If not for counterfactuals: negating counterfactuality in natural language

Robert Henderson
University of Arizona

Draft as of Jan 2014

Abstract

Based on previously unnoticed contrasts between standard counterfactuals and the non-canonical counterfactual construction *"if not for ϕ, ψ"* (hereafter NC, for "Not" Counterfactuals), this paper (i) presents new evidence for a distinction between epistemic and ontic readings of counterfactuals, (ii) develops a new account of this distinction, and (iii) argues that doing so requires enriching our models with causal laws (see e.g., Schulz, 2007). First, we show that NCs systematically reject inferences against the flow of causality, providing a point of contrast with standard counterfactuals, and supporting the existence of a separate epistemic reading for counterfactuals. We then show that NCs reject these inferences because, unlike standard counterfactuals, they presuppose the counterfactuality of their antecedents. After enriching the model with causal laws, we present an analysis in which NCs presuppose a fact, which fixes those facts lying causally upstream. This accounts for the new observation that NCs systematically reject non-causal epistemic inferences, while otherwise retaining their paraphrasability with standard counterfactuals.

If not for the courage of the fearless crew, the Minnow would be lost.
- Gilligan's Island

1 INTRODUCTION

The central obstacle to building a semantics for counterfactuals is to determine what it means to make a counterfactual assumption. Once the contribution of

the antecedent is calculated, truth hangs on the simpler question of whether the consequent is entailed. For example, consider Ramsey's (1950) famous intuition about the evaluation of counterfactuals, which spawned a series of approaches under the heading of Premise Semantics (Rescher, 1964; Veltman, 1976; Kratzer, 1981b, a.o.). Stalnaker summarizes it nicely (1968, p. 106):

> "first add the antecedent (hypothetically) to your stock of beliefs; second make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is true."

To see that all of the action is in the evaluation of the antecedent, just consider how trivial the final clause is compared to the challenge of maintaining a consistent set of beliefs in the face of accommodating a proposition that is by hypothesis counter-to-fact. In such a situation, one must be willing to give up some of what is known to be the case, but not too much in either quantity or quality. This is why Lewis (1973, p.73) famously characterized the relevant counterfactual worlds as those "where a small, localized, inconspicuous miracle" permits the admission of the counterfactual antecedent. This paper investigates how these miracles are made. Specifically, we consider two different species of counterfactual antecedents to see how the properties unique to each constrain the revision process, and thus the inferences that can be made with that type of counterfactual.

At first pass, it seems like Ramsey's intuition and Lewis's intuition about counterfactual revision are opposite sides of the same coin; the first is couched in terms of belief revision, while the second characterizes the relevant antecedent worlds as those where facts, not an agent's belief state, have changed. We might think that evaluating the consequent against minimally different belief states supporting the antecedent and minimally different worlds supporting the antecedent would yield equivalent results, but this is not the case, as shown by Katsuno & Mendelzon (1991) and Schulz (2007). They show that only in the latter case can the revision induced by the antecedent lead to derived inconsistences with laws the speaker takes to hold. That is, the local revision of worlds can allow facts to count more for similarly than laws, while global belief revision does not. While this is an interesting formal result, Schulz (2007) goes one step further and argues that natural language makes use of this distinction. Specifically, she argues that these two routes to counterfactual revision map onto two distinct readings for counterfactuals, *epistemic* and *ontic*.[1] Though we depart from Schulz's (2007) analysis,

[1]Though the terminology is different, this is similar to the epistemic-metaphysical distinction

the epistemic-ontic distinction lies at the core of this work. In particular, we argue that ontic readings are required when a counterfactual construction presupposes a fact that the antecedent must remove. Epistemic readings can arise when no such fact is presupposed.

Intuitively, epistemic readings concern what an agent could infer when in a different epistemic state based on the proposition given by the antecedent. On the other hand, ontic readings, which we will see correspond to the default interpretation of counterfactuals, are about the consequences if the facts were different so as to come into line (perhaps miraculously) with the antecedent. We can illustrate the difference with a classic example. Consider the context in (1), provided by Lewis (1973). In such a context, evaluating the antecedent leads to ambiguity in whether the past must change. Example (1a) shows that the counterfactual antecedent can return worlds where the past is different, while example (1b) presents a case where the past must not have changed. since the surprise would be due to the fact that the fight happened.

(1) Jim and Jack had a big quarrel yesterday, and Jim is still very angry.

   a. If Jim were to have asked Jack for help today, there would have to have been no quarrel yesterday.

   b. If Jim were to have asked Jack for help today, I would have been surprised.

Backtracking counterfactuals like (1a) classify as epistemic because they present what one could conclude given a new piece of information. The inference proceeds as follows in the case of (1a): Given that one does not ask for help from those they are angry at, one could infer that Jack and Jim had had no recent fight upon learning that Jim asked Jack for help. Crucially, it does not seem to mean that a consequence of Jim asking Jack for help is that the past must change so that there is no fight. That is to say, Jim asking Jack for help is a sign of whether there had been a fight, which an agent could use to make an inference, but it does not directly cause or prevent fights. In contrast, example (1b) can be used to characterize the effects of a miraculous change, which in this case lead me to be surprised. The counterfactual revision leaves the past unaffected, but leads to counterfactual futures that take into account the effects of the changes made to accommodate the antecedent.

Condoravdi (2002) draws between two types of readings for necessity and possibility modals. Kaufmann (2005) draws a similar distinction between predictive and non-predictive readings of indicative conditionals.

The problem presented by examples like (1) is how to account for the fact that the same antecedent in the same context can generate two different counterfactual updates. Previous authors have argued that counterfactual antecedents make use of two different revision mechanisms (Schulz, 2007), or have access to different ordering sources/premise sets due to the inherent vagueness of counterfactual revision (Kratzer, 1979, 1981a,b, 1989). Using new data from a counterfactual construction in English that has not yet been treated in the literature, this paper presents a different account of epistemic counterfactuals rooted in the old observation that standard counterfactuals only implicate the counterfactuality of their antecedents (Anderson, 1957; Stalnaker, 1975). Specifically, we investigate the counterfactual construction in (2) (henceforth NC for "not counterfactual"), compared to standard counterfactuals, exemplified in (3).

(2) If not for Mary going to the store, we wouldn't have salsa.

(3) If Mary hadn't gone to the store, we wouldn't have salsa.

While (2) and (3) appear to be good paraphrases of each other, NCs systematically differ from standard counterfactuals. First, unlike standard counterfactuals, which only implicate the counterfactuality of their antecedents, NC antecedents are non-defeasibly counter-to-fact. Secondly, NCs differ from standard counterfactuals in that they do not support certain non-causal inferences, including those based on backtracking and correlations, which only make sense under an epistemic reading.

Our analysis ties these two differences together. The core idea is that counterfactual antecedents make minimal revisions to worlds in models that incorporate causal relations (Halpern & Pearl, 2005a,b; Pearl, 2000; Schulz, 2007, i.a.).[2] The difference between standard counterfactuals and NCs is that only the latter presuppose a fact about the world their antecedents alter. We then provide an analysis where this fact fixes those facts that lie upstream with respect to the flow of causality. The result is that NC antecedents require a Lewis-style revision that suppresses epistemic inferences, while standard counterfactuals are correctly predicted to allow both epistemic and ontic readings, with a bias for ontic readings proportional to the strength of the counterfactuality implicature of the antecedent.

[2]Incorporating causality is important because, as the previous example shows, the differences between ontic and epistemic readings of counterfactuals come out most clearly after taking into account the temporal and causal relationships between antecedent and consequent. The reason is that when the counterfactual antecedent causally antecedes the consequent, the ontic and epistemic ambiguity makes no difference in the inferences available. This is not surprising because, as shown later, the effects of a change are precisely what an agent will be able to infer given the information that such a change happened.

Moreover, we tie the semantic differences between NCs and standard counterfactuals to their clear morphological differences. NC antecedents do not contain a proposition implicated to be counter-to-fact, but a presupposed eventive nominal. In this way NC antecedents presuppose a fact about the world that the antecedent can revise. Obligatory negation in NCs is interpreted as a model update function that removes facts, suppressing the correct class of non-causal epistemic inferences. Crucially, when the inference at issue is ontic and the antecedent is truly counter-to-fact, NCs will not differ from standard counterfactuals, which accounts for those cases where there is mutual paraphrasability, as in (2-3).

Beyond providing the first analysis of this non-canonical counterfactual construction in English, this work makes a number of contributions to our understanding of counterfactuals. First, the fact that NCs cannot be used to make epistemic inferences provides important evidence that the epistemic uses of standard counterfactuals are a distinct reading, which has been challenged (Rott, 1999; Veltman, 2005, a.o.). That is, to establish that a morpheme or construction has two readings, we need a second morpheme or construction that has one, but not the other. NCs provide just this point of contrast. A similar point can be made for the observation that standard counterfactuals implicate, but do not entail that their antecedents are actually counter-to-fact. While not as controversial as the existence of epistemic readings, the fact that NCs contrast strongly with standard counterfactuals in the tests for this implicature strengthens our confidence in this old result. Finally, in accounting for the distribution of NCs, we end up developing a new test for our intuitions about the flow of causality in counterfactual contexts. We will see that if we can reason between A and B with an NC, but not in the opposite direction, then A must causally antecede B. We will show how we can use this test to generate insights about classic counterfactual puzzles.

The analysis begins in §2 which details the morphosyntactic and semantic properties of NCs, especially where they diverge from standard counterfactuals. Section §3 presents the formal system, which enriches our models with causal relations. It also develops an analysis of the contribution of NC antecedents within such causal models and compares our approach to other accounts of the epistemic-ontic distinction. Section §4 concludes, considering areas for future work, both in other constructions and other languages.

2 INTRODUCING A NEW COUNTERFACTUAL

The goal of this section is to present the core morphological and semantic properties of NCs. The discussion is detailed because NCs have not yet been thoroughly described in the literature. Moreover, section §3 will provide an account of the semantic generalizations presented here using only the morphological resources considered now.

2.1 The Core Morphosyntactic Generalizations

The counterfactual construction that is the empirical focus of this work consists of a nominal embedded under *if not for* (4a-4b). As a consequence of accepting nominals, the construction also embeds clausal gerunds, and naturally occurring complex examples are easily found (5a-5b).

(4) a. If not for the Beatles, we wouldn't have CT scans.[3]

b. The heir of Arvedui, who would have become King of Arthedain if not for the destruction of his realm by the Witch-king . . ..[4]

(5) a. If not for finding your web site out there in cyberspace when I was searching on how to eliminate my bad breath, I would not have had my bad breath problem resolved.[5]

b. If not for this couple trying social nudism, . . . they would have almost no one to share their wedding day with.[6]

Although NCs allow both individual denoting nominals along with more complex eventive nominals, the analysis that follows will focus on NCs with clausal gerunds because doing so makes it easier to draw comparisons to standard counterfactuals which have clausal antecedents. That being said, restricting the analysis to NCs embedding clausal gerunds is not an unreasonable idealization. When using an individual-denoting nominal in an NC, one must always construct some relation that individual is in, yielding something approaching a clausal gerund meaning.

(6) If not for Mary, I would have failed the exam.

[3] http://thedecisiontree.com/blog/?p=265

[4] www.glyphweb.com/arda/c/chieftainofthedunedain.html

[5] http://www.breathcure.com/testimonials_b.html

[6] http://thenudelife.com/category/bride/

Example (6) requires the listener to construct a relation from the context, like *tutoring me*, *changing my grade in the gradebook*, etc.[7] While there is much more to be said about how to construct relations in context, which will not be done here, the important point is that NCs *syntactically* allow individual-denoting nominals in the antecedent, but *semantically* require an individual along with a relation, like that given by a clausal gerund.[8]

The second morphosyntactic property unique to NCs is the preposition *for*. Note that *for* has a non-trivial contribution, distinguishing NCs from minimally different conditionals with different interpretations, as in (7-8)

(7) a. If not the Corvette, what would I buy?
b. #If not for the Corvette, what would I buy?

(8) a. If not for Mary, how would we survive?
b. *If not Mary, how would we survive?

Notice that a nominal embedded under *not*, like (7a), can restrict the domain of the WH-word in the consequent, in this case, to values other than *the Corvette*. NCs have no such reading (7b). Similarly, NCs, like other conditionals, can be used in conditional questions (8a) (Hulstijn, 1997; Isaacs & Rawlins, 2008; Velissaratou, 2000), while the same antecedent without *for* is ungrammatical (8b). This can be attributed to the fact that *Mary* is not in the domain of manner questions. We see that *for* helps define NCs morphosyntactically, and the preposition will be given a non-trivial semantics in §3.

Another morphosyntactic fact about NCs that needs to be considered is that they come in two forms, one of which is less *elliptical* than the type discussed so far.[9]

[7]While we should be able to access the existence relation in principle, such readings are, in fact, difficult to construct. We can see that this is true from the following sentence, which does not easily generate the reading that I have three siblings and if my brother did not exist, I would only have two. Much stronger is the reading where my brother did something to ensure the survival of one of my siblings, including himself.

(1) If not for my brother, I would only have two siblings.

[8]See, for instance, the large literature on nominal complements of predicates like *begin*, which have been argued to require the accommodation of an event/relation (de Almeida, 2004; McElree et al., 2001; Pickering et al., 2005; Traxler et al., 2002, i.a.).

[9]Since no known ellipsis process can derive one from the other, the two constructions will not

(9) a. If not for Mary going to the store, we wouldn't have salsa.
b. If it weren't for Mary going to the store, we wouldn't have salsa.

There are no detectable semantic difference between (9a) and (9b) with regard to the properties discussed in this work, so the following analysis will treat them interchangeably. That being said, the distinction does provide an important lesson. Where the antecedents of conditionals can stand alone as a sentence, as in (10), this is not the case with NCs, shown in (11).

(10) a. If Mary hadn't gone to the store, we wouldn't have salsa.
b. Mary hadn't gone to the store.

(11) a. If not for Mary going to the store, we wouldn't have salsa.
b. *Not for Mary going to the store.

(12) a. If it weren't for Mary going to the store, we wouldn't have salsa.
b. *It wasn't for Mary going to the store.

The same facts hold for the less elliptical version in (12). Even though the material embedded under the conditional complementizer in (12a) has all of the trappings of a sentence radical, the antecedent cannot stand alone as such. It appears that the copula is only present in (12) to support negation, which forces expletive *it* to appear because finite English verbs require subjects. The ungrammaticality of (12b) is due to the fact that there is no such root clause expletive construction. This is clear when we consider the grammaticality of the surface similar sentence in (13-14) with referential *it* and a benefactive preposition.

(13) It wasn't for Mary to cook with.

(14) It wasn't for Mary's brother to buy.

Although similar to NCs, notice that (13-14) necessarily contain a gap coreferential with the pronoun that does not appear in NC antecedents.

---

be treated as syntactically related. One possibility is that English is in the process of grammaticalizing (9b), eliminating the morphology that has no critical semantic contribution, namely *be* and expletive *it*. English just happens to be a stage where both forms are available. The end result would be a distinct conditional complementizer made from the fusion of *if*, negation, and *for*. This might be a common path of grammaticalization. Ippolito & Su (to appear); Nevins (2002) report on Mandarin *yaobushi* counterfactuals which share properties with English NCs and are formed with a special complementizer consisting of a fusion of negation, subjunctive, and the conditional complementizer. Nevins (2002) reports similar facts for Tagalog.

What this discussion shows is that NCs, in both their forms, are constructional. There is a fixed piece of morphosyntax that can only occur in the antecedent of a conditional that allows a nominal to provide the semantic content of that antecedent. While the morphosyntax is fixed, it is not opaque. We will see that each piece of morphology makes sense from the general properties of NCs, and so they will be given a compositional treatment. Moreover, since it has been established that the morphology in the expanded NC is expletive, it will be ignored, and the analysis will only work with the morphology that is present in the reduced NC.

The final morphological fact to be considered is especially significant because it is crosslinguistically stable. While the morphology of NCs varies across languages, negation is always present. Example (15), from Kaqchikel (Mayan), shows that there are NCs with nominal antecedents but no preposition, while the example (16) from Spanish has a preposition like English (in this case *porque* 'because'), but allows full clausal complements in place of nominals (which English and Kaqchikel have). In all cases, though, negation is obligatorily present, as is clear from the (b) examples.

(15) KAQCHIKEL

a. *Wi **man** ta Maria, yi-b'ison ta.*
If **NEG** IRR Maria, INFL-sad IRR
If not for Maria, I would be sad.

b. *Wi ta Maria, ...

(16) SPANISH

a. *Si **no** fuera porque María ha ido a la tienda, no tendríamos salsa.*
If **NEG** be.SBJ because Maria had gone to the store, NEG have.COND.PST.1pl salsa
If not for Maria going to the store, we wouldn't have salsa.

b. *Si fuera porque María ha ido a la tienda, ...

The same is true for English NCs as in (17)

(17) a. If not for Mary, I wouldn't have passed the test.

b. *If for Mary, I would have passed the test.

c. If it weren't for Mary, I wouldn't have passed the test.

d. *If it were for Mary, I wouldn't have passed the test.

These data show that negation is a critical part of NCs, both in English, and crosslinguistically (see Ippolito & Su (to appear) and Nevins (2002) for similar data in Mandarin and Tagalog respectively). This fact will have an explanation after considering the semantic properties of NCs in the following section. In particular, the eventive nominals in NC antecedents will be shown to be factive. Since the nominal denotes a fact of the actual world, negation is necessary in order to reason counterfactually about the consequent.

2.2 The Core Semantic Generalizations

The previous section considered how the morphology of NCs differs from that of standard counterfactuals. This section presents the two core semantic generalizations that receive an account in §3. First, NCs, unlike standard counterfactuals, have antecedents that are non-defeasibly counter-to-fact. The second is that while standard counterfactuals are ambiguous between ontic and epistemic readings, NCs systematically lack epistemic readings.

2.2.1 Non-defeasibly counterfactual antecedents

Both NCs and standard counterfactuals generate inferences that the antecedent and consequent do not hold in the world of evaluation $w*$, as shown in (18-19).

(18) If Mary hadn't gone to the store, we wouldn't have salsa.

a. Inference: Mary went to the store in $w*$.

b. Inference: We have salsa in $w*$.

(19) If not for Mary going to the store, we wouldn't have salsa.

a. Inference: Mary went to the store in $w*$.

b. Inference: We have salsa in $w*$.

It is well known that the inferences that accompany standard counterfactuals, like those in (18), are only implicatures. The reason is that counterfactuals can be used to argue for the proposition in the antecedent, as well as to conduct informative modus tollens arguments (Anderson, 1957; Stalnaker, 1975). Both of these arguments fail for NCs. First, note that while the truth of the consequent can be used as evidence for the truth of the antecedent proposition in (20), this is not possible with NCs, as in (21).

(20) If John were not sick with the measles, he would have the same temperature he does now; therefore, I conclude he doesn't have the measles.

(21) #If not for John being sick with the measles, he would have the same temperature he does now; therefore, I conclude he doesn't have the measles.

The consequent in (20) cannot be counter-to-fact because its truth is taken as evidence that the antecedent is also true in the actual world. Since neither the antecedent nor the consequent of (20) are counter-to-fact, it is not possible to tell whether the infelicity of (21) is due to the non-defeasible counterfactuality of the antecedent or the consequent. The next example, though, shows that it is the antecedent that is strongly counterfactual. Modus tollens arguments with counterfactual conditionals involve denying the consequent (which amounts to asserting its counterfactuality), in order to argue for the counterfactuality of the antecedent. With standard counterfactuals, the denial of the consequent provides an argument for the proposition that would make the counterfactual antecedent false in the actual world. Since this proposition is only implicated, it can then be asserted without being uninformative as in (22). In contrast, the same sentence with an NC, shown in (23), does not pass for an argument. This would make sense if the proposition making the antecedent counterfactual is already in the common ground.

(22) If the butler hadn't carved the turkey, the knife would have been sharp. The knife was dull; therefore, the butler carved the turkey.

(23) #If not for the butler having carved the turkey, the knife would have been sharp. The knife was dull; therefore, the butler carved the turkey.

If it is true that NCs are only admissible if the antecedent is counterfactual with respect to a proposition already in the common ground, then it should not be possible to felicitously use an NC when the requisite proposition is missing. Example (24) shows that this is the case. An NC cannot be used in a modus tollens argument if the antecedent is still under discussion.

(24) A: John went to the store.
B: No he didn't.

a. A: (Yes, he did, since:) If John hadn't gone to the store, he would be home right now (and he isn't).

b. A: #(Yes, he did, since:) If not for John having gone to the store, he would be home right now (and he isn't).

The response in (24b) is infelicitous because it ignores the fact that whether or not John went to the store is under discussion. It cannot be in the common ground because of speaker B's denial.

To summarize, the generalization is that NCs are only licit if their antecedents are presupposed to be counter-to-fact. This conclusion is supported by the fact that NCs cannot be used when arguing for the truth of the antecedent, that they generate an informativity violation in modus tollens arguments, and that they are not admissible if the truth of the antecedent is under discussion. In §3 we will give an analysis of NC antecedents in which they presuppose a fact about the world that the antecedent can revise through negation. Furthermore, we will show that it is precisely this presupposition that blocks epistemic inferences with NCs, though first we lay out the empirical facts about availability of various types of inferences with NCs.

2.2.2 Missing epistemic inferences

To see the difference between epistemic and ontic inferences it helps to think about counterfactuals against the background of causal laws, where causal laws are functions from causes to effects, but not vice versa. A counterfactual inference that follows the flow of causality can be seen as reporting one of two things: (i) how the world would unfold with respect to the causal laws if the facts concerning the antecedent were different, and (ii) what one could infer given different information about the antecedent and knowledge of the causal laws. We call the first inference ontic and the second epistemic. Things are not very exciting in this situation, though, because what we could infer based on the causal laws is the same as tracking the causal effects of assuming the antecedent. The two readings only pull apart when considering an inference against the flow of causality. Here a counterfactual cannot felicitously report how the world would unfold with respect to the causal laws if the facts concerning the antecedent were different. The reason is that changing the antecedent facts has no effect on the consequent via the causal laws because the causal laws only run from cause to effect. The result is that to make an inference against the flow of causality requires an epistemic inference using the information in the laws, but not causal consequence itself. Schematically, given that only A causes B, if you know that B happened you can infer that A happened, but that does not mean that B caused A.

Since counterfactual inferences against the flow of causality do not follow under an ontic reading, it makes sense that many classic examples of epistemic readings for counterfactuals make use of contexts where the antecedent does not

causally antecede the consequent. But if counterfactual inferences against the flow of causality only go through under epistemic readings, and epistemic readings are not clearly distinguishable in inferences that follow the flow of causlity, there would be no reason to propose a separate ontic reading. The evidence that we need to is that, in fact, when successful, such an inference from effect to cause stands out, presenting a strong contrast with what seems to be the default case for counterfactual inferences. Suppose, for instance, that the children are full because they ate pizza. It seems fine to say (25), but (26) is infelicitous.

(25) If they hadn't eaten pizza, they wouldn't have been full.

   a. Ontic Paraphrase: If we changed the world so that they did not eat pizza, we could conclude that they would not be full.
   b. Epistemic Paraphrase: If we had learned that they did not eat pizza, we could conclude that they would not be full.

(26) ??If they hadn't been full, they wouldn't have eaten pizza.

   a. Ontic Paraphrase: If we changed the world so that they were not full, we could conclude that they would not have eaten pizza.
   b. Epistemic Paraphrase: If we had learned that they were not full, we could conclude that they had not eaten pizza.

We present (25-26) along with their epistemic and ontic paraphrases. Note that in the case of (25), both the epistemic and the ontic seem acceptable. That is, one of the consequences of changing the world so that we did not eat pizza is that it would unfold in a way leaving them hungry. Similarly, if we had learned, counter to fact, that they did not eat pizza, then it would be a valid conclusion that they would be hungry. In contrast, the ontic paraphrase of the backtracking counterfactual in (26) sounds just as odd as (26) itself. That is, changing the world so that they are not full does not require the world change so that they did not eat pizza recently, that is, it is a not a causal consequence of such a change. In constrast, the paraphrase of the epistemic reading seems sound, especially if we knew that the only thing filling on the menu were pizza. Crucially, (26) seems to completely lack such a reading.

In light of facts like these, we find ourselves in the opposite position as before concerning the existence of two readings for counterfactuals. If we cannot distinguish between the epistemic and ontic paraphrases of (25), and (26) does not have access to the epistemic reading that would make it felicitous, then we have no need for the epistemic reading. What we will come to see, though, is that there are

contexts where backtracking inferences like (26) are licit. The result is that counterfactuals cannot only have epistemic readings because sometimes inferences like backtracking are unavailable, yet counterfactuals cannot only have ontic readings because they are sometimes available. While this suggests that there are two readings for standard counterfactuals, what we will show is that NCs do not allow inferences against the flow of causality, even in extraordinary contexts. In particular, the two non-causal inferences to be considered here involve *backtracking* and *correlations*. Since such inferences only make sense under an epistemic reading, the conclusion is that NCs must lack this reading and only have ontic readings. We begin the discusion with backtracking.

As Arregui (2004) notes, backtracking requires a context supporting either a clear causal or analytic relation between the antecedent and consequent.[10] The reason is that backtracking works when information about the consequences of such a relation can be used to make an epistemic inference about its cause. For example, if we conceive of the marriage ceremony causally affecting an individual's bachelorhood status through the analytic relation supplied by the predicate *bachelor*, the backtracking inference in (27) is licensed.

(27) If John weren't a bachelor, he would have to have had a wedding.

The inference in (27) proceeds as follows: Upon learning counterfactually that John is not a bachelor, one can infer that he would have had a marriage ceremony due to the definition of bachelors as unmarried males.

Example (28) presents another case of backtracking, but this time it is supported by the causal connection between rain and the functionality of the car's wiring.

(28) Suppose that the car works perfectly except for after a rain, which always causes the wiring to short out. Suppose it didn't rain and the car started as usual.

   a. If the car hadn't started, it would have to have rained.

The reason we want to treat inferences like (28) as epistemic is that intervening to change the car's state does not cause it to have rained, but learning that the car

[10] In addition, many authors have shown that backtracking improves when the consequent contains the modal *have to* (Lewis, 1973; Arregui, 2004; Schulz, 2007). In fact, Schulz (2007) argues that felicity with *have to*-consequents is a diagnostic for epistemic readings. In the analysis that follows, we will use *have to*-consequents to bias epistemic readings of the relevant examples. We will see, though, that NCs reject epistemic readings of even these biased counterfactuals.

does not work can allow the inference that it had rained. This explains the contrast with (29), which does not license the backtracking reading.

(29) #If I had made the car not start, it would have to have rained.

When the speaker causally intervenes on the state of the car, the backtracking epistemic reading disappears.

Since backtracking is only possible under epistemic readings of counterfactuals, the possibility of backtracking can be used as a test for the availability of epistemic readings. The crucial observation is that NCs pattern with examples like (29), and not standard counterfactuals like (27-28), with respect to backtracking. This is presented in (30-31).

(30) #If not for John being a bachelor, he would have to have had a wedding.

(31) #If not for the car starting, it would have to have rained.

Examples (30-31) do not permit a backtracking reading. They can only be true in the implausible situation where John's bachelorhood prevented him from being obliged to marry and the car's starting had an effect on local weather patterns.

The backtracking test shows that NCs do not have an epistemic reading, but are obligatorily ontic. A second type of epistemic inference forces the same conclusion. As discussed by Schulz (2007), standard counterfactuals are systematically ambiguous between ontic and epistemic inferences between two correlated variables. Correlations occur when two effects are anteceded by the same cause. When reasoning epistemically between two such variables, the correlation is maintained; that is, learning something about one variable provides information about the other. This is not the case for ontic interpretation because intervening to change one of the variables destroys the correlation. Schulz (2007, p. 118) presents dialogues like (32), which show the ambiguity for standard counterfactuals.[11]

(32) Suppose an alarm sounds at the docks whenever there is an impending storm. Further suppose we took the bridge instead of the ferry because we heard the alarm and there was, in fact, a storm.

   a. Thank goodness, if the alarm hadn't gone off, we would have taken the ferry and we might have all drowned in the storm.

[11] Schulz (2007) uses a low barometer as the correlated variable. I changed this because my speakers found the contrast clearer with an alarm.

b. No no no, that alarm always works. If it hadn't gone off, there would have to have been no storm coming.

Dialogues like (32) are possible because standard counterfactuals are ambiguous between ontic and epistemic readings. Example (32a) presents the ontic reading, where the state of the alarm is changed, independent of the storm and other correlated variables. The epistemic inference in (32b) allows for the value of the storm to change because we consider, not how the world would be different if the alarm had been forced to be different, but what we would know had we not observed the alarm. In the latter case, we would know that no storm was coming.

The ambiguity of counterfactuals in the context of correlations provides another test for epistemic readings. If a counterfactual allows both an epistemic and an ontic reading, it should be able to appear in both sides of an argument like that in (32). If it does not have an epistemic reading, it should only be felicitous as the ontic argument. NCs instantiate the latter possibility as seen in (33).

(33) Suppose an alarm sounds at the docks whenever there is an impending storm. Further suppose we took the bridge instead of the ferry because we heard the alarm and there was, in fact, a storm.

a. Thank goodness, if not for the alarm going off, we would have taken the ferry and we might have all drowned in the storm.

b. #No no no, that alarm always works. If not for the alarm going off, there would have to have been no storm coming.

The reply in (33b) can only be true under the strange situation where the alarm *caused* the storm, that is, they are no longer merely correlated. Once again, we find that NCs and standard counterfactuals are mutually paraphrasable under an ontic reading (32a-33a), but that NCs have no epistemic reading. When placed in a context where such a reading is forced, they are either false or aberrant depending on whether a causal relation can be constructed between antecedent and consequent, allowing the possibility of an ontic reading.

We have focused on cases of backtracking and correlations because they clearly show the importance of a relation, what we take to be a causal relation, in distinguishing epistemic and ontic readings. That being said, NCs provide a general probe for epistemic readings. They are precisely those that disappear when substituting a standard counterfactual for an NC. Consider, for instance, the following example from Hansson (1989).[12]

[12]This is similar to Veltman's (2005) *duchess* example that we discuss in (108-109).

(34) Suppose that one Sunday night you approach a small town of which you know that it has exactly two snackbars. Just before entering town you meet a man eating a hamburger. Suppose now that after entering the town, you see that A is in open. Would you now accept the following conditional?

a. If snackbar A were closed, then snackbar B would have to be open.

Those who deny the existence of epistemic readings, like (Rott, 1999; Veltman, 2005, a.o.), claim that examples like (34a) are infelicitous. NCs can add to the argument because they present a stark constrast with standard counterfactuals on this point, as we have already seen with clearer cases of backtracking and correlations. While many epistemic readings are hard to get with standard counterfactuals, they are hopeless with NCs. Consider example (35) in the same context.

(35) #If not for snackbar A being open, then snackbar B would have to be open.

Clearly the availability of an epistemic reading for (35) is more remote than (34a). The NC in (35) can only be used in a context where the state of snackbar A causally antecedes that of snackbar B, that is, snackbar B is closed because snackbar A is open. Unlike (34a), it simply cannot be used to describe what can be infered about the state of snackbar B given information about the town, the man with the hamburger, and the state of snackbar A. The fact that NCs present a clear contrast with standard counterfactuals in contexts that require epistemic readings is evidence that there is, in fact, a distinct epistemic reading. Since we have two different morphological forms, defining at least two different readings, we have a strong argument against both those who deny that counterfactuals have epistemic readings, for example Veltman (2005) or Rott (1999),[13] as well as those authors who want to reduce all counterfactual inferences to epistemic readings (Morreau, 1992).

After accepting that NCs reject epistemic readings, they become a useful test, not only for epistemic inferences, but also to to diagnose the flow of causality within a context. Specifically, given a context where $P \leftrightarrow Q$, we can use NCs to determine whether our internal models treat P or Q as causally dependent. The inference from the dependent variable will be infelicitous with an NC. Consider, for instance, Kratzer's (1989) famous example about King Ludwig of Bavaria (via Goodman (1947)).

[13]Though he makes an exception for reductio arguments, which he claims work because the epistemic inference is so far-fetched.

(36) King Ludwig of Bavaria likes to spend his weekends at Leoni Castle. Whenever the Royal bavarian flag is up and the lights are on, the King is in the Castle. At the moment the lights are on, the flag is down, and the King is away.

a. If the flag had been up, then the King would have been in the castle.

b. If the King had been in the castle, then the flag would have been up.

Since the flag is up if and only if the King is in the castle, the standard counterfactual can be used felicitously for inferences in both directions. In contrast, NCs are only felicitous with the latter inference.

(37) #If not for the flag being down, the King would have been in the castle.

(38) If not for the King being away, the flag would have been up.

Example (37) can only be used if the flag is actually preventing the King from coming to the castle, in opposition to (38), which makes a good paraphrase of (36b). From the behavior of the NCs we can conclude that in this context the state of the flag is a function of the King's position. This important to know because it tells us that (36a) is actually a backtracking counterfactual in the causal domain, which is not immediately clear because there is not the usual inversed temporal relationship between antecedent and consequent that accompanies common cases of backtracking. A consequence of this observation is that we can immediately explain why the following counterfactual is false (Kratzer, 1989, ex. 23).

(39) If I had hoisted the flag, the king would have appeared in the castle.

Kratzer (1989) argues that the antecedent of (39) demotes a non-accidental generalization about the position of the flag and the king to an accidental generalization that cannot be added as a premise for reasoning about the consequent. The results of the NC diagnostic suggest a different explanation, namely it should be treated like other cases of failed backtracking, as in (29). Given that the flag's position is causally dependent on the king's location, we know that reasoning from the flag's position to the king's location requires an epistemic inference. While an agent may be able to deduce the king's location after learning that the flag is in a hypothetically new state, changing the flag's state will not cause the king to be in the castle because it is not a causal antecedent of the king's location. In contrast, we predict that intervening on the king's location should allow the flag's position to change, which seems to be the case.

(40) If I had brought the king back to Leoni castle, the flag would be up.

While only a quick case study, this example shows how the empirical results of our investigation of NCs can be used to identify causal backtracking and to verify the causal relations that hold in a given context. While NCs do not reveal why particular causal relations hold and not others, NCs could be used in future work to help determine the linguistic resources that bias a listener to assume a certain causal structure.

2.3 Summary

This section detailed the core morphological and semantic properties of NCs that distinguish them from standard counterfactuals. On the morphological side, they consist of a nominalized proposition, the preposition *for*, and obligatory negation. On the semantic side, NCs have non-defeasibly counterfactual antecedents and only permit ontic readings. We also fit into the wider theory of counterfactuals, both as an argument for the existence of a distinct epistemic reading and as a diagnostic tool. The next section §3 synthesizes these two domains, providing an analysis of the unique semantics of NCs using only their core morphological differences.

3 THE SEMANTICS OF NCS

This section uses the nonstandard morphology of NCs to build a semantics that captures the generalizations presented in the previous section. Specifically, it argues that the eventive nominal embedded in an NC antecedent denotes a set of situations that exemplify a proposition (Kratzer, 2002). Due to a presupposition contributed by the preposition *for*, one of these situations is presupposed to be contained in the world of evaluation, which explains the non-defeasible counterfactuality of NC antecedents. Finally, the obligatory and eponymous negation of NCs is treated as constituent negation that removes the fact denoted by the nominal. We will account for the fact that NCs reject certain non-causal epistemic inferences after enriching our models with causal laws. Importantly, the analysis is rooted in the independent fact that NCs presuppose the counterfactuality of their antecedents, which is only implicated by standard counterfactuals.

3.1 Worlds, Situations, and Causal Models

The analysis that follows assumes a general familiarity with possible world semantics. Instead of treating possible worlds as indices for an interpretation, they will be interpretations themselves. Situations are partial interpretations and will play a role, both in the semantics of eventive nominals, as well as in defining causal models.

(41) **Definition. (Worlds and Situations)**
Let $\mathcal{P}$ be a finite set of atomic sentences.

i. A *world* w is a function from $\mathcal{P} \rightarrow \{0, 1\}$.

ii. A *situation* s is a partial function from $\mathcal{P} \rightarrow \{0, 1\}$.

iii. Let W be the set of all worlds. Define I as the set of all partial function s such that $\exists w \in W$ where $s \subseteq w$.

We work with the language $\mathcal{L}$ that is the closure of $\mathcal{P}$ under negation and conjunction, giving the usual recursive truth definitions for complex formula.

(42) **Definition. (Models and Propositions)**

i. A *model* M for $\mathcal{L}$ is a set of worlds W. For any formula ϕ of $\mathcal{L}$, we write $M, w \vDash \phi$ in case ϕ is true with respect to M and w.

ii. We write ϕ^M for the set of worlds $w \in W$ such that $M, w \vDash \phi$, and call this the *proposition* that ϕ.

The discussion of the distinction between ontic and epistemic counterfactuals in the previous section showed the importance of the causal structure of the context. Since NCs are infelicitous in contexts where the causal structure requires an epistemic interpretation against the flow of causality, to successfully analyze NCs the models must contain causal laws.

While most theories of counterfactuals recognize the importance of laws in their interpretation, including the original similarity-based approaches (Stalnaker, 1968; Lewis, 1973; Veltman, 1976), the causal flow of laws is not usually modeled. In contrast, we want to capture the crucial difference between (*P* iff *Q*) and (*P* iff *Q* and *P* causes *Q*). This is because under the ontic reading, a counterfactual intervention that affects a cause will also affect its effects, but it is not the case that a cause will be affected if a counterfactual intervention changes something about its effects.

A simple way to model the flow of causality is in terms of functions that determine the value of a variable based on its causal inputs. This is the influential approach taken by Halpern & Pearl (2005a,b) and Pearl (2000), who model causality with systems of structural equations (see Goldberger (1972) for a formal and historical overview). As long as these functions have certain properties, for instance, non-circularity, they can capture our intuitions about various types of causal dependence (Galles & Pearl, 1998).[14] Counterfactual inference is one such intuition, and their approach gives counterfactuals an interpretation where assuming the antecedent means changing the causal laws to ensure its truth (possibly breaking other laws). A schematic example is presented in (43), which assumes a system of two variables P, Q such that $P \iff Q$ and P causally antecedes Q (represented by the arrow).

(43)

Both P and Q are true in (43). In Pearl's system, in order to assume counterfactually that Q is false the causal relationship between P and Q must be broken and Q must be stipulated as false. This is shown in (44).

(44)

This is precisely what is needed for the interpretation of NCs and ontic counterfactuals more generally. For instance, taking this approach immediately explains why backtracking is not possible with NCs since enforcing the counterfactuality of the antecedent breaks its connection with its causal antecedents.

While the structural equations approach would work, it suffers from a few problems (see Schulz (2007, p. 113-121)), not the least of which is the implausible assumption that interpreting a counterfactual requires that the hearer alter the causal structure of the model. It would be better to follow the intuition of Lewis (1973) that the laws are not removed on making a counterfactual assumption, but only locally ignored. Schulz (2007) does precisely this, showing that we can mimic the effects of Pearl's causal interventions by locally manipulating interpretation functions in a way that allows causal laws to be violated. Her approach is now presented, which will form the basis for interpreting the antecedents of

[14]Moreover, Galles & Pearl (1998) show that for the counterfactual interpretation of causality, causal models with a recursive axiomatization impose no restrictions over those of Lewis's possible-worlds similarity approach.

NCs, though in the end, we will make the distinction between epistemic and ontic readings in a different way than she does.

First, the models must be enriched with a *Causal Structure*. This is a partition of the atomic formula into sets of those that are causally dependent and causally independent, along with a set of functions that allows the values of the dependent atoms to be determined based solely on the values of the independent atoms. These functions can be thought of as encoding the causal relations.

(45) **Definition. (Causal Structures and Models)** (Schulz, 2007, p. 141)
Given a finite set of atomic sentences $\mathcal{P}$, a model is a set of worlds W along with a causal structure $C = \langle B, E, F\rangle$ where:

i. $B \subseteq \mathcal{P}$ are *exogenous* variables.

ii. $E = \mathcal{P} - B$ are *endogenous* variables.

iii. F is a function mapping elements Y of E to tuples $\langle Z_Y, f_Y\rangle$, where Z_Y is an n-tuple of $\mathcal{P}$ and f_Y is a partial truth function $f_Y : \{0,1\}^n \longrightarrow \{0,1\}$. F is *rooted* in B.

(46) **Definition. (Rootedness)** (Schulz, 2007, p. 141-142)
Let $\mathcal{P}$ be a finite set of proposition letters and $\mathcal{L}$ the language obtained when closing $\mathcal{P}$ under negation and conjunction. Let $C = \langle B, E, F\rangle$ be a causal structure. We introduce a binary relation R_F on the set of atomic formula $\mathcal{P}$. $R_F(X, Y)$ holds if X occurs in $F(Y)$. Let R_F^T be the transitive closure of R_F. The R_F minima of a letter $Y \in \mathcal{P}$, $Min_{R_F}(Y)$, is defined as follows: $Min_{R_F}(Y) = \{X \in \mathcal{P} | R_F^T(X, Y) \& \neg Z \in \mathcal{P} : R_F^T(Z, X)\}$. We say that F is *rooted* in B if R_F^T is acyclic and $\forall Y \in \mathcal{P} - B : Min_{R_F}(Y) \subseteq B$.

The structure of the causal relations is given by the final clause (45iii). Every endogenous atom is associated with a set of atoms that it causal depends on and a partial interpretation that gives its value depending on the values of those atoms that it causally depends on. Rootedness ensures that when looking at the variables associated with an endogenous variable by F, if they are not exogenous, their F relations can be followed all the way back to exogenous variables (see (Schulz, 2007, p. 142) for more discussion of the rootedness property). This is needed so that the causal relations are not cyclic; knowing the values of the exogenous atoms determines the truth value of every endogenous atom (and thus every non-atomic formula). This is because of the additional constraint in (47b), namely that an endogenous variable Y will have the same truth value as the function f_Y applied to its n-tuple of causal antecedents.

(47) a. $\phi^{M,w} = w(\phi)$, if $\phi \in B$
b. $\phi^{M,w} = f_\phi(Z_\phi)$, if $\phi \in \mathcal{P} - B$

The truth value of a formula from the set of background variables is simply its value at the world of evaluation as seen in (47a). The new case is presented in (47b). If ϕ is endogenous, then its truth value is a function of the causal laws f_ϕ applied to the causal antecedents of ϕ, Z_ϕ.

What the causal models just defined provide is a more fine grained notion of law against which counterfactuals can be evaluated. One prominent way to incorporate laws into the interpretation of counterfactuals uses premise semantics. Recall that in premise semantics (Veltman, 1976; Kratzer, 1981a), the core idea is to define a function that returns a set of premises at every index. A counterfactual is true if minimal revision to the premise set to accommodate the antecedent results in a set of propositions that entails the consequent. Veltman (2005) enriches this system with two sets of premises, the laws and the facts, the latter of which is more easily given up to accommodate the counterfactual antecedent. Since the laws are more easily given up than the facts, Veltman (2005) predicts that inferences requiring epistemic readings, like backtracking, are necessarily false.

By contrasting NCs and standard counterfactuals, we have argued for a distinct reading of counterfactuals in which the facts are given up to preserve laws, so we will have to part from Veltman (2005) on this point. That being said, we follow his lead in distinguishing the contribution of laws and facts by using the notion of the *basis* of a world, which is the smallest set of facts that along with the laws, derives all the other facts that characterize a world. The main different between Veltman's (2005) basis and ours is that we also encode the causal flow of the laws.That is, given a world with only two facts P and Q, and the law $P \rightarrow Q$, Veltman's approach would say that $P = 1$ is just as good a basis as $Q = 0$ because either allows one to derive the rest of the facts. In providing the causal basis of a world where $P \rightarrow Q$ and P causes Q, we want to allow $P = 1$ to be a basis, but not $Q = 0$. The reason is that we want the basis to be centered on those facts that are not causally dependent. Schulz (2007) defines such a basis as follows:

(48) **Definition. (Basis)** (Schulz, 2007, p. 144)

i. The basis b_w of a world $w \in W$ is the union of all interpretation functions $b \in I$ that fulfill the following two conditions:
   a. $b \subseteq w \subseteq \bar{b}$
   b. $\neg\exists b' : b' \subseteq w \subseteq \bar{b}$' & $b' \subset b$, where $\bar{b}$ is the closure of b under the causal laws.

ii. Let M be a causal model and $i \in I$ a partial interpretation of $\mathcal{P}$. The *causal closure* $\bar{i}$ of i is the minimal i' in I fulfilling the following conditions:
   a. $i \subseteq i'$
   b. $i' = \bigcap \{w \in W | i' \subseteq w\}$
   c. for all $P \in E$ with $Z_P = P_1, ..., P_n$ such that $i(P)$ is undefined the following holds: if for all $k \in \{1, ..., n\} : i'(P_k)$ is defined and $f_P(i'(P_1), ..., i'(P_n))$ is defined, then $i'(P)$ is defined and $f_P(i'(P_1), ..., i'(P_n)) = i'(P)$ (see (Schulz, 2007, p. 274) for the proof that the causal closure is unique).

Definition (48i) says that the basis of a world w is the smallest partial interpretation function that has w as its closure under the causal laws. Causal closure is defined in (48ii) as the interpretation that extends a partial interpretation function i with the interpretation of those variables that can be derived by the laws and the variables in the domain of i. A basis can be thought of as encoding the facts that characterize a world. It is the smallest interpretation function that, along with the causal laws, can derive all there is to know about a world.

When interpreting a counterfactual, the consequent must be verified only in those worlds that (i) satisfy the antecedent, (ii) have minimally different bases, and (iii) are minimally different in the laws that hold. Example (49) defines a notion of similarity that fulfills these requirements.[15]

(49) **Definition. (Similarity)** (Schulz, 2007, p. 145)
   i. Define $\leq$ mapping w to the order: for $w_1, w_2 \in W$: $w_1 \leq_w w_2$ iff
      a. $b_{w1} \cap b_w \supseteq b_{w2} \cap b_w$
      b. if $b_{w1} \cap b_w = b_{w2} \cap b_w$, then $b_{w1} - b_w \subseteq b_{w2} - b_w$.

The order defined by $\leq_w$ says that a world w_1 is more similar to w than w_2 iff its basis overlaps more with w or, in case they are equal in this respect, the part of its basis that does not overlap is smaller than that of w_2. The fact that it is smaller

[15] Schulz (2007, p. 145) defines a second order that is used to reckon similarity only after generating a set of similar worlds with respect to the order in (49). Specificially, this second order ranks worlds with respect to the derivable facts. We leave out this second round of similarity ordering because, for the examples considered here, it only matters that basis similarity counts more than derivable similarity, which is true whether we calculate derivable similarity second or not at all. In the way, the analysis of similarity is more like that in Veltman (2005), which also does not reckon similarity with respect to derivable facts.

means that more of the causal laws are left intact since fewer stipulated facts are needed to characterize the world.

Since similarity is defined in terms of a world's basis, and the basis is dependent on the flow of causal laws, locally changing upstream variables counts more for the purposes of similarity than those downstream. The truth table in (50) will help illustrate an example. It represents a situation with two variables P, Q, where $P \iff Q$ and P causally antecedes Q. The latter two facts are represented by the $\boxed{\text{boxes}}$, which encode the basis of each world. For example, the basis of w_1 is $P = 1$, since $Q = 1$ can be derived $P \iff Q$ and the fact that P causally antecedes Q. In contrast, both the values of P and Q are in the basis of world w_2 since the laws are violated in this world, so each fact has to be stipulated in the basis.

(50)

	P	*Q*
@w_1	$\boxed{1}$	1
w_2	$\boxed{1}$	$\boxed{0}$
w_3	$\boxed{0}$	$\boxed{1}$
w_4	$\boxed{0}$	0

Starting at w_1, consider the most similar worlds where $Q = 0$. There are two worlds where $Q = 0$, namely w_2 and w_4. According to clause (56i-a), which reckons the similarity of bases, the most similar world where $Q = 0$ is w_2 due to the fact that its basis overlaps with that of w_1. Now consider the closest worlds where $P = 0$. Once again, there are two options, namely w_3 and w_4. Notice that in this case, though, the closest world will be w_4, which alters the value of Q. This is because w_3 and w_4 are equal with respect to clause (56i-a), so the closest world with respect to w_1 is the world where more of the causal laws are left intact, per (56i-b). Thus, when Q is altered, P is not altered, but when P is altered, Q must change as well.

The crucial point is that variables that lie upstream with respect to the flow of causality are harder to alter because they count more for similarity. Employing a notion of similarity like (49) in the causal models presented here suppresses classic epistemic inferences like backtracking or inferences between correlated variables. This represents the usual case of counterfactual reasoning, which Schulz (2007) argues always takes place via local surgery on a single world within a similarity structure like those presented here. She argues that epistemic readings arise via a completely different method of counterfactual revision involving a separate definition of basis and similarity reckoned globally over sets of worlds in a be-

lief state. While we wait until §3.4 to compare approaches, we show now how to generate epistemic readings of counterfactuals by using the same definition of similarity, but relativizing the notion of basis to what a speaker takes to be the case. The result is that NCs will block epistemic readings because, unlike standard counterfactuals, they presuppose the antecedent is counter-to-fact.

3.2 Epistemic Bases and Two Readings for Counterfactuals

The notion of basis developed in the previous section is completely objective in that it does not depend on what the speaker knows.[16] That is, a world may or may not be compatible with what a speaker takes to be true, but its basis is not affected by these considerations. In order to capture the distinction between epistemic and ontic readings of counterfactuals, we introduce the notion of an epistemic, or *speaker's basis* of a world. Just the basis of a world is the set of facts, that along with the laws, derive everything there is to know about a world, the speaker's basis is the set of facts, that along with the laws, derive everything the speaker knows about a world. We define it in two steps.

First, we distinguish sets $C_\sigma \subseteq U \subseteq W$, where U is the set of all worlds compatible with the causal laws and C_σ is the set of all worlds compatible with what the speaker σ takes to be the case. The set C_σ, which we will call an epistemic state, is a subset of U because we assume that the speaker knows all of the laws. This is an idealization, but it is not crucial for the simple examples considered here, and we could always index U to various interlocutors and provide a method to update the known laws, as in Veltman (2005). Given that C_σ is a set of assignments, the *speaker's facts* in an epistemic state C_σ is the overlap of the worlds in C_σ, as defined in (51).

(51) **Definition**. (Speaker's Facts)

i. Call the situation s_σ the speaker's facts in an epistemic state C_σ iff
   a. $s_\sigma \subseteq w$ for all $w \in C_\sigma$
   b. there is no $s' \supset s_\sigma$ such that $s' \subseteq w$ for all $w \in C_\sigma$.

If a speaker knows everything there is to know, then the s_σ characterizing her facts in C_σ is simply the single w in C_σ, namely the actual world. With the help of the speaker's facts, we can define the promised *speaker's basis* as below. The strategy is to take the objective basis of each world in C_σ and restrict it to the

[16] Veltman's (2005) non-causal basis is as well, for that matter.

minimal interpretation whose causal closure makes the same sentences true as the speaker's facts, s_σ.

(52) Definition. (Speaker's Basis)
Let M be a causal model, C_σ the epistemic state of a speaker σ, and s_σ the speaker's facts in C_σ. For each w in C_σ, the ***speaker's basis*** b^σ_w of w is the minimal interpretation $i \in I$ meeting the following conditions:

i. $i \subseteq b_w$

ii. For every atomic sentence $P \in \mathcal{P}$, $\bar{i}(P) = s_\sigma(P)$ if $s_\sigma(P)$ is defined.

We can think of the speaker's basis as giving the initial conditions for each world in C_σ that, along with the causal laws, generate the speaker's facts. This is illustrated in the following examples. Suppose that we have three variables, P, Q, and R, where $(\neg P \iff Q) \iff R$, and P and Q causally antecede R. This corresponds to the lightswitch example from Lifschitz (1987), where some light R is on just in case two independent switches P and Q are in the different positions. Example (53) present the objective basis of each possible world in this situation.

(53)

	P	*Q*	*R*
w_1	1	0	1
w_2	0	1	1
w_3	1	1	0
w_4	0	0	0
w_5	1	1	1
w_6	1	0	0
w_7	0	1	0
w_8	0	0	1

Suppose that a speaker σ knows that the switch P is on, but nothing else. The speaker's epistemic state is greyed out in (54) and the speaker's basis is represented by underlining, which we now show how to derive.

(54)

	P	*Q*	*R*
w_1	1	0	1
w_2	0	1	1
w_3	1	1	0
w_4	0	0	0
w_5	1	1	1
w_6	1	0	0
w_7	0	1	0
w_8	0	0	1

The speaker's facts s_σ in (54) is the assignment $\{\langle P, 1\rangle\}$ since this is what is constant across the epistemic state. To get the speaker's basis for each world in C_σ, we take the smallest assignment contained in its basis that along with the causal laws, agrees with s_σ everywhere it is defined. Since s_σ is defined only for P and it has no causal antecedents, the speaker's basis for w_1 and w_3 is nothing more than the single basis fact $\{\langle P, 1\rangle\}$. In this case, the speaker's basis is smaller than the objective basis because the speaker does not know much, and what is known neither decides nor is decided by much.

Now consider a case where a speaker σ knows that the light is on, but does not know the position of the switches. The speaker's epistemic state is greyed out in (55).

(55)

	P	*Q*	*R*
w_1	1	0	1
w_2	0	1	1
w_3	1	1	0
w_4	0	0	0
w_5	1	1	1
w_6	1	0	0
w_7	0	1	0
w_8	0	0	1

The speaker's facts s_σ in (55) is clearly the assignment $\{\langle R, 1\rangle\}$. To get the speaker's basis for each world in C_σ, we take the smallest assignment contained in its basis that along with the causal laws, agrees with s_σ everywhere it is defined. Since s_σ is defined only for R, and its causal antecedents are P and

Q, the speaker's basis for w_1 and w_2 is the same as the objective basis, that is $\{\langle P,1\rangle, \langle Q,0\rangle\}$ and $\{\langle P,0\rangle, \langle Q,1\rangle\}$ respectively.

Finally, we need to define a new notion of similarity to take into account that the speaker's basis can differ from the objective basis of a world.

(56) **Definition. (Speaker Similarity)**

i. Define $\leq^\sigma$ mapping $w \in C_\sigma$ to the order: for $w_1, w_2 \in W$: $w_1 \leq_w w_2$ iff

a. $b_{w1} \cap b^\sigma_w \supseteq b_{w2} \cap b^\sigma_w$

b. if $b_{w1} \cap b^\sigma_w = b_{w2} \cap b^\sigma_w$, then $b_{w1} - b^\sigma_w \subseteq b_{w2} - b^\sigma_w$.

The order defined by $\leq^\sigma$ says that a world w_1 is more similar than w_2 to a world w compatible with what the speaker takes to be the case iff its basis overlaps more with the speaker's basis of w or, in case they are equal in this respect, the part of its basis that does not overlap is smaller than that of w_2. The fact that it is smaller means that more of the causal laws are left intact since fewer stipulated facts are needed to characterize the world.

With an understanding of how the epistemic, or speaker's basis of a world can differ from its objective basis, we now show how the epistemic-ontic ambiguity arises with standard counterfactuals. The result is contrasted with the analysis of NCs in §3.3, where we explain why epistemic readings cannot arise. The core idea is that NCs presuppose that their antecedents are counter-to-fact, necessitating a larger speaker basis against which similarity is calculated.

We start with a simple set of background laws like (57), which we have seen before. We have two variables P and Q, where the first causally antecedes the second and $P \iff Q$. That is, the model will associate the variable Q with the function f_Q that gives the value of Q as a function of P.

(57) P causally antecedes Q

$$f_Q(P) : \begin{cases} 1 \to 0 \\ 0 \to 1 \end{cases}$$

The space of possible values to P and Q is given in (58) along with the objective basis of each world.

(58)

	P	*Q*
w_1	**1**	1
w_2	**0**	0
w_3	**1**	**0**
w_4	**0**	**1**

We interpret a *would*-counterfactual, *If it has been the case that P, it would have been the case that Q*, with the revision operator in (59).

(59) *If it had been the case that P, it would have been the case that Q^M=1* iff $Rev(C_\sigma, P) \vDash Q$ where:
   i. $Rev(C_\sigma, P) = \bigcup_{w \in C_\sigma} Min(\leq_w^\sigma, P^M)$
   ii. $Min(\leq_w^\sigma, P^M) = \{w' : w' \in P^M \ \ \& \ \ \neg\exists w''(w'' \in P^M \ \ \& \ \ w'' <_w^\sigma w')\}$

That is, we take each world w in the speaker's epistemic state C_σ and calculate closest worlds relative to $\leq_w^\sigma$ where the antecedent is true. The counterfactual is true if the consequent is entailed by the union of these worlds. This is just the standard similarity semantics story. The difference is that in revising to accomodate the antecedent, as we have seen, the similarity measure requires preferential faithfulness to the basis facts, whatever those may be.

There is one more aspect to the evaluation of counterfactuals that needs to be discussed before considering a few examples. We saw in §2.2 that standard counterfactuals implicate the counterfactuality of their antecedents, but we did not discuss how that implicature arises. To accomplish this, we follow the analysis of Iatridou (2000). She argues that just like past tense asserts that the topic time excludes the utterance time, the fake past tense in counterfactual antecedents asserts that the topic worlds (those that satisfy the antecedent), exclude the worlds of the speaker.[17] Since the speaker is not willing to assert that the actual world is among the topic worlds satisfying the antecedent, the implicature is that it does not satisfy the antecedent, yet the speaker never asserts this. Translating her account, $Rev(C_\sigma, P)$ corresponds to the topic worlds, that is, worlds that satisfy the antecedent against which we evalutate the consequent. The set C_σ are the speaker's worlds, that is, those worlds compatible with everything the speaker takes to be the case. To say that the topic worlds exclude the speaker's worlds is equivalent to (60), which we take all counterfactuals to assert.

(60) $C_\sigma \nsubseteq Rev(C_\sigma, P)$

The condition in (60) holds both in cases where an antecedent P is false throughout C_σ or where P is undecided, but implied to be false. These are the two cases in which a counterfactual can be used, as we saw in §2.2. Crucially, each of these

[17] Note that the expanded NCs, just like standard counterfactuals have *fake* past tense for those speaker's without subjunctive *were*: *If it **was/*is** not for John being tall, he would have been able to dunk.*

two cases will yield a different speaker's basis. We can now see what happens when we interpret various counterfactuals with respect to these different sets of basis facts.

First consider a counterfactual that follows the flow of causality in a context where P causes Q. When a speaker utters *If it had been the case that P, it would have been the case that Q*, she implicates that she is in an epistemic state where P and Q are false. The speaker's epistemic state is grayed in (61). Notice that the objective basis of world w_2 and the speaker's basis, indicated by underlining, are the same. This is because the speaker has all of the facts about w_2.

(61)

	P	*Q*
w_1	1	1
w_2	0	0
w_3	1	0
w_4	0	1

We must now interpret $Rev(C_\sigma, P^M)$ to see if the resulting set of worlds entails Q. The derivation procedes as follows, where similarity reckoned with respect to the speaker's basis of each world in C_σ is given in (62). We only care about worlds that satisfy the antecedent, so we will suppress the similarity calculation for worlds that do not.

(62) $\mathbf{w_1} <^{\sigma}_{w_2} w_3$

(63)
i. $Rev(C_\sigma, P^M) \vDash Q$ iff
ii. $\bigcup_{w \in C_\sigma} Min(\leq_w, P^M) \vDash Q$ iff
iii. $\bigcup \{w_1\} \vDash Q$ iff
iv. ✓$w_1 \vDash Q$

The counterfactual is true because the closest world(s) to w_2 that (i) keep most of the speaker basis facts intact, and (ii) keep most of the laws intact, is w_1, where Q holds.

While the counterfactual is true in the epistemic state just considered, it is not the only state in which the counterfactual could be uttered. Since the counterfactuality of P and Q is only implicated, *If it had been the case that P, it would have been the case that Q* can also be interpreted in an epistemic state like (64).

(64)

	P	Q
w_1	1	1
w_2	0	0
w_3	1	0
w_4	0	1

Note here that the speaker's basis is empty. This is because neither P nor Q is settled in this epistemic state. Since there are no basis facts, similarity will be decided by those worlds that keep most of the causal laws intact. Similarity is calculated with respect to each world in C_σ as in (65-66).

(65) $\mathbf{w_1} <^{\sigma}_{w_1} w_3$

(66) $\mathbf{w_1} <^{\sigma}_{w_2} w_3$

To interpret the counterfactual, we use the similarity rankings in (65-66) to compute the closest world(s) to each world in C_σ that satisfy the antecedent to see if their union entails the consequent. In this case it does.

(67) i. $Rev(C_\sigma, P^M) \vDash Q$ iff
ii. $\bigcup_{w \in C_\sigma} Min(\leq_w, P^M) \vDash Q$ iff
iii. $\bigcup \{\{w_1\}, \{w_1\}\} \vDash Q$ iff
iv. ✓ $w_1 \vDash Q$

What these examples show is that way the counterfactuality implicature is resolved does not matter when the inference follows the flow of casuality. That is, we get the same result if the antecedent is taken as false, as the speaker implies, or if it is not settled, which the speaker does not rule out by asserting only $C_\sigma \nsubseteq Rev(C_\sigma, P)$.

This is not the case when the inference from antecedent to consequent goes against the flow of causality. If the antecedent is false, as implied, epistemic inferences are necessarily false. They can only be true in the latter case where the antecedent interpreted against a context in which its truth is undecided. We show this now with a universe the same as before, repeated in (68-69).

(68) P causally antecedes Q

$$f_Q(P) : \begin{cases} 1 \rightarrow 0 \\ 0 \rightarrow 1 \end{cases}$$

(69)

	P	*Q*
w_1	1	1
w_2	0	0
w_3	1	0
w_4	0	1

We want to interpret the backtracking counterfactual, *If it had been the case that Q, it would have been the case that P*. We begin with the case where the antecedent is false as implied. The speaker's epistemic state is in gray. The speaker's basis, that is, the smallest set of facts necessary to derive everything the speaker knows, is underlined, which is the same as the objective basis.

(70)

	P	*Q*
w_1	1	1
w_2	0	0
w_3	1	0
w_4	0	1

Interpreting the counterfactuals means finding the closest worlds to w_2 where the antecedent is true and then checking if the consequent is entailed. There are two worlds where the antecedent is true, namely w_4 and w_1. According to the definition of similarity in (49), w_4 is closer because it keeps more of the facts from the speaker's basis. But then the counterfactual must be false, as is clear from (71-72).

(71) $\mathbf{w_4} <^{\sigma}_{w_2} w_1$

(72)
i. $Rev(C_\sigma, Q^M) \vDash P$ iff
ii. $\bigcup_{w \in C_\sigma} Min(\leq_w, Q^M) \vDash P$ iff
iii. $\bigcup \{w_4\} \vDash Q$ iff
iv. ✗ $w_4 \vDash Q$

In fact, if the truth of the antecedent is settled, backtracking inferences in a model that incorporates causal laws and with a definition of similarity that privileges basis facts, will necessarily be false. The reason is that if the antecedent is settled, the facts that casually antecede it will be in the speaker's epistemic basis. Since basis facts are harder to give up than the laws, the closest worlds that satisfy the antecedent will be worlds where the consequent is false.

The situation is very different if the context where the speaker does not assume a fact of the matter concerning the antecedent. The epistemic state in (73) presents such a situation.

(73)

	P	*Q*
w_1	1	1
w_2	0	0
w_3	1	0
w_4	0	1

Interpreting the counterfactual *If it had been the case that Q, it would have been the case that P* means finding, for each of the worlds in the speaker's epistemic state, the closest worlds satisfying the antecedent. For the counterfactual to be true, the consequent must be entail by all of these worlds. Examples (74-75) show the most similar worlds to w_1 and w_2 given the speaker's epistemic basis.

(74) $\mathbf{w_1} <^{\sigma}_{w_1} w_4$

(75) $\mathbf{w_1} <^{\sigma}_{w_2} w_4$

Importantly, since the speaker has no facts concerning Q, the speaker's basis is empty. This means that w_1 is now closer to w_2 than w_4 since both are equivalent in terms of basis overlap with w_2, yet w_1 keeps more of the causal laws intact. Now the backtracking inference goes through.

(76) i. $Rev(C_\sigma, Q^M) \vDash P$ iff
ii. $\bigcup_{w \in C_\sigma} Min(\leq_w, Q^M) \vDash P$ iff
iii. $\bigcup \{\{w_1\}, \{w_1\}\} \vDash P$ iff
iv. $w_1 \vDash P$

The inference goes through because when the antecedent is not settled, there are fewer facts in the speaker's epistemic basis. When these facts are not fix, lawful worlds win out, allowing backtracking to procede.

To sum up, the four examples just considered cross two factors: (i) whether the antecedent is taken to be false or not, and (ii) whether the consequent lies upstream or downstream from the antecedent with respect to the causal laws. In an inference that flows downstream, whether the antecedent is counterfactual as the antecedent implies, or not settled, which is not ruled out by the implicature, does not affect the result. Only when reasoning against the flow of causality does the truth of the antecedent come into play. If the antecedent is truly counter-to-fact, an epistemic inference like backtracking is necessarily false. In contrast, if the truth value of the antecedent is not settled, which is compatible with the antecedent's implicature, then backtracking is possible.

A crucial factor in generating this asymmetry is the notion of epistemic basis, which is like Schulz's (2007) basis, except that that it is generated after considering what the speaker takes to be fact. We have seen that standard counterfactuals are ambiguous as to whether their antecedents are false in the actual world, or merely undecided. In this view, epistemic inference like backtracking come about in the later case, where an undecided antecedent fixes fewer basis facts that are resistent to change. In the next section, we so that this option is not available for NCs. By presupposing the counterfactuality of their antecedents, their use fixes those facts that causally antecede the antecedent. The result is that the consequent can not concern one of these upstream facts, blocking epistemic inferences like backtracking.

3.3 The Evaluation of NCs

This section presents an analysis of the evaluation of NC antecedents within the causal models developed in the previous section. We show that the semantic contribution of NC antecedents explains the restricted range of inferences available with NCs. At each point of the analysis, the semantic contribution of NC antecedents is closely tied to their unique morphology. The eventive nominal in NC antecedents is analyzed as denoting a fact that exemplifies a proposition. In conjunction with obligatory presence of negation, this explains the non-defeasible counterfactuality of NCs. Negation is interpreted as a type of constituent negation taking a fact and a world and returning the most similar worlds without that fact. Since this presupposed fact fixes its causal antecedents in the speaker's basis, epistemic inferences are blocked.

The first piece of morphology to be considered in building a semantics for NCs are the eventive nominals found embedded in the antecedent. There is a long tradition that interprets such nominals as denoting something that is smaller than a proposition, but with similar informational content (Chierchia, 1984; Portner, 1992, a.o.).[18] This approach is especially attractive in a Kratzerian Situation

[18]The strongest arguments are based on differences in the semantic distribution of eventive nominals and *that*-clauses, which plausibly denote propositions. For instance, *that*-clauses can be subjects, but unlike eventive nominals, they cannot have causal import (Asher, 1993, ch. 4).

(1) John singing upstairs gave me a headache.

(2) #That John was singing upstairs gave me a headache.

Similarly, *that*-clauses can be true or false, while eventive nominals can be neither (Asher, 1993, ch. 4).

Semantics framework (Kratzer, 1989), where eventive nominals can denote in the space of situations, which are parts of worlds. A similar account is pursued here in the formal system just developed where situations are partial interpretation functions.[19]

Following an idea from Kratzer (2002), eventive nominals in NCs will denote situations that *exemplify* a proposition. Such situations contain all and only the information relevant for the truth of a proposition. In order to define exemplification clearly, a few other useful notions must be defined, starting with *forces* (Veltman, 2005).

(77) **Definition. (Forces and Minimally Forces)**

i. Say that s *forces* ϕ iff $\forall w(s \subseteq w \rightarrow w \in \phi)$

ii. Say that s *minimally forces* ϕ iff

a. s forces ϕ

b. There is no $s' \subseteq s$ such that s' forces ϕ

Now we can define a situation that exemplifies a proposition as a situation that forces a proposition containing no superfluous information. We filter off this information as in (78).

(78) s *exemplifies* ϕ iff for all $s' \subseteq s$ such that s' does not force ϕ, there is an s'' such that $s' \subseteq s'' \subseteq s$ and s'' minimally forces ϕ.

The crucial aspect of (78) is that if s contains superfluous information, then there will be a subpart that does not force ϕ and which cannot be extended to an interpretation that minimally forces ϕ. We can now define the Kratzerian minimalization operator $\downarrow$ that takes a formula ϕ and maps it to the situations that exemplify the proposition ϕ.

(79) **Definition.** ($\downarrow$)
$\downarrow\phi = \{s | s \text{ exemplifies } \phi\}$

---

(3) That John is singing upstairs is true/false.

(4) John singing upstairs is #true/#false.

[19]This implementation, though mirroring the partiality of a Situation Semantics approach, is different in that our situations are not first order objects. That is, our situations are not reified into model-stuff that can enter into relations. None of the examples we consider hinge on this point, though it is important to make clear.

In what follows, the analysis assumes (though it is surely a oversimplification), that eventive nominals uniformly denote sets situations that exemplify a proposition. Thus, the operator $\downarrow$ can be thought of as the semantic counterpart to a nominalization operator, and along these lines, the nominalization of a proposition ϕ will be translated as $\downarrow\phi$.

Section §2 showed that NCs cannot be used in informative modus tollens arguments or in arguments for the antecedent. It was suggested that this was due to a presupposition of counterfactuality, but now that we have a semantics for the nominals in NC antecedents, the locus of this presupposition can be found and its introduction modeled explicitly. Specifically, we propose that the preposition *for* imparts a factive presupposition to its nominal complement.

The primary argument is that in other contexts where *for* takes an eventive nominal complement, the nominal must be interpreted factively as shown in (80). For example, (80a) requires that Mary has indeed spilt coffee all over the desk.

(80) a. I blame Mary for coffee spilling all over the desk.
b. They impeached the Mayor for his lying about the budget.
c. I am angry at Bill for the boat wrecking.

The fact that these inferences are presuppositions is confirmed by the examples in (81), which show that the factive inference projects out of standard presupposition holes.

(81) a. Who do you blame for coffee spilling all over the desk?
b. They didn't impeach the Mayor for his lying about the budget.
c. I am not angry at Bill for the boat wrecking.

The proposal that NCs contain a factive presupposition receives support from the fact that NCs in other languages also employ factive morphology. For instance, Spanish NCs, as exemplified in (16), contain the preposition *porque* 'because', which imposes a factive presupposition on its clausal complement. It is also interesting to note that *for* in the examples in (80), like *porque* 'because', has a causal/reason flavor.

A final argument that *for* contributes a factive presupposition is that NPIs are unavailable in NC antecedents (82). This is intially surprising given that negation is obligatorily present.

(82) a. *If not for John advising any students, he wouldn't have gotten tenure.
b. *If not for John borrowing a red cent, I would still have my money.

c. *If not for John ever arriving, we wouldn't have started the meeting.

We account for these facts under the assumption that NPIs are not licensed in factive contexts by negation outside the scope of factivity. For example, *for*-adverbials block NPI licensing from clausemate negation as in (83).

(83) a. *I didn't blame Mary for ever showing up.
b. *I am not angry at John for advising any students.
c. *They didn't impeach the mayor for stealing a red cent.

Given that *for* contributes a factive presupposition, NPIs are predicted to be unable to appear in NCs as well.[20]

These considerations support an analysis where *for* is a presupposition trigger, taking a predicate of situations and appending a factive presupposition. In the framework developed here, for a gerund with a factive presupposition to be admissible, a situation in its denotation must be a speaker fact, that is, a part of every world in the speaker's epistemic state. Since every situation in the denotation of the gerund exemplifies a proposition, this proposition will be presupposed to be true.

An important consequence of this analysis is that it explains why NCs are non-defeasibly counter-to-fact, whereas standard counterfactuals only implicate their counterfactuality. The factive presupposition on the gerund ensures that the proposition the gerund exemplifies is true in the actual world, which is not cancellable. NCs become counterfactual when negation is appended, which negates the gerund's meaning, but does not block the presupposition. The result is that the consequent is interpreted against worlds where the gerund does not hold, which must be counterfactual worlds since the gerund characterizes a fact of the actual world.

In fact, we can show a stronger result, namely that negation is a necessary property of these counterfactuals, which we saw was a crosslinguistically stable property of these non-standard counterfactual constructions. What would happen if we had an NC with no negation, that is, a counterfactual with a presupposed

[20] In considering Mandarin NCs, Ippolito & Su (to appear) argue that the factive presupposition is contributed by light negation (Schwarz & Bhatt, 2006). For the purposes of this paper, both approaches work equally well, and so we will not try to distinguish them here. Their approach potentially better generalizes across languages since all NCs contain negation, though we have an explanation for this fact as well. Moreover, the fact that NCs sometimes contain "reason" prepositions, which are often factive, is still telling and left unexplained in accounts that put the presupposition in negation.

antecedent and without a higher-scoping negation? Recall that a counterfactuals requires the topic worlds to exclude the speaker worlds, that is, $C_\sigma \nsubseteq Rev(C_\sigma, P)$. But if the antecedent were presupposed and there were no higher scoping negation, then the closest worlds to C_σ where P were true would just be C_σ. Thus $C_\sigma \subseteq Rev(C_\sigma, P) = C_\sigma$ and $C_\sigma \nsubseteq Rev(C_\sigma, P) = C_\sigma$, which is a contradiction. The result is that a counterfactual that presupposes its antecedent content, like NCs, must have a higher scoping negation to remove the presupposed fact, generating counterfactual worlds against which to evaluate the consequent. We now present our proposal for NC negation.

First, note that negation in English NCs cannot be the standard propositional operator defined via complementation; there is no proposition for it to operate on and negation via complementation does not work well with partial worlds or partial interpretations (see, for instance, Kratzer (1989, p. 643-651)). Moreover, negation must provide a bridge from the denotation of eventive nominals to possible worlds. The reason is that only a set of worlds can entail the consequent, and since NC consequents are indistinguishable from the consequents of standard counterfactuals, we should make the standard assumption that the truth of a *would*-counterfactual depends on whether the consequent is entailed by the context after it is updated with the antecedent. With these concerns in mind, NC negation is interpreted as the model update function in (84) that takes a set of worlds and a fact that exemplifies a proposition and returns the most similar worlds where that fact has been removed. Similarity is calculated with the help of the function Min, which, as defined in (59ii), takes a set of worlds and generates the minimal subset with respect to an ordering relation.[21]

(84) **Definition. ($\textbf{Remove}_\textbf{M}$)**

a. $Remove_M(C_\sigma, \downarrow\phi) = \bigcup_{w \in C_\sigma} Min(\leq^\sigma_w, \{w' \mid \forall s(s \in \downarrow\phi^M \rightarrow s \nsubseteq w')\})$

The definition in (84a) says that ***Remove***$_M$ is a function from an epistemic state and the facts exemplifying ϕ to a set of worlds without these facts that are, due to the definition of $\leq_w$, maximally similar to the worlds in C_σ.

Since NCs obligatorily contain negation and a nominalization, NC antecedents will necessarily require the update defined in (84). We give their interpretation now. For simplicity, we will not allow for arbitrarily embedded NCs.

[21] Nothing crucial hinges on rolling an evaluation of similarity into the defintion of negation, though it makes for simplier formula later on. We could just have easily had negation give the following proposition $\{w' \mid \forall s(s \in \downarrow\phi^M \rightarrow s \nsubseteq w')\}$ when applying to event nominals, which we could then feed to the revision operator we defined for standard counterfactuals, that is $Rev(C_\sigma, \{w' \mid \forall s(s \in \downarrow\phi^M \rightarrow s \nsubseteq w')\})$.

(85) **Definition. (Language with *would*-NCs)**
Give a set of propositional letters $\mathcal{P}$ and our language $\mathcal{L}$, which is the closure of $\mathcal{P}$ under negation and conjunction, the language $\mathcal{L}^{\succ}$ is the union of $\mathcal{L}$ with the set of sentences of the form $\downarrow\phi \succ \psi$ for all $\phi, \psi \in \mathcal{L}$.

We interpret the language of NCs in our causal models as defined in (45).

(86) **Definition. (Interpretation of *would*-NCs)**
Let M be a causal model and $w \in W$ a possible world.

a. $M, w \vDash \downarrow\phi \succ \psi$ iff $Remove_M(w, \downarrow\phi) \vDash \psi$

It is now possible to see how this correctly predicts the restricted class of readings that are available with NCs. All epistemic readings that require an inference against the flow of causality will be necessarily suppressed.

Recall that NCs reject backtracking inferences, like that in (87), repeated from (28). The standard counterfactual allows an inference about the state of rain at some previous time based on the state of the car (87a). This reading is necessarily false with the NC (87b); it seems odd because it can only be true under the strange reading where the car starting prevented it from raining in the past.

(87) Suppose that the car works perfectly except for after a rain, which always causes the wiring to short out. Supposed it didn't rain and the car started as usual.

a. If the car hadn't started, it would have to have rained.

b. #If not for the car starting, it would have to have rained.

The background situation in (87) is such that the rain R and whether the car starts S are causally connected. This state of affairs can be represented as in (88). The model will associate the variable S with the function f_S that gives the value of S as a function of R.

(88) R causally antecedes S

$$f_S(R) : \begin{cases} 1 \rightarrow 0 \\ 0 \rightarrow 1 \end{cases}$$

The truth table in (89) presents the universe that covers the logically possible truth values of R and S. Recall that the $\boxed{\text{box}}$ indicates the basis of each world given the causal structure. For instance, the basis of w_2 is $\langle R, 1\rangle$ since we can derive $\langle S, 0\rangle$ from the causal laws. World w_4 where the causal laws are broken must

concomitantly have a larger basis. The pair $\langle S, 1\rangle$ has to be represented in the basis because it cannot be derived from the laws due to the fact that this set of facts violates the laws.

(89)

	R	*S*
w_1	0	0
w_2	1	0
w_3	0	1
w_4	1	1

By using an NC in the scenario above, the speaker presupposes that the car started. Thus, the speaker's state C_σ includes just w_3 because it has the fact $\langle S, 1\rangle$ and C_σ must be a subset of lawful worlds. Since rain causally acts on the state of the car, and the car's state is a speaker fact, the speaker's basis consists of $\langle R, 0\rangle$, since we can derive everything the speaker knows from this fact alone.

To interpret the counterfactual, one must consider maximally similar worlds where the car did not start. There are two such worlds, namely w_1 and w_2. Notice that w_2 supports the backtracking inference, while w_1 does not. Under an epistemic reading, the evaluation of the antecedent would pick out world w_2 as the most similar. Under the ontic reading, the consequent would be interpreted in world w_1. Crucially, the analysis of NC antecedents correctly predicts that w_1 is the closest. The reason is that w_1 retains more facts from the speaker's basis, per clause (56i-a) in the definition of similarity.

Specifically, interpreting the antecedent $Remove_M(C_\sigma, {\downarrow}S)$ returns the set of worlds that are closest to w_3 with respect to $\leq^\sigma$ that do not contain situations exemplifying S, namely $\langle S, 1\rangle$. The worlds where this holds are ranked as in (90).

(90) Similarity: $w_1 <^\sigma_{w_3} w_2$

The derivation in (91) shows that the backtracking inference cannot go through with NCs.

(91) If not for the car starting, it would have to have rained$^{M,C_\sigma}$ = 1 iff

a. $M, C_\sigma \vDash {\downarrow}S \succ R$ iff
b. $Remove_M(C_\sigma, {\downarrow}S) \vDash R$ iff
c. $\bigcup_{w \in C_\sigma} Min(\leq_w, \{w' \mid \forall s(s \in {\downarrow}\phi^M \rightarrow s \not\subseteq w')\}) \vDash R$ iff
d. $\bigcup \{w_1\} \vDash P$ iff
e. ✗ $w_1 \vDash R$

The analysis not only predicts that backtracking inference is necessarily unavailable, it also predicts that the only true reading is the strange case where the car starting causally antecedes it raining. Notice that switching the causal relationship between R and S, as in (92), generates the new bases in (93).

(92) S causally antecedes R

$$f_R(S): \begin{cases} 1 \rightarrow 0 \\ 0 \rightarrow 1 \end{cases}$$

(93)

	R	*S*
w_1	0	0
w_2	1	0
w_3	0	1
w_4	1	1

Using an NC still presupposes that the car started. The difference is that with this causal structure, the speaker's basis is now the fact that $\langle S, 1\rangle$, because this fact permits all the speaker's facts to be derived. While only a minor change, it makes all the difference for the interpretation of the revision function contributed by NC antecedents. Now the minimally different world without the fact $\downarrow S$ is w_2.

(94) Similarity: $w_2 <^{\sigma}_{w_3} w_1$

The reason is that both w_1 and w_2 are equally bad with respect to preserving the speaker's basis, but w_2 is more similar than w_1 to the actual world because its basis is smaller, that is, more of the causal laws are preserved (see (56i-b) for the formal calculation). The inference now goes through.

(95) If not for the car starting, it would have to have rained^{M,w_3} = 1 iff

a. $M, C_\sigma \vDash \downarrow S \succ R$ iff
b. $Remove_M(C_\sigma, \downarrow S) \vDash R$ iff
c. $\bigcup_{w \in C_\sigma} Min(\leq_w, \{w' \mid \forall s(s \in \downarrow \phi^M \rightarrow s \not\subseteq w')\}) \vDash R$ iff
d. $\bigcup \{w_2\} \vDash P$ iff
e. ✓ $w_2 \vDash R$

The prediction is that NCs in backtracking context can be true, only if the context is restructured so that the antecedent causally antecedes the consequent, eliminating the backtracking. The NC will be felicitous in so much as such this causal structure is plausible. The prediction is born out, and (95) is inflecitous because

it can only be true if the state of the car's engine can causally affect the weather against time's arrow.

To sum up, the crucial fact is that standard counterfactuals allow the backtracking reading, while NCs do not. This is predicted if NCs presuppose a fact that is then retracted within a causal model. The reason is that the speaker's basis will necessarily contain those facts that causally antecede the fact the antecedent presupposes. Since basis facts take precedence when computing similarity, counterfactual inferences about them will always be false.

Having examined a case of backtracking, now consider a slightly more complex example where the counterfactual concerns correlated variables. Correlations present a clear example of the epistemic-ontic distinction because, as opposed to backtracking, it is easy to construct natural arguments about how things would be different if one of the correlated variables were different. The ontic counterfactual leaves the correlated variable in the consequent unchanged. This is given in (96a), where the storm still happens and everybody drowns. Notice that both NCs and standard counterfactuals can be used to make this argument. This presents a sharp contrast with the epistemic rebuttal in (96b); only the standard counterfactual is felicitous here. The only reading the NC has is where the alarm actually *caused* there to be storm, which is not the case.

(96) Suppose an alarm sounds at the docks whenever there is an impending storm. Further suppose we took the bridge instead of the ferry because we heard the alarm and there was, in fact, a storm.

  a. i. Thank goodness, if the alarm hadn't gone off, we would have taken the ferry and we might have all drowned in the storm.

  ii. Thank goodness, if not for the alarm going off, we would have taken the ferry and we might have all drowned in the storm.

  b. i. No no no, that alarm always works. If it hadn't gone off, there would have to have been no storm coming.

  ii. #No no no, that alarm always works. If not for it going off, there would have to have been no storm coming.

In this scenario, the alarm A and the storm S are correlated, let's say through low pressure LP, which causally antecedes both. Thus, both A and S will be associated with a function, determining their value as a function of LP, as in (97).

(97) LP causally antecedes S and LP causally antecedes A

$$f_S(LP) : \begin{cases} 1 \rightarrow 1 \\ 0 \rightarrow 0 \end{cases}$$

$$f_A(LP): \begin{cases} 1 \rightarrow 1 \\ 0 \rightarrow 0 \end{cases}$$

Supposing that these are the only variables and laws, the bases are as in (98). In uttering a NC counterfactual about the state of the alarm, the speaker presupposes that the alarm, in fact, went off. Since the speaker's epistemic state must be lawful, C_σ must contain the single world w_1, where there was low pressure that caused both the alarm to go off and a storm. The minimal fact, that, along with the causal laws, derives everything the speakers knows is the single fact $\langle LP, 1\rangle$.

(98)

	LP	*A*	*S*
w_1	1	1	1
w_2	1	1	0
w_3	1	0	1
w_4	0	1	1
w_5	1	0	0
w_6	0	1	0
w_7	0	0	1
w_8	0	0	0

Now consider which are closest worlds where the alarm did not go off. Under an epistemic reading, world w_8 is closest; learning that the alarm went off allows one to reasonably infer that the pressure is also not low and so there is no storm. But also note that such an interpretation requires changing many facts about the world, including more facts from the basis. Under the ontic reading, these facts are left alone and w_3 is closest. Now consider the prediction made the analysis of NC antecedents.

Revising w_1, the only world in C_σ, by removing the fact that the alarm went off ($\downarrow A$) yields a set of worlds ordered by similarity as in (99). World w_3 is closest because it overlaps most with the basis of a world in the speaker's state, as clause (56i-a) requires.

(99) Similarity: $w_3 <^\sigma_{w_1} w_5 <^\sigma_{w_1} w_8 <^\sigma_{w_1} w_7$

Since w_3 is the closest world, the NC is predicted to be necessarily false, since the storm remains at that world.

(100) If not for the alarm going off, there would have to have been no storm coming$^{M,C_\sigma}$ = 1 iff

a. $Remove_M(C_\sigma, \downarrow A) \vDash \neg S$ iff

b. ✗ $w_3 \vDash \neg S$

To summarize, the analysis of NCs in (100) correctly predicts the data as presented in (96). The first inference (96a) goes through, since as shown, local update with an NC leaves the storm untouched. The epistemic reading is unavailable in (96b) for the same reason. Once again, the only way to get the purported epistemic counterfactual to be true is to change the causal laws so that they link the antecedent and consequent. This is why (96b-ii) only has a true reading where the alarm is causally implicated in the appearance of the storm.

There is now a complete account of the semantic facts that motivated a split between standard counterfactuals and NCs. We saw that NCs, unlike standard counterfactuals, cannot be used to make inferences against the flow of causality. For instance, NCs do not make good backtracking counterfactuals, requiring an inference from effect to cause, nor can they be used to make inferences between correlated variables, which requires changing their shared cause. To account for this, we showed how NC antecedents, in opposition to standard counterfactuals, presuppose a fact that fixes a subset of the basis in each world in C_σ. When calculating the result of counterfactually supposing the antecedent, it is better to momentarily suspend a particular law than change facts in the speaker's epistemic basis.

We now consider two more correct predictions that the analysis makes concerning causality and the interpretation of NCs. The first involves concessives, which require the denial of a necessary casual relationship between the antecedent and consequent. The second involves analytic inferences which make use of no causal relationships. What is shown is that NCs behave exactly as predicted given the analysis just presented.

Concessive counterfactuals explicitly deny that the antecedent is a necessary condition for the consequent. For instance, example (101) requires that John studying not be a necessary condition for him passing, since he would have passed had he not studied.

(101) Even if John hadn't studied, he would have passed the test.

At first pass, it seems like NCs require their antecedents to be a necessary condition for their consequents because they sound odd in concessives out of the blue (102). This would be stronger than what we saw previously, where NCs are ungrammatical in epistemic readings where their antecedents do not causally antecede their consequents.

(102) ?Even if it weren't for John studying, he would have passed the test.

A consequence of the analysis of NCs is that the oddness of (102) can be derived from a model that incorporates causal laws, like the one just presented. It does not have to be attributed to a stronger condition. First, consider the fact that (102) is consistent with a few different situations concerning the connection between the antecedent and consequent. The antecedent could be sufficient for the consequent, but in a context where there are other sufficient causes at play, or it could be causally unconnected to its consequent. Concessive NCs become much better when the context is enriched so that the former option holds. Consider the following naturally occurring concessive NCs. They all occur accompanied by the other sufficient causes that mask the effect of removing a fact contained in the NC antecedent.

(103) a. Giuliani would have problems with conservatives even if it weren't for his position on abortion because he is twice-divorced and was a supporter of gun control.[22]

b. Even if it weren't for his shaved head and long scar, Schoenburg would be easy to recognize; his tall frame and honest face stand out anywhere.[23]

c. Her life would be impressive even if not for all this. A single parent, she spends three hours a day in the gym, before picking up her son Luis, 4, from school, writing essays or driving to Oxford Brookes to attend lectures.[24]

The natural proposal is that since NCs reject non-causal epistemic readings, concessives are marked unless it is made clear in context that the antecedent truly does causally antecede the consequent. Earlier data showed that when epistemic readings fail, a grammatical ontic reading can be rescued if the causal structure is modified so that the antecedent can be construed as causally anteceding the consequent. Concessives are good precisely when the context makes it clear that such a relation holds, namely that the NC antecedent is only one of many variables that causally govern the consequent. Finally, it is important to note that the fact that there are NC concessives resolves a descriptive puzzle that arose when detailing the properties of NCs. It was established that NC antecedents are non-defeasibly

[22] http://pewforum.org/news/display.php?NewsID=14307

[23] http://www.yaledailynews.com/magazine/magazine-cover/2007/11/02/renting-your-mind

[24] http://women.timesonline.co.uk/tol/life_and_style/women/the_way_we_live/article70

counter-to-fact, but now it is clear that the same is not true for NC consequents since they are defeasible in the presence of concessive morphology. We can now say that all of the non-standard properties of NCs are localized in the antecedent, which is a positive result because only the antecedent contains non-standard morphology.

A final piece of data that supports the proposal is that analytic inferences are felicitous with NCs, even though they are not causal in character (104).

(104) a. If not for six being even, it wouldn't be divisible by two.
b. If not for six being divisible by two, it wouldn't be even.

This is surprising at first, since many of the cases that have been considered show that NCs are infelicitous when the antecedent and consequent are not causally linked. Nevertheless, the grammaticality of (104) provides strong evidence that NCs are sensitive to the causal flow of laws. The reason is that NCs only interrupt non-causal reasoning when causal laws are broken by assuming an NC antecedent. If there are no causal laws to be broken, then the inference will go through unaffected in both directions, as with analytic inferences.[25]

When modelling analytic connections, it is not right to say that they amount to causal links between variables. Instead, they are inviolable laws, and if they do not hold, each part must be stipulated as a (strange) fact in the basis of a world. Thus, due to the that being even (E) is defined as being divisible by two (D), we get the picture in (105).

(105)

	D	E
w_1	1	1
w_2	1	0
w_3	0	1
w_4	0	0

The fact that the actual world has an empty basis shows that the truth value of D is equivalent to E, in this case true, by law; the value of one is not derived from the other within the causal structure of the model. Now consider what happens when evaluating a counterfactual like: *if not for six being even, it wouldn't be divisible by two*. The speaker presupposes that $\langle E, 1\rangle$, but this fixes no basis facts

[25] Sam Cumming (p.c.) points out that the class of non-causal inferences allowed with NCs should be widened to include, not just analytic inferences, but all *in-virtue-of* relations. For instance, we can say both *If not for its being carbon, it wouldn't have this chemical structure*, as well as, *If not for its chemical structure, this wouldn't be carbon.*

because there is no causal conection between E and D. The function $Remove_M$ will return the closest worlds to w_1 that do not contain the situation $\downarrow E$. Only worlds w_2 and w_4 are in competition. Since the basis of the only world in C_σ is , all worlds are equivalent with respect to the amount their basis overlaps with the worlds in the speaker's epistemic state. Instead, similarity is decided by how much is derivable by law, that is, which world has the smallest basis. This is world w_4, and the inference goes through as shown in (107).

(106) Similarity: $w_4 <^{\sigma}_{w_1} w_2$

(107) If not for six being even, it wouldn't be divisible by two$^{M,C_\sigma}$ = 1 iff
   a. $Remove_M(C_\sigma, \downarrow E) \vDash \neg D$ iff
   b. ✓$w_4 \vDash \neg D$

The analysis correctly predicts the fact that NCs support analytic inferences, but not other epistemic inferences. This supports the path taken here where the non-standard behavior of NCs is due to the way that the evaluation of the antecedent interacts with the causal structure of the model. In those situations where laws come into play, but those laws are non-causal, NCs again begin to behave like standard counterfactuals.

Finally, this analysis of NCs allows us to draw conclusions about the distribution of epistemic readings with standard counterfactuals. In particular, not only does it explain why NCs reject inferences that require an epistemic interpretation, it explains why epistemic readings are dispreferred with standard counterfactuals. That is, NCs reject backtracking counterfactuals because they require an epistemic inference, yet this is not possible because NCs presuppose that their antecedents are false. The reason why backtracking is dispreferred with standard counterfactuals is that while they do not presuppose that their antecedents are false, they do imply it. This means that standard counterfactuals should be biased towards ontic readings because overcoming this implicature is a prerequisite for backtracking and other clearly epistemic readings. One prediction that seems to be borne out is that epistemic inference should improve when the truth of the antecedent is in doubt. For instance, Veltman (2005, p. 174) presents the following example of a counterfactual with an infelicitous epistemic reading.

> The duchess has been murdered, and you are supposed to find the murderer. At some point only the butler and the gardener are left as suspects. At this point you believe

(108) If the butler did not kill her, the gardener did.

Still, somewhat later — after you found out convincing evidence showing that the butler did it, and that the gardener had nothing to do with it — you get in a state, in which you will reject the sentence

(109) If the butler had not killed her, the gardener would have.

We agree with Veltman that the epistemic reading is hard to get here, but he thinks that it can be accessed if "reference is made to some previous epistemic state, in this example the state you were in when only two suspects were left. Thinking back one can say that if it had not been the butler, it would have been the gardener Veltman (2005, p. 174)." This is precisely the intuition our analysis captures. We can also improve the inference presented above if we never get into a state where the fact of the matter is settled, for example if you and your partner continue to disagree. Only he gardener had keys to the herbarium where the murder weapon was taken from, yet the butler's prints were all over the crime scene. It could only be the butler or gardener, but you are conviced that it was butler and she is conviced that it was the gardener. You might very well try to convice your partner as follows:

(110) The butler must have done it! If it hadn't been the butler, it would have to have been the gardener, yet we didn't find his prints at the scene of the crime.

She could retort.

(111) No, the gardener must have done it! If it hadn't been the gardener, it would have to have been the butler, yet he didn't have access to the garden shears.

The result is that we can now understand Veltman's (2005) intuition here. When the fact of the matter concerning the antecedent is up for grabs, for example, if there is an argument about the antecedent, or if we think back to a time when things were not settled, even epistemic inferences without a causal relation between antecedent and consequent become available. This is predicted if epistemic readings require overcoming an implicature that the antecedent is counter to fact, that is, if the antecedent can be interpreted against a context in which it is not settled.

To summarize, this section developed the core analysis of NCs. The eventive nominals in NCs denote a set of situations (partial interpretation functions) that

exemplify a proposition. The preposition *for* contributes a factive presupposition on its nominal complement. Negation is obligatory because, due to the factive presupposition on the gerund, the only way to move to counterfactual worlds for the evaluation of the consequent is to eliminate the facts denoted by the nominal in the antecedent. This motivated an analysis of negation as a model update function, taking a set of situations, and returning the closest worlds that do not contain said situations. When interpreted over a model enriched with causal relations, the analysis correctly predicts the class of non-causal epistemic inferences that are unavailable with NCs. Crucially, the account is rooted in the independent fact that NCs presuppose the counterfactuality of their antecedents, while standard counterfactuals only imply it. In the next section, we compare our analysis to the epistemic-ontic distinction to other approaches in light of the new data from NCs.

3.4 Comparing Approaches

The previous section showed that NCs lack epistemic readings because NCs presuppose a fact that negation removes, allowing causal laws to be broken. Our proposal is that under counterfactual revision, this fact fixes those facts that lie causally upstream, blocking epistemic inferences. We argued that these missing readings are available with standard counterfactuals because while they imply that there is a fact concerning the antecedent, this is not entailed. The result is that standard counterfactuals have both epistemic and ontic readings, while NCs are only paraphrasable with standard counterfactuals under their ontic reading.

While we tie the epistemic-ontic distinction to the existence of a fact concerning the antecedent in the speaker's basis, this is the not the only approach. Previous accounts attribute the two readings to an inherent vagueness in the content of a counterfactual's ordering source/premise set (most prominently Kratzer (1979, 1981a,b, 1989), though see Kaufmann (2005) for an account of epistemic readings of indicative conditionals), or to different ways of doing counterfactual belief revision (Schulz, 2007). In this section we consider how the data from NCs bear on the these approaches to the epistemic-ontic distinction. We show that our account stakes out area in the middleground. In particular, we attribute the epistemic-ontic distinction to a distinction in the structure of the speaker's basis, which we can think of as encoding various ordering sources/premise sets. That being said, using the data from NCs our account allows us to pinpoint the source of vagueness. In order to do so, though, we argue like Schulz (2007) that counterfactuals have an essential causal character and reckon similarity via causally defined bases. We start by examining the Schulz's (2007) approach in relation to

ours because the underlying the formalism is closer to that used here.

As discussed in the introduction, Schulz (2007) makes a distinction between local and global revision, where the former is revision with respect to individual worlds in a belief state, and the latter is revision with respect to a belief state in toto. In the former case, the result of revision does not have to be maximally consist with all beliefs, since similarity is reckoned against the particular content of a world. Beliefs in the form of laws that govern the distribution of facts across worlds can be ignored. In the case of global revision, similarity is reckoned over the entire belief state, including the laws.

To give a simple intuitive example from Katsuno & Mendelzon (1991), consider the case where an agent believes that p or q, but not both. Such a state, which is labeled in gray in (112), would contain p-worlds and q-worlds, but not worlds that are both p-worlds and q-worlds.

(112)

w_1	w_2	w_3
p,$\neg q$	$\neg p$,q	p,q

A local revision function that updates w_1 with q will map this world outside the belief state to world w_3 because it is the most similar world where q holds. Crucially, w_2 is not considered in making the revision, which along with w_1, encodes the agent's belief that p or q, but not both.

Global belief revision prevents such derived inconsistencies. The reason is that it relativizes similarity to belief states, which forces it to consider all of the possibilities encoded on a state when revising a world in that state with new information. Once again consider a belief state like that in (112). A global revision function would update the entire belief state in (112) with q, and unlike its local counterpart, it will map it onto w_2. The reason is that w_2 is more similar than world w_3 with respect to the belief state bescause it is actually in the belief state.

Schulz (2007) then argues that the antecedents of standard counterfactuals are ambiguous with respect to which type of revision they instantiate. Since local revision allows derived inconsistencies with the laws, it generates ontic readings, while global revision generates epistemic readings. Specifically, Schulz (2007) argues that counterfactuals are interpreted with respect to the same causal models used here, but similarity is calculated with respect to only the objective basis a world. Consider again the backtracking example repeated from (87-89), where in the actual world the car started and it did not rain. As usual, the objective basis each world is represented by a box.

(113) If the car hadn't started, it would have to have rained.

(114)

	R	*S*
w_1	0	0
w_2	1	0
$@w_3$	0	1
w_4	1	1

Interpreting the antecedent of the counterfactual amounts to identifying the closest worlds to w_3 where the antecedent is true. Under local revision, the closest such world is w_1 because it makes the antecedent true and changes the fewest basis facts. That being said, such a revision leads to a world that is inconsistent with respect to the causal laws, but as we have seen, derived inconsistences are the hallmark of local belief revision. The problem is that we only consider w_3 when making the revision required by the antecedent. Under global revision, though, we have to take into account that the speaker knows the laws and belief revision cannot generate a state inconsistent with them. Now the closest world satisfying the antecedent is w_2 because, although it alters basis facts, it is a lawful world. That is, it does not lie outside the space of worlds that could have been the actual world given the laws.

Note now that w_2 satisfies the counterfactual above, while w_1 does not. That is to say, epistemic inferences under local revision will always be false, yet are possible under global revision. To account for NCs, which do not permit epistemic inferences, an analysis in line with Schulz (2007) would have to say that NCs, unlike standard counterfactuals, only instantiate local revision. The problem with such an analysis is that it is stipulative unless obligatory local revision could be grounded in a independent property of NCs. The clearest candidate is that NC antecedents are non-defeasibly counterfactual, yet there is nothing inherent to local/global revision preventing an antecedent from presupposing a fact that makes it false, while at the same time performing global revision with respect to that fact. The fact that NCs are both non-defeasibly counterfactual and block epistemic inferences would be accidental. The analysis presented here is better because it unifies these two primary ways that NCs differ from standard counterfactuals. In doing so, though, we no longer maintain the distinction between global and local revision in accounting for the epistemic-ontic distinction.

What we maintain from Schulz's (2007) analysis is the idea that causal laws play a crucial role in the interpretation of counterfactuals and should be represented explicitly. For Schulz (2007), the causal laws define those variables that are harder to change under local revision, that is, those variables that lie upstream from the variable the antecedent revises. We also make use of the fact that epis-

temic inferences go against the flow of causality. For us, non-defeasibly counterfactual antecedents presuppose a fact that fixes those variables that lie causally upstream. This is what suppresses epistemic inferences, not obligatory local revision.

The second primary account of the empistemic-ontic distinction for counterfactuals places the ambiguity in the ordering source/premise set against which the antecedent is interpreted (Kratzer, 1979, 1981a,b, 1989). The idea is that counterfactual consequents are interpreted relative to antecedent worlds that are maximally similar with respect to "what is the case" in the world of evaluation. For instance, (Kratzer, 1981b), following an equivalent proposal in Kratzer (1979), argues that counterfactuals are interpreted relative to an empty modal base and a realistic ordering source, that is, some set of propositions, each of which the world under consideration is a member of. Just like "what is the case" is a vague notion, the membership of the ordering source is also vague. It's this vagueness that allows for the counterfactuals to support a variety of inferences.

In particular, the availability of epistemic readings like backtracking is attributed to whether the law connecting the two variables is added to the ordering source/premise set. Kratzer (1989) makes this explicit in her treatment of *lumping* and counterfactual inference in a premise semantics framework.[26] Her idea is that non-accidental generalizations are true in either every situation of a world or none, meaning they will be *lumped* by any proposition true in a world. Thus, when we build our premise set, if we add any proposition at all, we have to add this generalization. For instance, suppose we know that P and Q, and have the non-accidental generalization that whenever P happens, Q happens. When we build our premise set we have to add $\neg Q$, but since this proposition will lump our generalization, we have to add the law $\neg P \vee Q$. At this point we cannot add P to our premise set because $\{\neg Q, \neg P \vee Q, P\}$ is inconsistent. Moreover, the set of premises we are left with, namely $\{\neg Q, \neg P \vee Q\}$, entails $\neg P$. The result is that given a non-accidental generalization corresponding to our causal laws, we can make a backtracking inference from $\neg Q$ to $\neg P$. Without the non-accidental generalization, the inference will not go through. What is left vague is whether various generalizations that could be made are accidental or not. The answer to this question determines whether a particular counterfactual in a particular context will have an epistemic reading or not.

[26] *Lumping* is defined as follows: For all propositions p and $q \in P(S)$ and all $w \in W$: p lumps q in w if and only if the following conditions hold: (i) $w \in p$ (ii) For all $s \in S$, if $s \leq p$, then $s \leq q$.

This type of analysis of epistemic readings suffers from the same probem as Schulz's (2007) account. NCs show a clear connection between non-defeasibly counterfactual antecedents and a lack of epistemic readings, yet the defeasiblity of the antecedent is completely unrelated to whether or not a particular generalization is taken to be non-accidental. Given what NCs show, an analysis that is based on vagueness in the strength of a lawlike generalization misses an important generalization about the epistemic-ontic distinction. That being said, there is a clear connection between the analysis developed here and those that attribute the epistemic-ontic distinction to vagueness concerning the composition of the ordering source/premise set. We have proposed that standard counterfactuals are vague as to whether there is a settled fact concerning the antecedent, while NCs are not, and that this accounts for why NCs do not have epistemic readings. In effect, the data from NCs allow us to see exactly where the vagueness lies and how it has an effect on the composition of the ordering source/premise set.

The analysis of counterfactual ambiguity that is closest in spirit to the one developed here is Kaufmann's (2005) account of the distinction between what he calls *predictive* and *non-predictive* readings of indicative conditionals, which is itself closely connect to the account of epistemic and metaphysical readings of necessity and possibility modals presented in (Condoravdi, 2002).

(115) Kaufmann (2005, ex. 1)

a. If he submits his paper to a journal, we won't include it in our book. *predictive,*

b. If he submitted his paper to a journal, we won't include it in our book. *non-predictive*

(116) Condoravdi (2002, ex. 41)

a. He may get the flu. *epistemic/metaphysical*

b. He may have the flu. *epistemic only*

For example, the predictive counterfactual in (115a) expresses an inference about the how the world will have to be given some background uncertainty about how it unfolds with respect to the antecedent. That is, the speaker cannot say whether the paper will be included because the facts concerning the antecedent have not yet happened. In contrast, the non-predictive conditional in (115b) has only an epistemic reading. That is, the speaker cannot say whether the paper will be included because the truth of the antecedent is unknown, even if there is a fact of the matter. Similarly, the contrast between the interpretation of the future modal and the present modal in (116a-116b) reduces to whether there is a fact of the matter

concerning the propostion the modal operates on. If there is not, as in (116a), then the modal can express metaphysical possibility, that is, the possibility that the world unfolds in a certain way, given that it could evolve in many different ways. In contrast, (116b) only has an epistemic interpretation where it indicates the speaker's uncertainty about proposition given their belief state, even if the fact of the matter is settled. The speaker could learn more and come to conclude that he must have the flu, yet by the time (116b) is uttered, he has the flu or not and nothing can change that.

Kaufmann (2005) and Condoravdi (2002) account for the availability of predictive and metaphysical readings respectively within world-time modals (Thomason, 1984), where the past is fixed and the future is indeterminant, or *future-branching*. The core idea is that modals and conditionals will be ambiguous as to whether they use an epistemic or metaphysical modal base, yet the latter will be unavailable if the tense of the proposition the modal operates on requires that it be settled history, that is, if there are no open metaphysical alternatives. In these approaches, the march of time fixes facts, which in turn restricts the available readings for modals and indicative conditionals. In the analysis of NCs developed here, the temporal properties of the antecedent do not fix certain propositions as settled, but a factive presupposition does. Just as in the accounts of Kaufmann (2005) and Condoravdi (2002), the settled proposition then restricts the available readings of the relavant construction, here NCs. While the contribution of tense in counterfactuals is hotly debated, we think our account is on the right track in that, unlike Kaufmann (2005) and Condoravdi (2002), the temporal properties of the antecedent do not play a role in determining the availability of various readings. For instance, in contrast to examples (115-116), there are no clear tense difference between epistemic and ontic readings of standard counterfactuals. Moreover, there is not even overt tense morphology in the NC antecedents we have considered most closely here.

To summarize, we considered couple of different apporaches to the epistemic-ontic distinction for counterfactuals. We argued that the analysis developed by Schulz (2007), which specifies two different types of counterfactual update, does not capture the generalization that non-defeasible counterfactuality is correlated with a lack of epistemic readings in NCs. We saw that certain approaches to the epistemic-ontic distinction, for example, the *lumping* approach developed by Kratzer (1989), suffers from the same problem. Here, the strength of the generalization supporting the epistemic inference determines whether it must be added to the modal base/ordering sorces, which is a problem because NCs show that it is the fact concerning the antecedent that matters. That being said, the data from

NCs is not an argument against all accounts in terms of different modal basis. In particular, we saw that Kaufmann (2005) and Condoravdi (2002) draw a distinction between epistemic and metaphysical modal bases that pull apart in temporal models when the speaker does not have information about a proposition that is settled at the time of utterance. This is very similar to our idea that we need to make a distinction between the objective basis of a world an the speaker's basis of a world, which can pull apart when the counterfactual antecedent implied to be false, but not presupposed. While Kaufmann (2005) and Condoravdi (2002) are not immediately interested in capturing the epistemic-ontic distinction in counterfactual conditionals, their analyses are the closest analogues to the account of NCs we develop, and provide supporting evidence that settledness of a proposition can restrict the available readings for modalized expressions.

4 CONCLUSIONS

In this work we have shown that there exists a species of counterfactual in English, namely the NC, which systematically differs from standard counterfactuals in resisting non-causal epistemic inferences. This led us to propose a new account of the epistemic-ontic distinction, though rooted in the work of Schulz (2007) on casual approaches to counterfactuals, that could account for why NCs only have one reading. Specifically, we argued that epistemic readings arise because standard counterfactuals only implicate that their antecedents are false, while NC antecedents presuppose a fact that negation removes by minimally altering the world. First, since NC antecedents carry a factive presupposition, we can account for the fact that their antecedents are non-defeasibly counterfactual. Second, we capture the fact that NCs have only a subset of the readings available with standard counterfactuals by showing how the antecedent's presupposition fixes the facts upstream in a causal model (Schulz, 2007). When there is an antecedent that is counter-to-fact, and an inference that does require changing facts against the flow of causality, NCs and standard counterfactuals will be mutually paraphrasable, which is exactly what the data show.

Finally, while this paper only considers English NCs, as noted in the text, there are similar constructions in other languages. That is, non-canonical counterfactual constructions with obligatory negation and non-defeasibly counterfactual antecedents, like Mandarin (Ippolito & Su, to appear), Tagalog (Nevins, 2002), Spanish, and Kaqchikel. What is not known is whether they behave as NCs with respect to the available inferences, though we would predict that they do. While

studying NCs across languages is clearly important for understanding the variety of counterfactual updates, it would also be interesting to see if the analysis can be extended to other non-canonical counterfactual constructions. A prime example is the counterfactual reading of *without* adjuncts. First, note that counterfactual *without*, like NCs, does not license NPIs, suggesting that it, too, might instantiate fact removal through constituent negation (118a).

(117) I was able to buy the car without borrowing a red cent.

(118) a. *I wouldn't have been able buy the car without borrowing a red cent.
b. I wouldn't have been able to buy the car without borrowing money.

Once again, just like NCs, epistemic inferences like backtracking with counterfactual *without* are odd (119).

(119) Suppose that the car works perfectly except for after a rain, which always causes the wiring to short out. Suppose it didn't rain and the car started as usual.
a. Without the clear skies, the car wouldn't have started.
b. #Without the car starting, it would have to have rained.

At first pass it seems like counterfactual *without* behaves exactly as predicted if it were another operator that locally remove a fact, and therefore amenable to an analysis like the one developed here for NCs.

The data from *without* only underline a point made throughout this work, namely that natural language counterfactuals are not monolithic. Standard counterfactuals have various readings, and there are morphologically distinct counterfactuals that support only restricted classes of readings. The fact that non-standard counterfactuals have morphologically distinct antecedents provides strong evidence that morphology can constrain the evaluation potential of the antecedent, which in turns affects the class of available inferences. In particular, this work has argued that the epistemic-ontic distinction should be accounted for as a split in whether there is a fact of the matter concerning the antecedent proposition. To support this conclusion, it was shown that NCs suppress non-analytic epistemic inferences because their morphology restricts the interpretation of their antecedents so that they presuppose a fact that must be removed via negation. In this way, the analysis not only provides an account of NCs, but also presents the first piece in a larger typology of counterfactual constructions and the update potential of their antecedents.

Acknowledgements: First above all I need to thank Donka Farkas for spending many hours talking with me about these counterfactuals in English and other languages. I also want to thank Pranav Anand, Scott AnderBois, Ryan Bennett, Adrian Brasoveanu, Daniel Büring, Ivano Caponigro, Sandy Chung, Lily Clark, Cleo Condoravdi, Sam Cumming, Amy Rose Deal, Olga Kagan, Ed Keenan, Sven Lauer, Chris Potts, Kyle Rawlins, and Luis Vicente for helpful comments on this work. I also need to thank audiences at UC Santa Cruz S(yn/em)-Circle (2008), CUSP I at UCLA (2009), and the N3C Workshop at Stanford (2009) for their deep and lively commentary about the data and analysis presented here. Luis Vicente deserves credit for the Spanish data, and Magda Sotz Mux helped with the Kaqchikel facts. That being said, all remaining mistakes are my own.

References

de Almeida, R.G. 2004. The effect of context on the processing of type-shifting verbs. *Brain and Language* 90(1-3). 249–261.

Anderson, AR. 1957. A theory of subjunctive conditionals. *The Journal of Symbolic Logic* 22(3).

Arregui, Ana. 2004. *On the accessibility of possible worlds*: PhD dissertation, UMass Amherst dissertation.

Asher, Nicholas. 1993. *Reference to abstract objects in English*. Kluwer Academic Press.

Chierchia, Gennaro. 1984. *Topics in the syntax and semantics of infinitives and gerunds*: University of Massachusetts, Amherst dissertation.

Condoravdi, C. 2002. Temporal interpretation of modals: Modals for the present and for the past. *The Construction of Meaning* 59–88.

Galles, D. & J. Pearl. 1998. An axiomatic characterization of causal counterfactuals. *Foundations of Science* 3(1). 151–182.

Goldberger, A.S. 1972. Structural equation methods in the social sciences. *Econometrica: Journal of the Econometric Society* 979–1001.

Goodman, Nelson. 1947. The problem of counterfactual conditionals. *Journal of Philosophy* 44. 113–28.

Halpern, J.Y. & J. Pearl. 2005a. Causes and explanations: A structural-model approach. Part I: Causes. *British Journal for the Philosophy of Science* 56(4). 843–888.

Halpern, J.Y. & J. Pearl. 2005b. Causes and explanations: A structural-model approach. Part II: Explanations. *The British Journal for the Philosophy of Science* 56(4). 889.

Hansson, Sven O. 1989. New operators for theory change. *Theoria* 55(2). 114–132.

Hulstijn, Joris. 1997. Structured information states. Raising and resolving issues. In *Proceedings of mundial97, university of munich*, .

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31(2). 231–270.

Ippolito, Michela & Julia Su. to appear. Counterfactuals, negation, and polarity. In *Proceedings of nels*, .

Isaacs, James & Kyle Rawlins. 2008. Conditional questions. *Journal of Semantics* 25(3). 269.

Katsuno, H. & A. Mendelzon. 1991. On the difference between updating a knowledge base and revising it. *KR* 387–394.

Kaufmann, S. 2005. Conditional truth and future reference. *Journal of Semantics* 22(3). 231.

Kratzer, Angelika. 1979. Conditional necessity and possibility. *Semantics from Different Points of View: With 15 Figures* 117.

Kratzer, Angelika. 1981a. *The notional category of modality* 38–74. Berlin: Walter de Gruyter.

Kratzer, Angelika. 1981b. Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic* 2. 201–216.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12(5). 607–653.

Kratzer, Angelika. 2002. Facts: Particulars or information units? *Linguistics and Philosophy* 25(5). 655–670.

Lewis, David. 1973. *Counterfactuals*. Harvard University Press.

Lifschitz, V. 1987. Formal theories of action (preliminary report). In *Proc. of ijcai*, vol. 87, 966–972.

McElree, B., M.J. Traxler, M.J. Pickering, R.E. Seely & R. Jackendoff. 2001. Reading time evidence for enriched composition. *Cognition* 78(1). 17–25.

Morreau, Michael. 1992. Epistemic semantics for counterfactuals. *Journal of Philosophical Logic* 21(1). 33–62.

Nevins, Andrew Ira. 2002. Counterfactuality without past tense. In Masako Hirotani (ed.), *Proceedings of north east linguistic society 32*, vol. 2, 441–450.

Pearl, Judea. 2000. *Causality: Models, reasoning, and inference*. Cambridge UP.

Pickering, M.J., B. McElree & M.J. Traxler. 2005. The difficulty of coercion: A response to de Almeida. *Brain and Language* 93(1). 1–9.

Portner, Paul. 1992. *Situation theory and the semantics of propositional expressions*: University of Massachusetts-Amherst dissertation.

Ramsey, Frank P. 1950. Foundations of Mathematics and other Logical Essays, chapter General Propositions and Causality. *New York* 237–257.

Rescher, Nicholas. 1964. *Hypothetical reasoning*. North-Holland Amsterdam.

Rott, Hans. 1999. Moody conditionals: Hamburgers, switches, and the tragic death of an American president. *Gerbrandy, Marx, de Rijke, and Venema, editors, JFAK. Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday, Vossiuspers AUP, Amsterdam* .

Schulz, Katrin. 2007. Minimal models in semantics and pragmatics: Free choice, exhaustivity, and conditionals. *ILLC Dissertation Series* .

Schwarz, Bernhard & Rajesh Bhatt. 2006. Light negation and polarity. *Crosslinguistic research in syntax and semantics: negation, tense, and clausal architecture* 175.

Stalnaker, Robert. 1968. A theory of conditionals. *Studies in Logical Theory* .

Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5(3). 269–286.

Thomason, R.H. 1984. Combinations of tense and modality. *Handbook of philosophical logic* 2. 135–165.

Traxler, M.J., M.J. Pickering & B. McElree. 2002. Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language* 47(4). 530–547.

Velissaratou, Sophia. 2000. Conditional questions and which-interrogatives. *MSc in Logic thesis, ILLC, University of Amsterdam* .

Veltman, Frank. 1976. Prejudices, presuppositions and the theory of conditionals. In *Proceedings of the first Amsterdam colloquium: Amsterdam papers in formal grammar, vol. 1*, 248–281.

Veltman, Frank. 2005. Making counterfactual assumptions. *Journal of Semantics* 22(2). 159.