to many participants (e.g., *If the butter is heated, then it melts; the butter melts; therefore, the butter was heated*). Effects of this sort are similar to those of conditional perfection in that they encourage a reading like that of *A iff B*, but they arise for reasons other than the pragmatic considerations that we have discussed so far. Because our intent in these experiments is to test the pragmatic account of Section 2, we chose conditionals with antecedents that don't already suggest necessity.⁹

⁹ For related reasons, the research we describe here is neutral with respect to probabilistic theories of conditionals. A number of theorists have proposed that whether a conditional sentence is appropriate depends on the conditional probability of its consequent given its antecedent (e.g., Adams, 1965; Evans & Over, 2004). If these theories are correct, then we should expect conditional perfection when both the conditional probability of the consequent given the antecedent and that of the antecedent given the consequent are sufficiently high (as they are in the butter heating example). In addition to constraining the semantics to a lesser degree, the pragmatic theory we explore here maintains that whatever the success of these probabilistic accounts, a further source of conditional perfection is a

Experimenting with (Conditional) Perfection / 9

A more pertinent set of studies have examined conditional promises (e.g., (6)-(7)) and threats (e.g., *If you continue to disrupt the class, you'll have to leave the room*). These studies have shown greater acceptance of AC and DA for promises and threats than for ordinary indicatives (e.g., *If there is an electrical failure, school will be closed*, Markovits & Lesage, 1990; *If the student is doing economics then he is a socialist*, Newstead, Ellis, Evans, & Dennis, 1997). Similarly, participants are more willing to accept the inference from *If A, B* to *If not A, not B* under the same circumstances (Fillenbaum, 1975). The goal of a promise or threat is usually to get the addressee to perform some action (e.g., mow the lawn) or to refrain from one (disrupting the class). For these speech acts to be effective, speakers presumably intend not to provide the promised reward if the action is not taken and not to carry out the threatened punishment if the infraction is not committed. These presumptions can be overridden in the right circumstances, as we've already noted in Section 2. But by default, promises and threats convey an exhaustive interpretation, as the experiments we've just cited suggest. Our aim in the present experiments is to manipulate participants' impression of exhaustivity by means of [CONS?] questions, so we picked conditionals for our experiments that (unlike promises and threats) don't by themselves convey perfection.

The second guiding idea in shaping our tasks was that we needed to create matched items for purposes of experimental control. One version of each item used [CONS?] questions; the other either used no question at all or [ANT?] questions—depending on the experiment. In the following, we call this variable *question type*. Here is an example of one of our [CONS?] items (*modulo* some differences in exact wording across experiments):

John has taken a test on Chapters 4-6 that has not been graded yet.
[You ask Mary, "Did John do well on the test?"]
Mary says, "If John understood Chapter 5, then John did well on the test."
Assume that Mary's response is true and that John did well on the test. Given this information, then, does Mary's statement imply that John understood Chapter 5?

When the [CONS?] question was a polar question (i.e., a yes/no question), the matching item did not have a question at all (so it is simply the result of removing the bit in square brackets).

Other experiments involved more complicated [CONS?] questions, such as:

You ask Mary, "What are all the ways John could manage to do well on the test?"

Here, the matching item replaced the line in which we asked the [CONS?] question with a line in which we asked the [ANT?] question:

conversational demand for exhaustivity, which can perfect conditionals even if the conditional probability of the antecedent given the consequent is not initially high.

Experimenting with (Conditional) Perfection / 10

You ask Mary, “What are all the things that could happen in case John understood Chapter 5?”

Note also that the sample item above is testing for Affirming the Consequent. For each vignette (there were 16 of them), we had items that tested Modus Ponens, Tollens, Affirming the Consequent, and Denying the Antecedent.

In total, this means that we associated each vignette with eight possible items (four inferences for each of the two possible questions). Participants saw the vignettes presented one-at-a-time on a computer screen, in a new random order for each participant. They responded by clicking on one of two options (e.g., “implies” vs. “does not imply”).

Participants were college students enrolled in an introductory psychology course, and they completed the experiment as part of a course requirement. No participant took part in more than one experiment.

Note that [CONS?] questions like “What are all the ways John could manage to do well on the test?” appear more complex, at least syntactically, than other possible [CONS?] questions, such as “Did John do well on the test?”. In most of our experiments, we have privileged the lengthier question because it reduces the permissibility of *mention-some* answers, which are partial answers. Consider the question:

(12) Q: Where can I buy Stephen King novels?

A: At Powell’s Books.

In (12), we do not expect the answer to be an exhaustive catalog of the places where the questioner can buy Stephen King novels. It is important to avoid this interpretation, for the account of perfection we sketched above breaks down if the answer is understood to be partial.

To have a better chance of ruling out *mention-some* answers (that is: for the question to set up the presumption that any answer would be exhaustive), we might ask a different sort of question, such as:¹⁰

(13) Q: What are all the places where I can buy Stephen King novels?

A: At Powell’s Books.

This is why most of our experiments use “What are all the ways in which *B*?” as the [CONS?] question. The Online Appendix lists all 16 vignettes with the full [CONS?] and matched [ANT?] questions for the problems testing Denying the Antecedent. These are the items as they

¹⁰ Perhaps, even the question in (13) sometimes permits a *mention-some* answer. We discuss the implications of this point in Section 5.

appeared in Experiment 2. We mention experimental variations in wording as they come up in the relevant parts of Section 4.

While this is a good reason to run experimental tests with the more complex phrasing, it is not a good reason *not to* test the more natural [CONS?] questions. It is for this reason that our leading experiment involves a simple polar question, as in our initial example above (in which “you” ask Mary, “Did John do well on the test?”).¹¹

4. The sequence of experiments.

A quest for perfection drives the experiments we report here. In each of them, we present participants with a series of problems, variations on the standard set that we described earlier. These problems vary the inference type (Modus Ponens, Tollens, Affirming the Consequent, and Denying the Antecedent) and question type ([CONS?] and [ANT?]) in the hope that a demand for an exhaustive set of reasons for the consequent—[CONS?] but not [ANT?—will lead participants to perfect the conditional. According to the exhaustivity hypothesis, perfection should be manifested by increased endorsement of the “fallacies,” Denying the Antecedent and Affirming the Consequent. As mentioned, these arguments switch from invalid to valid under a perfected conditional, which should increase the likelihood that participants will accept them. The same manipulation, however, should have only a weak effect, if any, on the endorsement of Modus Ponens and Modus Tollens, since these inferences are already valid under the “unperfected” reading of the conditional and remain so under the perfected reading. This, then, is the perfection pattern of our quest: Greater endorsement of Affirming the Consequent and Denying the Antecedent under [CONS?] questions but not [ANT?] questions, but little change in endorsement of Modus Ponens and Modus Tollens.

Experiment 1: Polar Questions

We start with our barest experiment.¹² As described, this experiment straightforwardly contrasted a polar [CONS?] question with the same item with the question removed.

John has taken a test on Chapters 4-6 that has not been graded yet.

[You ask Mary, “Did John do well on the test?”]

Mary says, “If John understood Chapter 5, then John did well on the test.”

Assume that John did not do well on the test.

¹¹ The reason why we did not use an [ANT?] question as a comparison in this case is that it is implausible to target the antecedent with one such question (say, “Did John understand Chapter 5?”). Except for some *recherche* contexts, the conditional “If John understood Chapter 5, then John did well on the test” is not an acceptable answer to the question: “Did John understand Chapter 5?”.

¹² Although we lead with this experiment, it is not temporally the first experiment we ran. It was suggested to us by Kai von Fintel as a simplification of some later experiments.

Experimenting with (Conditional) Perfection / 12

Given this information, does Mary's statement imply that John did not understand Chapter 5?

The design of the experiment included the standard set of items that we described earlier. The Online Appendix lists the full set of vignettes in their Denying the Antecedent versions. Although a given participant saw a specific vignette only once (i.e., in only one of its eight versions), a participant saw each combination of inference type and question type, instantiated in two different vignettes. So a participant received 16 trials in all. Across participants, each vignette version appeared equally often. We tested 32 participants from the population mentioned earlier, choosing this number based on the earlier experiments summarized in Figure 1 (see Evans et al., 1993, Table 2.4). The experiments we report later share this same structure.

At the beginning of the session, a computer presented written instructions informing participants that they would see on each trial several sentences about a particular conversation. They were asked to read the problem carefully and then to ask themselves "whether the last of these statements is implied within the conversation." They clicked on a button labeled "implied" or on one labeled "not implied" to record their decision. We did not provide an explication of "implied," since we are interested in participants' natural understanding of whether the conclusion of the inferences followed from the information in the vignette. "Implied" has an ordinary language sense—roughly the dictionary sense of conveying something indirectly—that includes both the semantic and pragmatic components that are at stake in theories of perfection.

The proportion of "yes" responses showed no difference between the [CONS?] question and no question at all. As Figure 2 illustrates, we did find the typical difference in endorsements as a function of inference type, $F(3,24) = 11.35, p < .001$. As the figure suggests, however, we found no reliable overall difference due to the [CONS?] question ($F(1,15) < 1$), and no differential effect of the [CONS?] question on the rate of endorsement for the individual inference types ($F(3,24) < 1$).¹³

¹³ The statistical tests in this and the following experiments are based on a generalized linear mixed model for binomial data. The models were "maximal" (in the sense of Barr, Levy, Scheepers, & Tily, 2013) in including as random effects: (a) the main effects of participants and vignettes, and (b) all interactions of participants and vignettes with the fixed effects of interest (question and inference type).

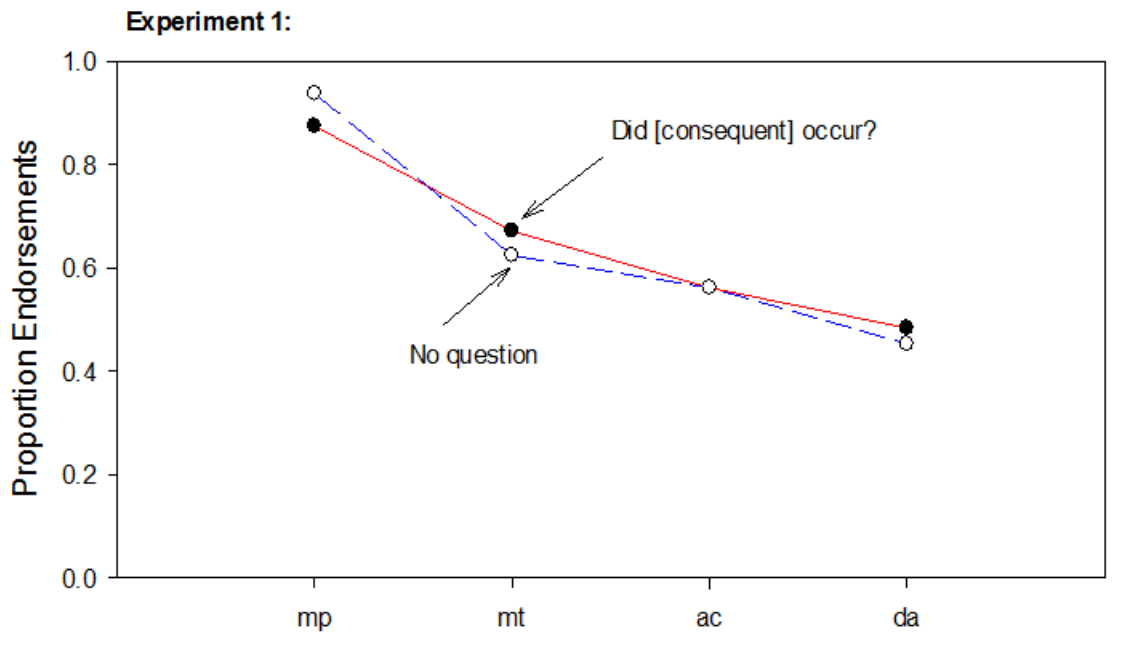


Figure 2. The effect of polar question vs. no question on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 1.

We also checked whether any of the individual vignettes showed an effect of the [CONS?] question. The analysis used the data from just Affirming the Consequent and Denying the Antecedent, since the theory under consideration predicts a difference only for these. However, we found no overall effect of vignettes and no interaction of the vignettes and the [CONS?] versus no question variation ($F(15,224) < 1$ in both cases). Planned comparisons found no effect of question for any of the individual vignettes (for the largest of these, $F(1,224) = 2.16, p = .14$).¹⁴

Experiment 2: Explicit Demands for Exhaustive Answers

Experiment 2 involved the more complicated [CONS?] questions, ones that attempt to emphasize the demand for exhaustivity in the question to Mary:

John has taken a test on Chapters 4-6 that has not been graded yet.
 You ask Mary, "What are all the ways John could manage to do well on the test?"
 Mary responds, "If John understood Chapter 5, then John did well on the test."
 Assume that John did well on the test.
 Given this information, does Mary's statement then imply that John understood Chapter 5?

¹⁴ This analysis was similar to the one just reported but treated vignettes and their interaction with question as fixed effects rather than as random effects. Denominator degrees of freedom were estimated using the Satterthwaite method.

As mentioned, the [CONS?] question was intended to make *mention-some* answers as impermissible as possible. The [ANT?] question was also introduced to make [ANT?] items parallel to the [CONS?] items, controlling for the length and complexity of the vignettes. Specifically, in the “Chapter 5” case, we used the [ANT?] question: "What are all the things that could happen in case John understood Chapter 5?".

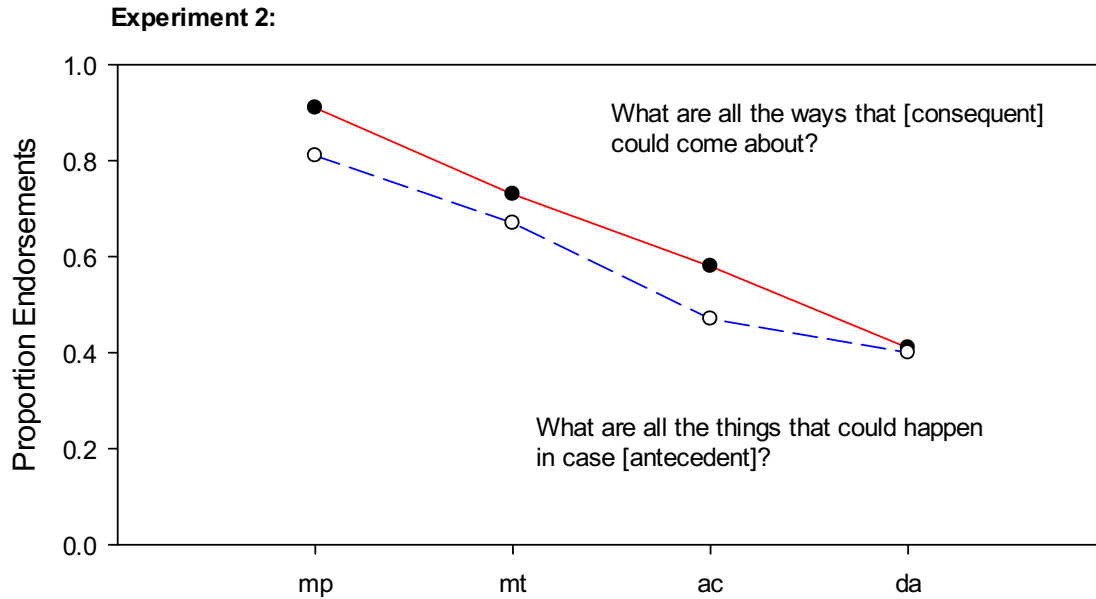


Figure 3. The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 2.

None of these changes, however, led to greater endorsement of Affirming the Consequent and Denying the Antecedent for [CONS?] questions relative to [ANT?] questions. Figure 3 graphs the data from 32 participants and shows the usual decrease in endorsement rates from Modus Ponens to Tollens to Affirming the Consequent to Denying the Antecedent, $F(3,24) = 12.46$, $p < .001$. However, we found no significant difference favoring problems with [CONS?] over those with [ANS?] ($F(1,15) = 2.73$, $p = .120$), and no significant difference appeared in the shape of these two functions, $F(3,24) < 1$. Explicit demands for an exhaustive answer to how the consequent could occur do not seem to elevate Affirming the Consequent and Denying the Antecedent to a greater extent than they did Modus Ponens or Modus Tollens.

Experiment 3: Checks for Memory for the Questions

Could it be that our participants were ignoring the question? After all, one could perform the task while skipping the question “you” are asking Mary and reading just the conditional and the minor premise. That might explain our difficulty in obtaining the perfection pattern.

To explore this hypothesis, we ran a variant of Experiment 2. After answering each item, participants viewed a new screen that prompted them to recall which question, [CONS?] or [ANT?], they had seen earlier, and they picked one of them by clicking on it. For the sample vignette, the choice was between “What are all the ways John could manage to do well on the test?” and “What are all the things that could happen in case John understood Chapter 5?” If participants could not remember what question had been asked, we inferred that the question did not play a role in their reasoning. Although they could correctly recognize the question without using it in reasoning, a correct answer would at least suggest that the question was available for them to reason with. In addition, since participants saw multiple items, this forced choice encouraged them to attend to the question as they were working through the later items in their allotted sequence.

Alas, the results did not fit the perfection pattern. Figure 4 plots the proportion of endorsements from 33 participants after we removed all trials on which the participants made a memory error (i.e., selected the [CONS?] question when they had actually seen [ANT?] or the reverse error). These errors ranged from 9% to 18% across the eight conditions shown in the figure. As the figure suggests, we found the typical effect of inference types, $F(3,23) = 10.42$, $p < .001$. But neither the effect of the question nor the interaction between the question and the inference types was statistically reliable (both F 's < 1).

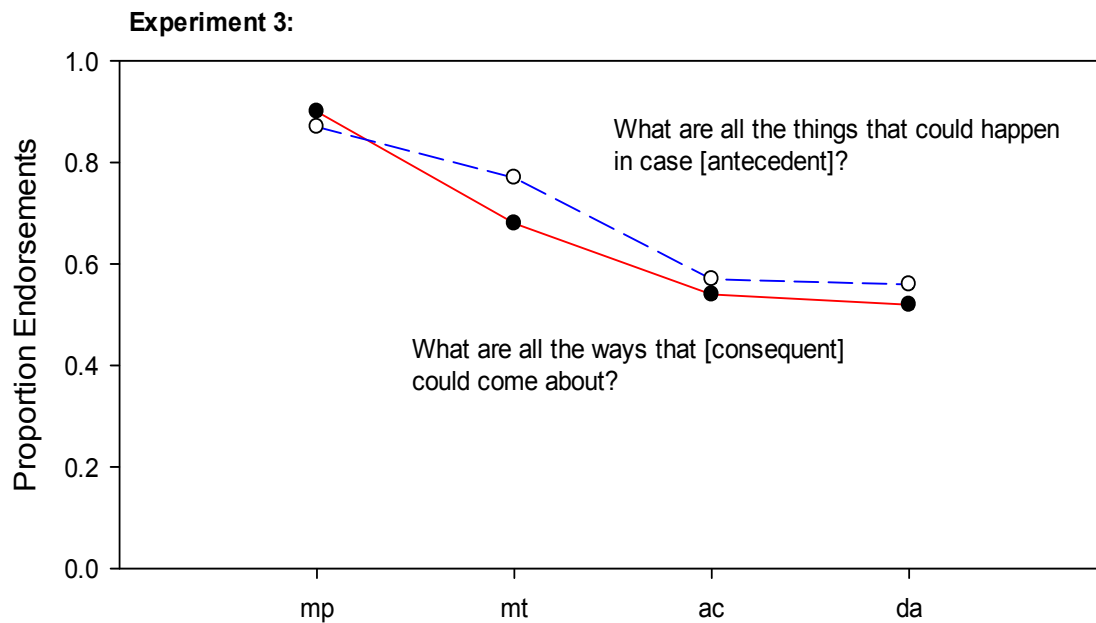


Figure 4. The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da). Experiment 3, with memory checks.

Given these results, one might become quite skeptical. Either there is a problem in the theoretical proposal or in the particular way we had sought to test it.

Experiment 4: Explicit Specification of the Antecedent as the Only Condition

Are there cases where we *do* get the perfection pattern using a similar experimental set up? The next idea, then, was to get very close to forcing a biconditional reading by semantic means. In the items for the new experiment, Mary does not just answer the question with the conditional: she overtly asserts information corresponding to one of the paths to perfection (to make things easier, we underlined the new bit below: it was not underlined in the stimuli presented to participants).

John has taken a test on Chapters 4-6 that has not been graded yet.

You ask Mary, "What are all the ways John could manage to do well on the test?"

Mary responds, "If John understood Chapter 5, then John did well on the test.

That is the ONLY way John could have done well on the test."

Assume that John did well on the test.

Given this information, does Mary's statement imply that John understood Chapter 5?

In the [ANT?] variant, in addition to the usual difference in question (i.e., "What are all the things that could happen in case John understood Chapter 5?"), Mary's response changes to:

Mary responds, "If John understood Chapter 5, then John did well on the test.

That is the ONLY thing that could have happened if John understood Chapter 5."

Note that the continuation in these [ANT?] variants does not go any distance towards conveying biconditional information (for it does not rule out that John's doing well on the test might have come about by some other means).

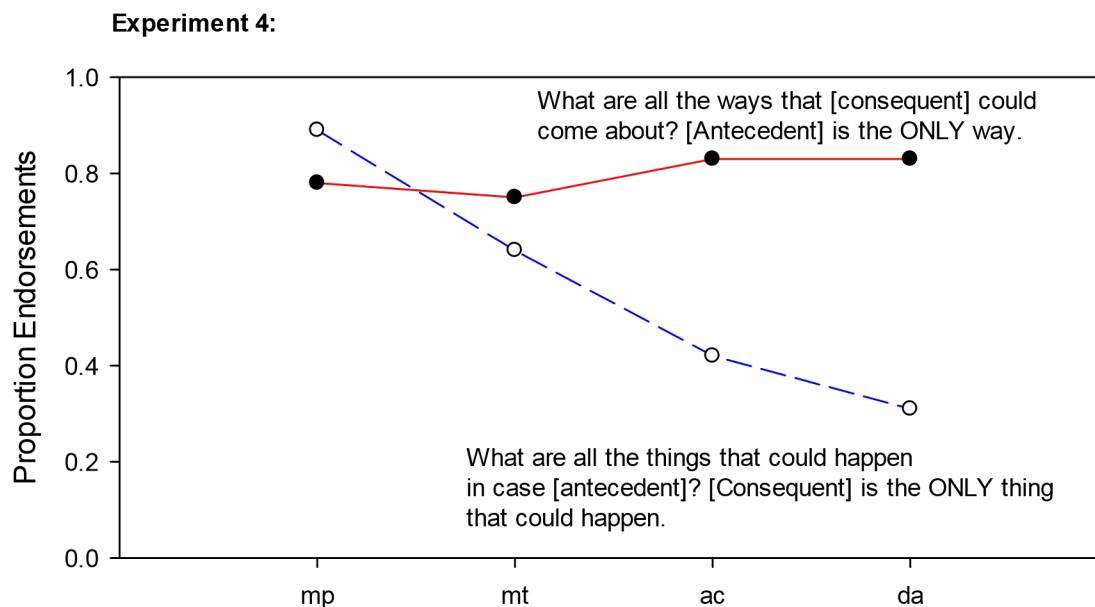


Figure 5. The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 4, with explicit indication of exhaustivity.

At last, these explicit changes did yield a perfection pattern, as shown in Figure 5 (based on 32 participants). Not only did we find a significant effect of inference type ($F(3,24) = 4.70, p = .010$), we also found a significant effect of question type ($F(1,15) = 13.38, p = .002$), and crucially an interaction of the two ($F(3,24) = 10.56, p < .001$). This last effect is the perfection result, apparent in the difference in the shape of the curves in the figure.

Essentially, if we explicitly provide the strengthening that is supposed to be conveyed by pragmatic means, we get precisely the pattern we would expect. This suggests that our procedure is sensitive to (at least some kinds of) information that can get participants to interpret a conditional as a biconditional.

Experiment 5: Speaker's Knowledge of the Answers and Willingness to Relate Them

The biconditional interpretation (predictably) arises if we have an explicit continuation with *only*, as in Experiment 4. So why were we unable to find a similar pattern in Experiments 1-3? We have already attempted to rule out that the questions were not sufficiently clear in their demand for an exhaustive answer and that participants were inattentive to the questions. Another possibility is that we have not completely eliminated a "mention some" answer. Although our [CONS?] question asks Mary for all the ways the consequent could come about, her response may reflect only her partial knowledge of these ways or her limited willingness to produce them. Participants may have interpreted Mary's statement in a way that allows for

Experimenting with (Conditional) Perfection / 18

these constraints. If so, they may have doubted whether her answer does indeed convey all the ways the consequent could happen.

To encourage participants to think that Mary's response was exhaustive, we tried "loading up" the context. The items in Experiment 5 looked like this (we underlined the new bit; it was not underlined in the text that participants saw):

John has taken a test on Chapters 4-6 that has not been graded yet.

You ask Mary, "What are all the ways John could manage to do well on the test?"

In fact, Mary knows all the ways and is willing to relate them. Mary responds, "If John understood Chapter 5, then John did well on the test."

Assume that John did well on the test.

Given this information, does Mary's statement imply that John understood Chapter 5?

The same underlined information was also inserted in the [ANT?] version of the problem.

We expected that stipulating that Mary knows all the ways in which John could do well on the test and is willing to relate them would have a similar effect to saying explicitly that [Consequent] is the ONLY thing that could happen.

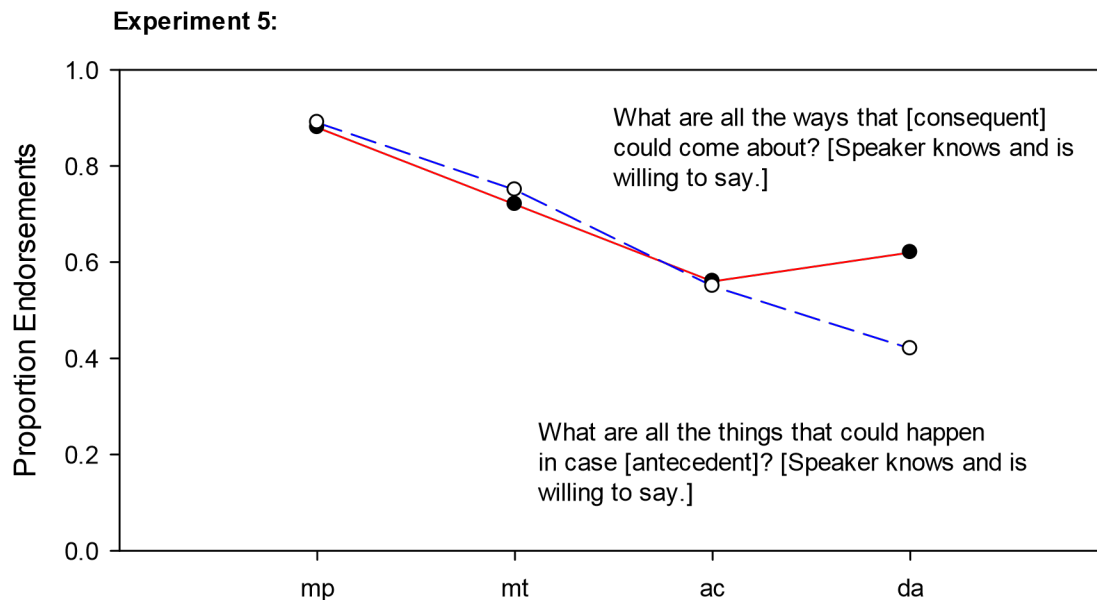


Figure 6. The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 5, with information about speaker's knowledge and willingness.

However, this expectation was not met. Endorsement rates for the four inference types again differed significantly ($F(3,24) = 8.50, p < .001$), based on data from 32 participants. But neither the overall difference between questions ($F(1,15) < 1$) nor the interaction between question and inference type ($F(3,24) = 1.52, p = .235$) are significant. Figure 5 shows a trend toward higher

Experimenting with (Conditional) Perfection / 19

endorsement of Denying the Antecedent for the [CONS?] questions, but this difference was not fully significant, $F(1,24) = 3.93$, $p = .059$. Moreover, the [CONS?] question did not affect endorsement rates for Affirming the Consequent ($F(1,24) < 1$). Note, too, that even for Denying the Antecedent, the boost in endorsement rates is smaller than what we have seen in the full-blown perfection of Experiment 4.

Experiment 6: Explicit Specification of Exhaustiveness

Even when participants know that the speaker of a conditional is (a) under explicit pressure to produce an exhaustive answer to the question of how the consequent could come about, and (b) knows all the ways it could come about and is willing to relate them, they do not produce the full perfection pattern. This failure may be due to a residual unwillingness on the participants' part to believe that the speaker really has produced all the ways. Although Mary may know all the ways and is willing to tell you about them, she may nevertheless give you just a sample, perhaps because the list is too long, too complex, or too unrelated to present concerns. These considerations suggest that we might be able to reinstate full perfection if Mary explicitly states that she is in fact giving all the ways when she asserts the conditional.

To check this prediction, we used the vignettes from Experiment 5, but added Mary's assertion that she was giving all the ways the consequent could come about:

John has taken a test on Chapters 4-6 that has not been graded yet.

You ask Mary, "What are all the ways John could manage to do well on the test?"

In fact, Mary knows all the ways and is willing to relate them. Mary responds, "Here are ALL of them: If John understood Chapter 5, then John did well on the test."

Assume that John did well on the test.

Given this information, does Mary's statement imply that John understood Chapter 5?

The corresponding [ANT?] version likewise included the "Here are ALL of them" prefix in Mary's answer.

The results from 32 participants appear in Figure 7 and show that the new "ALL of them" clause was enough to produce the perfected interpretation.

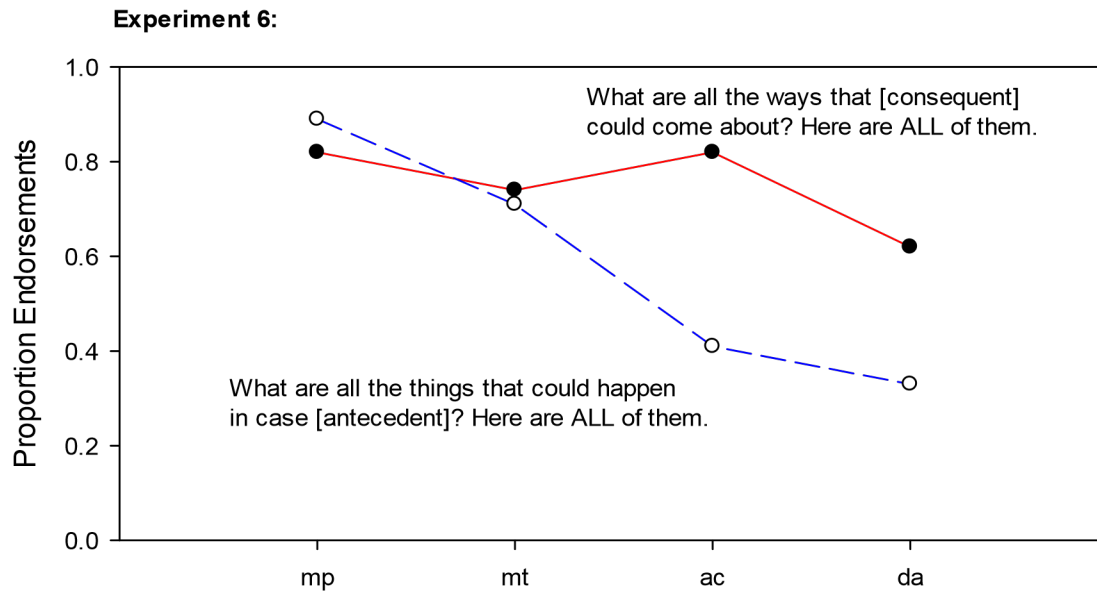


Figure 7. The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 6, with explicit exhaustivity.

When the speaker was asked to give all the ways the consequent could occur, and the speaker then made it clear that she was providing all the ways, participants were likely to endorse each of the inference types. When the speaker was asked about all the things that could happen if the antecedent occurred, however, we find the typical decreasing pattern across inference types. This produced significant effects of inference type ($F(3, 24) = 8.29, p < .001$), question type ($F(1, 15) = 9.67, p = .007$), and an interaction between the two ($F(3, 24) = 6.80, p = .002$). Unlike the results of Experiment 5, the difference due to the question is significant for both Affirming the Consequent ($F(1, 24) = 22.60, p < .001$) and Denying the Antecedent ($F(1, 24) = 9.81, p = .004$).

Experiment 7: Conditional Transformations

The studies we've reported so far suggest that demands for exhaustive answers fail to produce perfection. We get the expected pattern only when the speaker explicitly states that her answer is exhaustive. A gap seems to exist between requesting an exhaustive response and accepting the response as exhaustive. This gap may have been widened, though, by some aspects of our procedure. For one thing, the question we put to participants was whether the conditional statement (e.g., "If John understood Chapter 5, then John did well on the test"), together with the given assumption (e.g., "John did well on the test") implied the conclusion ("John understood Chapter 5"). But one might argue that pragmatic effects of the kind we're seeking depend on people's understanding of what the speaker implied rather than what her statement implied. Emphasis on the statement might have led participants to think that only what explicitly appears in that statement matters for the inference. This would help explain the difference between Experiments 4 and 6, where the speaker claims her answer is exhaustive, and Experiments 1-3 and 5, where she doesn't.

A second reason why the latter experiments may have failed to produce perfection has to do with the type of inference we asked participants to assess. As in traditional experiments on conditional reasoning, participants had to integrate information from the speaker's conditional statement with the minor premise that we asked them to assume. Combining these statements may have required cognitive effort that suppressed the effect of interest. Of course, this complexity would not explain the obtained differences between endorsement rates among the four inference types, since all used the same format. For the same reason, complexity would not explain the differences between Experiment 1-3 and 5, on one hand, and Experiments 4 and 6, on the other. However, it is possible that complexity masked a weak effect of exhaustivity in the former experiments.

A third possibility we entertained is that participants may feel that the problems are too abstract. Our items involve judgments about a character's (Mary's) conditional claims which in turn are about decisions by third parties (e.g., John's teacher). We sought to simplify this dynamic by turning the participant into *both* the decision maker and the interpreter of the conditional sentence.

To examine these factors, we revised the vignettes. First, instead of asking whether the conditional and the further assumption implied a conclusion, we asked whether the speaker intended to convey by her conditional a second conditional: either *If A, B* (the identical statement), *If not B, not A* (the contrapositive), *If B, A* (the converse), or *If not A, not B* (the obverse). These correspond to Modus Ponens, Modus Tollens, Affirming the Consequent, and Denying the Antecedent when we conditionalize the conclusion of the latter arguments on their minor premise. So we can usefully compare the acceptance rates for the new problems to their counterparts in the earlier experiments. Additionally, we changed the setup so that the participant would also be the decision-maker. For example, the converse inference for the test-taking vignette appeared like this (in its [CONS?] version):

John is a high school student, and he is taking *Introduction to Calculus*. John pays attention in class, and he studies with his friend Sarah for all of the exams. John has taken a test on Chapters 4-6 that has not been graded yet.

As a high school advisor, you need to give advice to John on whether or not he should take Calculus II based on his current performance. **You want to decide how well John has performed in the class so far.**

You ask Mary, "What are all the ways John could manage to do well on the test?"

In fact, Mary knows all the ways and is willing to relate them. Mary responds, "If John understood Chapter 5, then John did well on the test."

Does Mary intend to convey, among other things, that if John did well on the test then John understood Chapter 5?

Notice that the problem asks whether Mary intended to convey the conclusion rather than whether her statement implies it.

However, with these changes, we found that [CONS?] questions produced an increase in the endorsement rates for the converse, which was not fully significant ($F(1,23) = 4.04, p = .056$), and, this time, for the contrapositive ($F(1,23) = 5.80, p = .024$). But there was no increase for the obverse ($F(1,23) < 1$). Figure 8 shows this trend in data from 32 participants. An analysis parallel to that of the earlier experiments found the expected difference due to the type of inference ($F(3,23) = 12.44, p < .001$) and also an effect of question type ($F(1,15) = 5.84, p = .029$), but the interaction between question type and inference was not significant ($F(3,23) < 1$).

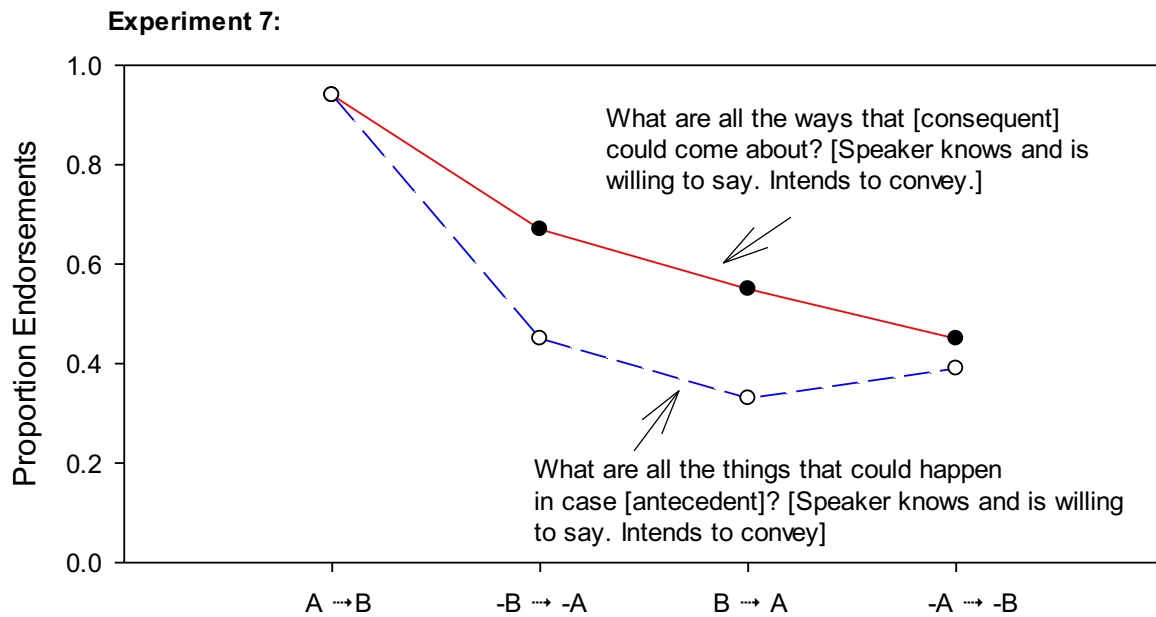


Figure 8. The effect of [CONS?] versus [ANT?] questions on endorsement of a conditional, its contrapositive, its converse, and its obverse, Experiment 7.

Another way to assess these findings is to compare them to the results of Experiment 4, which showed the full perfection pattern. To do this, we considered just the data from the [CONS?] questions from these two studies, treating experiment and inference type as the factors of interest. As in the earlier analyses, inference type had four levels, aligned across experiments according to the correspondence that we mentioned earlier (e.g., Affirming the Consequent in Experiment 4 was paired with the inference to the converse in Experiment 7).¹⁵ This analysis found a significant effect of inference type ($F(3,24) = 4.10, p = .018$) and a significant interaction between experiment and inference type ($F(3,179) = 7.42, p < .001$). These effects can be seen by comparing the red solid lines in Figures 5 and 8. The main effect of experiment was not

¹⁵ We again used a generalized linear mixed model for binomial data, including as random effects: (a) the main effects of participants (nested within experiments) and vignettes (crossed with experiments), and (b) the interactions of participants with inference type, vignettes with inference type, and vignettes with experiments.

Experimenting with (Conditional) Perfection / 23

significant, $F(1, 8) = 2.59, p = .146$. Planned comparisons showed that participants endorsed affirming the consequent in Experiment 4 more often than they endorsed the converse inference in Experiment 7 ($F(1,179) = 8.46, p = .004$) and denying the antecedent more often than the obverse inference ($F(1,179) = 13.68, p < .001$). If an inference from one conditional to another (e.g., from a conditional to its converse) is simpler than a conditional syllogism (e.g., Affirming the Consequent), then these differences are not due to simplicity.

Despite our efforts to clarify the structure of the problems, Experiment 7 failed to get participants to perfect conditionals to the same extent as Experiment 4. Moreover, it produced a near-significant difference in the acceptance of the converse but not the obverse, which is not predicted by any extant model. A clear statement that the antecedent represents the only way of bringing about the consequent produces the biconditional reading. But merely stating a conditional in response to a demand for an exhaustive answer does not fully do so. One could maintain that the positive results for Experiments 4 and 6 and the null results for the remaining experiments simply show that the former experiments provide the right ingredients for obtaining the exhaustivity effect. But although Experiments 4 and 6 do provide evidence for a biconditional reading, they don't provide evidence that the reading was due to the pragmatics of exhaustivity. To obtain the biconditional effect in these two experiments (i.e., increased endorsement of DA and AC), the vignettes added semantic material ("This is the ONLY way [the consequent could come about]," "Here are ALL the ways [the consequent could come about]") that signaled the necessity of the antecedent. What's missing is support for the idea that people will perfect a conditional when it is simply a response to a question about how the consequent could happen.

It is possible, of course, that further tinkering with the vignettes or with the instructions could succeed where Experiments 1, 2, 3, 5, and 7 did not. At this point, though, we think it worthwhile to consider some possible reasons for their lack of perfection.

5. Theoretical Discussion

The positive results in Experiments 4 and 6 suggest that true conditional perfection can, in fact, be linked to exhaustivity. However, the negative results in Experiments 1-3 and (perhaps to a lesser extent) Experiments 5 and 7 require us to put this finding in perspective.

Specifically, a common assumption is that a strong relationship exists between questions and exhaustivity. This relationship includes at least the idea that some overt questions trigger exhaustive readings in conditionals. The experimental evidence we have considered suggests that this is not quite right. In order to trigger exhaustive readings of conditionals, we need more than just questions.

In closing, we briefly consider how our experiments qualify the relation among questions, exhaustivity, and perfection. But first we consider a similar experiment that raises some of the same issues.

5.1. Relation to a prior experiment

There is some similarity between the question-variations we used in our experiments and an earlier experiment by Farr (2011). Farr gave her participants vignettes like this:

Monika sells seafood on the market. She gets 1 euro for a crab, 2.50 euros for an eel, 15 euros for a lobster and 2.50 euros for a pike. Kerstin, an employee of Monika, cannot remember the prices. Since she does not want to ask Monika again, she asks Sahra, who also works for Monika. Sahra knows the prices exactly.

At this point in Farr's experiment, there is a dialogue between Kerstin and Sahra. This dialogue starts with one of two questions:

what-if-p Kerstin: What happens if I sell an eel?

when-q Kerstin: When do I get 2.50 euros?

Sahra: If you sell an eel, you get 2.50 euros.

Participants are then asked:

Did Sahra answer Kerstin's question sufficiently? [Yes] [No]

Farr (2011) found that when the *what-if-p?* question preceded the conditional, participants more often responded that Sahra's answer was sufficient than when the *when-q?* question preceded it. Farr's interpretation is that *when-q?* demands an exhaustive answer with respect to the conditional's consequent (What are all the cases in which I get 2.50 euros?) and triggers a perfected reading of the conditional (if and only if you sell an eel do you get 2.5 euros). Because the conditional mentions only one of the two ways to get 2.50 euros as given in the background story, participants see it as insufficient. The *what-if-p?* question also demands an exhaustive answer, but to a different question (What are all the things that happen if I sell an eel?) and does not perfect the conditional. So participants see Sahra's answer to this question as sufficient.

Despite the similarities in the setup to the present studies, Farr's experiment did not end up testing whether Sahra's conditional gets perfected. As noted, participants learn from the story that there are two ways to get 2.50 euros but only a single price for an eel. So even a simple (nonconditional) answer like "An eel costs 2.50 euros" is complete with respect to questions like "How much is an eel?" or "What happens if I sell an eel?" But the same simple answer is incomplete with respect to "What costs 2.50 euros?" or "When do I get 2.50 euros?" It seems possible, then, that participants' dissatisfaction with the answer after the *when-q?* question does not depend on their perfecting the conditional but instead on their sense of Sahra's lack of full disclosure. For this reason, we think it important to have a more direct indicator of perfection, such as participants' willingness to accept the conditional's converse or to accept an inference like Affirming the Consequent.

5.2. How the Experiments Constrain Explanations of Perfection

At first sight, our experiments appear to refute the idea that questions trigger perfection. Even explicit questions that demand exhaustive answers about the consequent (e.g., *What are all the ways [the consequent] could come about?*) don't always yield a perfected conditional (*antecedent iff consequent*).

According to the von Fintel/Cornulier account that we have been pursuing, questions about some event B set up the expectation that a conditional answer of the form *If A , B* implies that A is the only way B could come about. Thus, *if $\sim A$, $\sim B$* . Together, *If A , B* and *If $\sim A$, $\sim B$* , yield the perfected interpretation, *A iff B* . For example, questions like *Did John do well on the test?* or *What are all the ways John could do well on the test?* imply that the answer *If John understood Chapter 5, then he did well on the test* supplies all ways he could do well. So John did well if and only if John understood Chapter 5. Our experiments, however, failed to produce this pattern of reasoning. In principle, then, this failure could come about either (a) because participants failed to infer a perfected conditional from (what they perceived as) an exhaustive answer, or (b) because they failed to interpret the conditional as exhaustive, in the first place. Let's consider these two possibilities in turn.

5.2.1. Do people infer perfected conditionals from exhaustive answers?

Perhaps our negative results are partial evidence that people do not reason from (i) to (ii):

- (i) For each alternative D to the antecedent A , \sim (*If D , B*)
- (ii) *If $\sim A$, $\sim B$*

In the case of our experiments, failure to infer (ii) from (i) amounts to the idea that participants understood the antecedent (e.g., *John understood Chapter 5*) as the only way that the consequent (*John did well on the test*) could occur, and yet did not infer that if John did not understand Chapter 5, he did not do well.

But on the contrary, Experiments 4 and 6 show that given discourse that basically entails claims of the form of (i), people will reason their way to a claim of the form of (ii).¹⁶ This suggests that the failure to find the perfection pattern in the remaining experiments is due to participants failing to interpret the conditional (e.g., *If John understood Chapter 5, he did well on the test*) as an exhaustive answer to the question (*What are all the ways John could do well on the test?*).

¹⁶ Note, incidentally, that if one rejects the Competitor assumption we sketched in Section 2, this would be indirect, and admittedly very defeasible, experimental evidence that people reason with something like Conditional Excluded Middle.

Experimenting with (Conditional) Perfection / 26

To back up this possibility, we asked participants in a further experiment to decide whether the conditionals mentioned all the ways the consequent could come about.¹⁷ The experiment was very similar to Experiment 5 (in which Mary is said to know all the ways and is willing to relate them), but in addition to asking whether participants agreed with the inference, we also asked them, “Did Mary’s response mention all the ways?” (Half the participants answered the inference question first and half answered the “all the ways” question first, though the order had no statistically reliable effect on the results.) For [CONS?] questions (e.g., *What are all the ways John could do well on the test?*), participants believed that the conditional response (*If John understood Chapter 5, he did well on the test*) mentioned all the ways on only 14.5% of trials.

Of course, our results do not mean that questions *never* produce the presumption that a conditional answer is exhaustive. Rather, the results suggest that even very explicit questions of the proper sort don’t always trigger an exhaustive reading. Something more is needed to ensure it.

5.2.2. Why don’t people believe conditional answers to questions are exhaustive?

Questions don’t always yield exhaustive answers. Still, we might to be able to recover the spirit of the Cornulier/von Fintel proposal on conditional perfection by supposing that questions yield exhaustive answers by default and giving an independent explanation for why this default pattern did not emerge in Experiments 1-3, 5, and 7.

Perhaps the explanation is that some property of the vignettes in these experiments encouraged a *mention-some* reading. Consider this dialogue:

(14) Q: Who are all the people who came to the party?

A: John came.

Even though the question is phrased as demanding a complete list, a possible *mention-some* interpretation of the answer seems available. The respondent rejects the burden of providing a complete answer to the question and volunteers instead whatever information she is able to provide (perhaps expecting that other conversational participants will be able to fill out the rest of the party-goers list). In the specific case of (14), this sort of interpretation might even be invited by the background knowledge that only extremely unusual parties have only one attendee. Similarly, it is possible that, when we ask, “What are all the ways in which [consequent] could come about?” the respondent’s answer is given a *mention-some* reading.¹⁸

¹⁷ We thank Robert Stalnaker for suggesting this experiment.

¹⁸ Demands for exhaustive answers often call for memory searches that exceed people’s abilities, especially in the context of an on-going conversation. In such situations, speakers may be thrown back on a satisfactory answer that is informative, but partial. Hearers may likewise make allowances for this kind of satisficing by leaving open the possibility that the speaker’s answer is all she can come up with at the moment—that it is temporarily exhaustive, rather than exhaustive period. In the case of (14), for example, this interpretation is enhanced if the speaker indicates

Though we do not think that this assessment is without merit, it has too many surprising consequences to be plausible. What we found is that, unless the respondent explicitly avows providing a complete answer, there is no significant pattern of perfection. Though there are slight increases in endorsement rates for the fallacies Affirming the Consequent and Denying the Antecedent in some experiments, they are typically not enough to meet standard significance thresholds. One might have expected that in the cases where a *mention-all* interpretation is possible but not mandated, we should have seen some participants reach for it. In Experiments 5 and 7, we even tried as much as possible to “load” up the context so as to invite a *mention-all* answer, but without much success. This is especially striking because *mention-all* interpretations of answers to questions are not unusual. According to the received wisdom (see, for example, van Rooij & Schulz, 2003), they should be preferred unless they are contrary to expectations (as in (14)). More generally, much recent work on implicatures argues that the computation of implicatures happens by default (see e.g., Chierchia, 2013).

Of course, it is still possible that participants read the [CONS?] questions as having open-ended answers that respondents were unlikely to answer exhaustively. In the case of our running example, participants may have taken the question, *What are all the ways John could manage to do well on the test?* as placing an impossible demand on the addressee, given the many ways John could do well (e.g., cheating, bribing the instructor, divine intervention, lucky guessing,...). However, some of the vignettes in our experiments were explicit in listing alternatives for the antecedent. For example, one of the vignettes began with the sentence *Someone has put a fertilizer, either Easy-gro or Bloom-builder, on the plants.* This was then followed by the [CONS?] question *What are all the ways the plants could manage to grow quickly?* and the conditional *If Easy-gro was put on the plants, then the plants grew quickly.* Although items like these are not completely immune to *mention-some* interpretations, they seem at least less open to these interpretations than those in which the alternatives are unspecified. The initial sentence seems to limit the relevant alternatives to just a few (in this case, Easy-gro or Bloom-builder); so the speaker of the conditional should find it less of a burden to provide an exhaustive list. However, a re-analysis of Experiment 1 shows that participants were not more likely to endorse the inferences for the items with explicitly-provided alternatives than for the remaining open items (Vignettes 1, 4, 12, and 14 in the Online Appendix have unspecified alternatives, and the remainder specified ones). Considering just the items with [CONS?] questions, we find that participants endorsed Affirming the Consequent on 54% of trials for the items with alternatives and 62% for the open items. Similarly, they endorsed Denying the Antecedent on 48% of trials for items with alternatives and 50% for the open items. These differences are in the wrong direction to explain the results. We should be cautious here about drawing strong conclusions, since we did not design the experiment with this difference in mind. Still, it provides some presumptive evidence against an explanation based on response burden.

some hesitancy: “Well...John came.” Viewed in this way, what our results suggest is that this temporarily exhaustive reading is more available than one might have expected.

It appears then that pragmatic reasoning based on background questions is not enough to trigger the relevant exhaustive readings.¹⁹ What turns out to be necessary—in our experiments, at least—is linguistic material that explicitly directs hearers towards an exhaustive interpretation.

A very austere development of this idea would be to claim that the biconditional interpretation requires that such material always be explicitly represented. This amounts to the claim that there is no distinctive pragmatic phenomenon of conditional perfection. Biconditional interpretations arise as entailments when an utterance of *If A, B* is conjoined with additional claims that are strong enough to entail their converses.

The immediate problem with this explanation is that several previous experiments have found evidence of perfection stemming from the conditional's status as a promise or threat (e.g., *If you disrupt the class, you'll have to leave the room*). (See the sources cited in Section 3.) These conditionals don't entail their obverses—they merely suggest them—so perfection is obtainable without explicit entailment of *If ~A, ~B*. What is the case is that questions, in particular, are not always enough to yield the obverse.²⁰

¹⁹ It is important to acknowledge in this context that there are theories like Franke (2009, section 5.2.2) which maintain that there are two, quite distinct kinds of perfection inferences and that alternative-based strengthening is only one of them. According to Franke some perfection inferences arise on account of expectations of normality in context. We think this is quite possible, and if so, there may be ways to construct vignettes (other than those we used here) that would produce perfection. However, we still require an explanation for why alternative-based strengthening failed to yield this pattern in our studies.

²⁰ Herburger's (2015) "Whole Truth" account of conditional perfection provides one possible route to deriving perfection without the need for explicit exhaustification devices. Her preliminary statement of the theory is:

Conditional Perfection and upper-bounding inferences arise as logical entailments when a sentence *S* is silently conjoined with *only S*, resulting in the conjunction *S and only S*. *S and only S* is then taken to express 'the truth and the whole truth'. (Herburger, 2015, p. 6)

Note that Herburger strikes through '*and only S*' to indicate the fact that it may not be pronounced. If it is possible to interpret utterances of *If A, B* as utterances of *If A, B and only If A, B*, then that's when we should expect perfection inferences to arise. According to this intermediate take, perfection inferences can arise due to overt or covert exhaustification.

This sort of approach is difficult to evaluate, experimentally at least, absent some systematic ideas about when we are allowed to supplement *S* with the silent *and only S*. Since it is part of the proposal that there are no systematic principles that connect the questions made salient by the discourse to the availability of the strengthened interpretation, it is hard to see what such principles might look like.

But perhaps a non-experimental argument is available. If perfection is achievable with overt exhaustification devices (as our experiments demonstrate), and if there is precedent for covert exhaustification, we should expect that nothing prevents an exhaustified interpretation of conditionals. The remaining question, once again, is why we did not find much trace of these exhaustified interpretations in our first experiments.

On the positive side, then, we strongly suspect that there's more than one path to perfection. One way to obtain it is through threats, promises, permissions, and obligations, which convey via practical reasoning that if the addressee does not perform the key action the reward, punishment, or some other kind of normative status will not be forthcoming. A second path to perfection is through background information that the converse of the conditional is true. Section 3 reviewed some evidence for these possibilities. A third route is through explicit exhaustification devices like the ones we used in Experiments 4 and 6. Pluralism about perfection may be due to the absence in natural language of expressions as simple as *if* that mean *iff*, leaving *if* to cover for *iff* in any of the variety of situations (outside math classes) in which *iff* would be more exact. Provided that there is some contextual reason to think that *If B, A* is true, then an utterance of *If A, B* may suggest *A iff B*. Because there are many reasons why *If B, A* could be true, there are many ways to achieve perfection.

Here's where we've got to: Questions that explicitly request exhaustive answers don't seem sufficient to produce perfection. That's the experimental finding. This seems to be because the answers aren't read as exhaustive rather than because an exhaustive answer doesn't yield perfection. Although it's possible that the question-to-exhaustive-answer link is the normal case and something about our experimental materials militated against it, the usual suspects—the respondent's lack of knowledge, uncooperativeness, and response burden—seem to be ruled out by the results. There may be some unusual suspects responsible for blocking the normal route to interpreting the answer as exhaustive, but what could these be? Instead, it seems more likely that an exhaustive interpretation of an answer requires more than just a question demanding one. This something more could be an explicit avowal that the response is exhaustive, but it seems likely that people could settle on exhaustive readings from weaker evidence. Perhaps what listeners require is some reason to think that an exhaustive conditional answer is in the respondent's interest. Otherwise, the listeners' experience with their own communicative foibles may make them hesitant to think they've gotten the full story.

Bibliography

- Adams, E. (1965). The logic of conditionals. *Inquiry*, 8, 166-197.
- van der Auwera, J. (1997). Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics*, 27, 261-274.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effect structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.
- Bonnefon, J.-F., & Politzer, G. (2010). Pragmatic conditionals, conditional pragmatics, and the pragmatic component of conditional reasoning. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 233-249). Oxford, UK: Oxford University Press.
- Cariani, F. (2021). *The modal future*. Cambridge, UK: Cambridge University Press.
- Cariani, F., & Rips, L. J. (2017). Conditionals, context, and the suppression effect. *Cognitive Science*, 41, 540-589.
- Cariani, F., & Santorio, P. (2018). Will done better: Selection semantics, future credence, and indeterminacy. *Mind*, 127, 129-165.
- Chierchia, G. (2013). *Logic in grammar*. Oxford, UK: Oxford University Press.
- Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, 19, 274-282.
- de Cornulier, B. (1983). If and the presumption of exhaustivity. *Journal of Pragmatics*, 7, 247-249.
- Evans, J. St.B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning*. Hillsdale, NJ: Erlbaum.
- Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, UK: Oxford University Press.
- Farr, M.C. (2011). Focus influences the presence of conditional perfection: Experimental evidence. In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn und Bedeutung*, 15, 225-239.

- Fillenbaum, S. (1975). If: Some uses. *Psychological Research*, 37, 245-260.
- von Fintel, K. (1997). Bare plurals, bare conditionals, and *only*. *Journal of Semantics*, 14, 1–56.
- von Fintel, K. (2001). *Conditional strengthening: A study in implicature*. MIT. <http://mit.edu/fintel/fintel-2001-condstrength.pdf>.
- Franke, M. (2009). *Signal to act: Game theory in pragmatics*. [Doctoral dissertation, Amsterdam University]. <https://www.illc.uva.nl/Research/Publications/Dissertations/DS-2009-11.text.pdf>
- Geis, M., & Zwicky A. (1971). On invited inferences. *Linguistic Inquiry*, 2, 561–566.
- Groenendijk, J., & Stokhof, M. (1984). *Studies in the semantics of question and the pragmatics of answers*. CiteSeerX. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.867.7509&rep=rep1&type=pdf>.
- Herburger, H. (2015). Conditional perfection: the truth and the whole truth. In S. D'Antonio, M. Moroney, & C. R. Little (Eds.), *Proceedings of SALT 25* (pp. 615-635).
- Horn, L. (2000). From IF to IFF: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics*, 32, 289–326.
- Kratzer, A. (1991). Conditionals. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 651-656). Berlin: de Gruyter.
- Kratzer, A. (2012). *Modals and conditionals*. Oxford, UK: Oxford University Press.
- Lilje, G. W. (1972). Uninvited inference. *Linguistic Inquiry*, 3, 540–542.
- Marcus, S. L., & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior*, 18, 199-223.
- Markovits, H., & Lesage, C. (1990). Pragmatic reasoning schemas for conditional promises: Context and representation. In J.-P. Caverni, J.-M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 183-192). Amsterdam: North-Holland.
- Newstead, S. E., Ellis, M. C., Evans, J. St.B. T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning*, 3, 49-76.

Experimenting with (Conditional) Perfection / 32

Oaksford, M, & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, UK: Oxford University Press.

van Rooij, R., & Schulz, K. (2003). *Exhaustification*. CiteSeerX.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.13.1744&rep=rep1&type=pdf>

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in Logical Theory* (pp. 98-112). Oxford, UK: Blackwell.

Stalnaker, R. (1981). A defense of conditional excluded middle. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs: Conditionals, belief, decision, chance, and time* (pp. 87-104). Dordrecht: Reidel.

Staudenmayer, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 55-79). Hillsdale, NJ: Erlbaum.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22, 742-758.

Van Canegem-Ardijns, I., & Van Belle, W. (2008). Conditionals and types of conditional perfection. *Journal of Pragmatics*, 40, 349-376.