# How Strong is a Counterfactual?

[Authors redacted for anonymous review]*

## 1 Introduction

There are two leading theories about the meaning of counterfactuals like (1):

(1) If David's alarm hadn't gone off this morning, he would have missed class.

Both say that (1) means, roughly, that David misses class in all of the closest worlds where his alarm doesn't go off. They disagree about how this set of worlds—the *counterfactual domain*—is determined. The *Variably Strict Analysis* (VSA) says that the domain is a function of the antecedent of the counterfactual; it *varies* from antecedent to antecedent. The *Strict Analysis* (SA) says that the counterfactual domain is not a function of the antecedent of the counterfactual; it does not vary from antecedent to antecedent.[1]

VSA and SA validate different inference patterns. Perhaps most famously, they disagree about a principle known as *Antecedent Strengtheing*. SA validates the principle; VSA does not. Early VSA theorists, such as Stalnaker (1968) and Lewis (1973), believed that certain apparent counterexamples to Antecedent Strengthening—now known as *Sobel Sequences*—refuted SA. More recently, defenders of SA have responded by enriching SA with certain *dynamic* principles governing how context evolves. They argue that Sobel sequences are not counterexamples to a *Dynamic Strict Analysis* (Dynamic SA).

But Antecedent Strengthening is just one of a family of strengthening principles. In this paper, we focus on a weaker principle, which we'll call *Strengthening*

*Thanks to [redacted]. The authors made equal contributions to the paper.

[1]For defenses of VSA, see Stalnaker (1968), Lewis (1973), Kratzer (1981a), Kratzer (1981b), Moss (2012), and Lewis (2017). For defenses of SA, see von Fintel (2001) and Gillies (2007).

*with a Possibility*, and give a counterexample to it. We show that the dynamic features attributed to *would*-counterfactuals and *might*-counterfactuals by proponents Dynamic SA are of no help when it comes to counterexamples to Strengthening with a Possibility, unlike counterexamples to Antecedent Strengthening itself.

We develop a new version of VSA that incorporates a Kratzerian *ordering source* into the meaning of counterfactuals, and we show how to model counterexamples to Strengthening with a Possibility on our account. In the final section, we address a worry about our account concerning counterfactuals with disjunctive antecedents.

## 2 Two Theories of Counterfactuals

Both VSA and SA assume that context supplies a comparative closeness ordering on worlds, represented by $\preceq_{c,w}$, which compares any two worlds with respect to their similarity to a benchmark world $w$.[2] We assume that $\preceq_{c,w}$ is a comparative *similarity* relation, so that $w_1 \preceq_{c,w} w_2$ means that $w_1$ is at least as similar to $w$ as $w_2$ is.

VSA uses a *selection function*, a contextually-determined function $f_{\preceq_c}$ that takes an antecedent $A$, and a world $w$, and returns the set of closest $A$-worlds to $w$, according to $\preceq_{c,w}$. A world $w_1$ is among the closest $A$-worlds to $w$ just if there's no other $A$-world $w_2$ that's closer to $w$ than $w_1$ is. Where $A \mathrel{\Box\!\!\rightarrow} C$ stands for the counterfactual conditional with antecedent A and consequent C, VSA's semantic entry for the counterfactual runs as follows:

(2) **Variably Strict Account (VSA)**. $[\![A \mathrel{\Box\!\!\rightarrow} C]\!]^{c,w} = 1$ iff $\forall w' \in f_{\preceq_c}(A, w)$ : $[\![C]\!]^{c,w'} = 1$.

$A \mathrel{\Box\!\!\rightarrow} C$ is true relative to a context $c$ and world $w$ just if, for every world $w'$: if $w'$ is among the closest $A$-worlds to $w$, then $C$ is true in $w'$.[3]

[2] $\preceq_{c,w}$ is transitive (for all $w_1, w_2, w_3$ if $w_1 \preceq_{c,w} w_2$, and $w_2 \preceq_{c,w} w_3$, then $w_1 \preceq_{c,w} w_3$), reflexive (for all $w_1$, $w_1 \preceq_{c,w} w_1$), and antisymmetric (for all $w_1, w_2$, if $w_1 \preceq_{c,w} w_2$ and $w_2 \preceq_{c,w} w_1$, then $w_1 = w_2$). It is also at least *weakly centered*: there is no $w_1$ such that $w_1 \prec_{c,w} w$.

[3] This statement of VSA makes the *limit assumption*, the assumption that, for all worlds $w$ and propositions $A$, there is at least one closest $A$-world to $w$. This is purely for ease of presentation; VSA can be stated without making the limit assumption, but it's more of a mouthful. Our statement

SA replaces the selection function with a contextually-determined accessibility relation $min_{\leq_c}$ that takes a world $w$ and returns the set of closest worlds to $w$, according to $\leq_{c,w}$. (Unlike the selection function $f_{\leq_c}$, $min_{\leq_c}$ does not take the antecedent as argument.) Here's the semantic entry given by SA:

(3) **Strict Conditional Account (SA)**. $[\![A \mathrel{\Box\!\!\rightarrow} C]\!]^{c,w} = 1$ iff $\forall w' \in min_{\leq_c}(w) \cap [\![A]\!]^c$: $[\![C]\!]^{c,w'} = 1$.

$A \mathrel{\Box\!\!\rightarrow} C$ is true relative to a context $c$ and world $w$ just if, for every world $w'$, if $w'$ is among the closest worlds to $w$, and $A$ is true in $w'$, then $C$ is true in $w'$.

The difference between SA and VSA is significant—whether the evaluation domain is variable or fixed matters to the *logic* of the counterfactual. VSA and SA validate different inference patterns. Here is one principle they disagree about:

(4) **Antecedent Strengthening**. $A \mathrel{\Box\!\!\rightarrow} C \vDash (A \wedge B) \mathrel{\Box\!\!\rightarrow} C$

SA validates Antecedent Strengthening, whereas VSA does not. It's not hard to see why. SA doesn't allow the evaluation domain to vary. If all the closest worlds where $A$ is true are worlds where $C$ is true, then *a fortiori*, all of the closest worlds where *$A$ and $B$* are true are worlds where $C$ is true. On the other hand, VSA allows the evaluation domain to vary from antecedent to antecedent: The closest $AB$-worlds need not be the closest $A$-worlds. So, what's true in all of the closest $A$-worlds may be false in some of the closest $AB$-worlds.

## 3 Sobel Sequences and Dynamic SA

Antecedent Strengthening seems subject to counterexample. Consider the sequence in (5) below, a *Sobel sequence*:

(5) a. If I had struck the match, it would have lit.

of VSA is neutral about the *uniqueness assumption*, the assumption that for all $w$ and $A$, there is a unique closest $A$-world to $w$. See Stalnaker (1968) and Stalnaker (1981) for defenses of both the limit assumption and the uniqueness assumption. See Lewis (1973) for arguments against both principles.

b. But of course, if I had struck the match *and* it had been soaked in water last night, it wouldn't have lit.

Sentences (5-a) and (5-b) seem consistent. Indeed, in most ordinary match-striking scenarios, (5-a) and (5-b) are both true. Typically, when you strike a match, the match is dry and lights easily. Soaking-wet matches aren't as cooperative—if you strike a match drenched in water, it doesn't light.

But if Antecedent Strengthening is valid, Sobel sequences like (5) are not consistent. If it's true that if I'd struck the match, it would have lit, then, by Antecedent Strengthening, it follows that if I'd struck the match *and* the match had been soaked in water, the match would (still) have lit.

That it validates Antecedent Strengthening would seem to be a clear strike against SA. But things aren't quite so simple. As von Fintel (2001) shows, a suitably sophisticated strict conditional analysis *can* account for the sequence in (5). His strategy is to appeal to the dynamic effects of counterfactuals in conversation: Though (5-b) isn't true when (5-a) is uttered, *asserting* (5-b) changes the context so that it comes out true.

Here's how that works. Counterfactuals presuppose that their domains contain some worlds where the antecedent is true. When that presupposition is not met, the context is minimally altered to ensure that it is. Suppose a speaker utters a counterfactual $A \mathrel{\Box\!\!\rightarrow} C$. If the domain contains $A$-worlds, nothing changes; if it doesn't, it *expands* to include the closest $A$-worlds. $A \mathrel{\Box\!\!\rightarrow} C$ is true in this *new* context just in case all of the $A$-worlds in the expanded set are $C$-worlds.

Let's apply von Fintel's *dynamic strict conditional account* (Dynamic SA) to our example. When the speaker asserts (5-a), the domain expands to include the closest worlds where she strikes the match. (5-a) is true. So in all of these worlds, the match lights. But any world where the match lights is one where the match is *dry*. So the presupposition of (5-b) isn't satisfied. When the speaker *utters* (5-b), the context changes to accommodate the presupposition—the domain expands to include the closest worlds where she strikes the match and the match is soaking wet. Since all of these worlds are ones where the match doesn't light, (5-b) comes out true.

Now that we have an informal understanding of how Dynamic SA works, it's time to state the view more precisely.[4] Let $h_{\leq_c}$ a contextually-supplied accessibility relation. Then:

(6) **Dynamic Strict Conditional Account (Dynamic SA)**.

a. $[\![A \boxright C]\!]^{c,w}$ is defined only if $h_{\leq_c}(w) \cap [\![A]\!]^c \neq \emptyset$

b. Where $[\![A \boxright C]\!]^{c,w}$ is defined, $[\![A \boxright C]\!]^{c,w} = 1$ iff $\forall w' \in h_{\leq_c}(w) \cap [\![A]\!]^c$: $[\![C]\!]^{c,w'} = 1$.

There are two parts to Dynamic SA: *definedness conditions* and *truth-conditions*. The definedness conditions, which are stated in (6-a), take the form of a semantic presupposition about the contextually-supplied accessibility relation. The presupposition is a *compatibility presupposition*: the accessibility relation $h_{\leq_c}$ must assign to the world of evaluation $w$ at least some worlds where $A$ is true. In other words, $h_{\leq_c}(w)$ must be compatible with $A$. If $h_{\leq_c}(w)$ is not compatible with $A$, $[\![A \boxright C]\!]^{c,w}$ is undefined. The truth-conditions for the counterfactual are stated in (6-b). They tell us when the counterfactual is true if it is defined. Specifically, they say that if $[\![A \boxright C]\!]^{c,w}$ is defined, then $[\![A \boxright C]\!]^{c,w}$ is true just in case the following holds: for every world $w'$, if $w'$ is among the closest worlds to $w$, and $A$ is true in $w'$, then $C$ is true in $w'$.

But we don't yet have the full story about the meaning of the counterfactual. Two questions are left open. First, what is the identity of the initial accessibility relation $h_{\leq_c}$? And second, what happens when $[\![A \boxright C]\!]^{c,w}$ is undefined—that is, when $h_{\leq_c}$ does not assign to the world of evaluation any worlds where $A$ is true? We said that when the compatibility presupposition of the counterfactual is not met, the context is minimally altered to ensure that it is. But what, exactly, does that mean?

von Fintel provides answers to both questions. He assumes a trivial initial accessibility relation, which assigns to any evaluation world $w$ the singleton set of that world itself $\{w\}$. His answer to the second question is more detailed, but the basic idea is that when a counterfactual $A \boxright C$ is undefined relative to a context

[4]Note that von Fintel adds the dynamic effects directly into the semantics, building a dynamic context-change potential into the semantic entry for the counterfactual. This version of the view is more complicated to state so we prefer to work with the formulation of Dynamic SA that puts the dynamic effects in the pragmatics. Nothing we say about Dynamic SA hangs on this choice.

$c$ and world $w$, asserting the counterfactual updates the context by adding to the counterfactual domain *all of the closest* $A$*-worlds to* $w$, according to $\leq_c$.[5] The counterfactual is true in this updated context just in case all of the $A$-worlds in the expanded domain are $C$-worlds.

It's important to note that Antecedent Strengthening is not *classically* valid on Dynamic SA. An inference is classically valid just in case its conclusion is true whenever its premises are. Dynamic SA says that Antecedent Strengthening is merely *Strawson valid*: Whenever $A \mathrel{\Box\!\!\rightarrow} C$ is true, and $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ is *defined*, $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ is true, too.[6] Dynamic SA allows contexts where $A \mathrel{\Box\!\!\rightarrow} C$ is true yet $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ is undefined. (This will happen whenever the domain contains $A$-worlds, but no $B$-worlds, and all of the $A$-worlds are $C$-worlds.) This is critical to its account of Sobel sequences. It is the fact that (5-b) is undefined, rather than simply false, that forces the context to change when (5-b) is asserted so that (5-b) comes out true.

We have seen that Sobel sequences like (5) do not refute Dynamic SA, even if they refute the simplest strict analysis. But von Fintel isn't just trying to even the score. He thinks we have reason to *prefer* Dynamic SA to VSA. He observes that when we reverse the order of (5-a) and (5-b), the sequence sounds much worse:

(7) a. If I had struck the match and it had been soaked in water last night, it wouldn't have lit.
b. But if I had struck the match, it would have lit.

VSA doesn't predict any difference between (5) and (7). As far as it's concerned, *both* sequences should sound fine. (5-a) and (5-b) are consistent; reversing the order can't change that.

Dynamic SA, on the other hand, has a persuasive account of the infelicity of (7). Find the closest worlds to actuality, according to the standards of similarity in the context. Among those worlds, identify the ones where the speaker strikes the

[5] Nichols (2017) objects to this feature of Dynamic SA and furthermore argues that any plausible way of accommodating the presupposition leads to unwelcome consequences. Our problem is separate from those he discusses, however.

[6] More generally, the inference from $[\![A]\!]^c$, $[\![P_1]\!]^c$, ... , $[\![P_n]\!]^c$ to $[\![C]\!]^c$ is Strawson-valid iff for any $w$ such that $[\![A]\!]^{c,w}$, $[\![P_1]\!]^{c,w}$, ..., $[\![P_n]\!]^{c,w}$ and $[\![C]\!]^{c,w}$ are all defined and such that $[\![A]\!]^{c,w} = [\![P_1]\!]^{c,w} = \ldots = [\![P_n]\!]^{c,w} = 1$, $[\![C]\!]^{c,w} = 1$ also.

match and it is wet. (7-a) says that all of these are worlds where the match doesn't light; (7-b) says that all of them are worlds where the match does light. It's no wonder that (7) is infelicitous—(7-a) and (7-b) contradict each other.

VSA theorists haven't conceded, of course. Moss (2012) argues that what Dynamic SA explains *semantically*, VSA can explain by appeal to warranted assertability.[7] Though semantically consistent, Moss argues, reverse Sobel sequences are pragmatically inappropriate.

We think the debate has reached something of stalemate. We have two explanations of the same data—one pragmatic, one semantic. Neither has a clear empirical edge.[8] We move to change the debate. Antecedent Strengthening is just one of a family of strengthening principles—and indeed, it is the strongest member of that family. By Strawson-validating Antecedent Strengthening, Dynamic SA predicts that a whole host of strengthening principles are Strawson-valid, too. We argue that this prediction is unwelcome. We focus on one strengthening principle—which we call *Strengthening with a Possibility*—and present a counterexample to it. Furthermore, we show that Dynamic SA *classically* validates this principle, rather than (merely) Strawson-validating it. This is significant, for it means that Dynamic SA's dynamic resources are of no help when it comes to counterexamples to Strengthening with a Possibility.

## 4 Strengthening with a Possibility

We can think of a strengthening principle as a principle that allows us to move from a counterfactual $A \mathrel{\Box\!\!\rightarrow} C$, along with certain auxiliary premises, to a counterfactual with a strengthened antecedent $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$. More formally, where $n \geq 0$, we have:

[7] Moss is not the only one who has defended VSA against reverse Sobel sequences. Lewis (2017) has also offered an alternative story within the VSA framework, according to which raising skeptical possibilities to salience renders examples like (7-b) false.

[8] Proponents of VSA might point out that VSA has an easier time predicting the felicitous reverse Sobel sequences noted by Moss (2012). But, on the other hand, proponents of SA might point out that their theory has an easier time explaining why NPIs are licensed in the antecedents of counterfactuals (as noted by von Fintel (2001)). Thus, while these theories are clearly not empirically equivalent, we think that the literature so far has not shown that one theory has a clear edge.

(8) **Strengthening Principle**. $A \mathrel{\Box\!\!\rightarrow} C, P_1, ..., P_n \vDash (A \wedge B) \mathrel{\Box\!\!\rightarrow} C$

Antecedent Strengthening is the instance of (8) where $n = 0$. It says that *no* further premises are needed to strengthen the antecedent of a counterfactual. This makes it the strongest strengthening principle: A semantics that validates it validates *every* strengthening principle. Classical validity is monotonic: *Adding* premises never turns a valid inference into an invalid one. If $A \mathrel{\Box\!\!\rightarrow} C$ entails $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ *all by itself*, $A \mathrel{\Box\!\!\rightarrow} C$ still entails $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ when combined with other premises. Similar reasoning shows that *Strawson*-validating Antecedent Strengthening Strawson-validates every other strengthening principle—Strawson-entailment is also monotonic.[9]

But there are weaker strengthening principles. These principles allow us to strengthen the antecedent of a counterfactual not with just *any* conjunct, but only those that satisfy some auxiliary premises. We're interested in one of these weaker principles, which we'll call *Strengthening with a Possibility*.

Before we state the principle, a few preliminary remarks are in order.

## 4.1 'Might' Counterfactuals

So far we've been talking about English *would*-counterfactuals like (1):

(1) If David's alarm hadn't gone off this morning, he would have missed class.

But *would*-counterfactuals are not the only kind of counterfactual. There are also *might*-counterfactuals. For example:

(9) If David's alarm hadn't gone off this morning, he might have missed class.

We need to say what *might*-counterfactuals like (9) mean.

We will write $A \mathrel{\Diamond\!\!\rightarrow} B$ for the *might*-counterfactual with antecedent $A$ and consequent $C$. We will assume that $A \mathrel{\Diamond\!\!\rightarrow} B$ is the *dual* of $A \mathrel{\Box\!\!\rightarrow} B$. What does that mean? The dual of an operator $O$ is a term which has the meaning of $\neg O \neg$. To say

[9] **Proof**: Suppose that $[\![A]\!]^c, [\![P_1]\!]^c, ..., [\![P_n]\!]^c \nvDash_{Str} [\![C]\!]^c$. There there must be some $w$ such that $[\![A]\!]^{c,w} = [\![P_1]\!]^{c,w} = ... = [\![P_n]\!]^{c,w} = 1$ but $[\![C]\!]^{c,w} = 0$. But then, since $w$ itself is a world where $[\![A]\!]^{c,w} = 1$ but $[\![C]\!]^{c,w} = 0$, we have $[\![A]\!]^c \nvDash_{Str} [\![C]\!]^c$. Contraposing, if $[\![A]\!]^c \vDash_{Str} [\![C]\!]^c$ then $[\![A]\!]^c$, $[\![P_1]\!]^c$,...$[\![P_n]\!]^c \vDash_{Str} [\![C]\!]^c$.

that $A \Diamond\!\!\rightarrow B$ is the dual of $A \Box\!\!\rightarrow B$ is to say that $A \Diamond\!\!\rightarrow B$ has the same meaning as $\neg(A \Box\!\!\rightarrow \neg B)$. For instance, (9) has the same meaning as (10):

(10) It's not the case that if David's alarm hadn't gone off this morning, he *wouldn't* have missed class.

The assumption that English *might*-counterfactuals are the duals of *would*-counterfactuals is called *Duality*.[10] Note that Duality is controversial: it's denied by many, including various defenders of Counterfactual Excluded Middle.[11] We assume Duality merely for ease of exposition. Our central counterexample can be stated without it. (See footnote 12 for further details.)

If we assume Duality, then we can state the semantics for *might*-counterfactuals given by Dynamic SA as follows:

(11) **Dynamic SA** ($A \Diamond\!\!\rightarrow C$)

a. $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w}$ is defined only if $h_{\leq_c}(w) \cap [\![A]\!]^c \neq \emptyset$

b. Where $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w}$ is defined, $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w} = 1$ iff $\exists w' \in h_{\leq_c}(w) \cap [\![A]\!]^c$: $[\![C]\!]^{c,w'} = 1$.

The semantics for $\Diamond\!\!\rightarrow$ has two parts, just as it does for $\Box\!\!\rightarrow$. First we have the definedness condition stated in (11-a): $A \Diamond\!\!\rightarrow C$ is defined only if $h_{\leq_c}$ assigns to the world of evaluation at least some worlds where $A$ is true. (This part is the same for $\Diamond\!\!\rightarrow$ as it is for $\Box\!\!\rightarrow$.) And second we have the truth-conditions, which are set down in (11-b). Where $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w}$ is defined, $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w}$ is true just in case there is some world $w'$ such that $w'$ is among the closest worlds to $w$, $A$ is true in $w'$, and $C$ is true in $w'$.

That's the semantics given by Dynamic SA. What about VSA? Assuming Duality, VSA gives $\Diamond\!\!\rightarrow$ the following semantics:

(12) **VSA** ($A \Diamond\!\!\rightarrow C$). $[\![A \Diamond\!\!\rightarrow C]\!]^{c,w}=1$ iff $\exists w' \in f_{\leq_c}(A, w) : [\![C]\!]^{c,w'}=1$.

[10] Note that von Fintel, as well as Gillies (2007), another proponent of Dynamic SA, seem to accept Duality. Moreover, Duality falls out out of widely-accepted restrictor analysis of conditionals in Kratzer (1986): on this analysis, the 'might' will only quantify over worlds that make the antecedent true and so *might*-counterfactuals will have the truth-conditions of $\Diamond\!\!\rightarrow$.

[11] See, among others, Stalnaker (1981) and Williams (2010) for arguments against Duality.

VSA says that $A \diamondto C$ is true relative to a context $c$ and world $w$ just if there's some world $w'$ that's among the closest $A$-worlds to $w$ that is also a $C$-world.

## 4.2 Strengthening with a Possibility and Dynamic SA

We're now in a position to state the strengthening principle that we're interested in.

(13) **Strengthening with a Possibility**. $(A \boxto C) \wedge (A \diamondto B) \vDash (A \wedge B) \boxto C$

(13) says that one can strengthen an antecedent with any proposition with which that antecedent is *counterfactually consistent*. Take an example. Suppose it's true that if I'd taken modal logic next semester, I would have passed. Does that mean that I would have passed had I taken the class and the class was taught by Joe? According to Strengthening with a Possibility, that depends on whether Joe *might* have been the teacher, had I taken the class. If Joe couldn't have taught the class—say, because he was on leave—I can truly say that I would have passed even if I would have bombed a class taught by Joe. On the other hand, if Joe might have taught the class, then I can't truly say that I would have passed unless I would have passed Joe's class, too.

We said that Antecedent Strengthening is the strongest strengthening principle. So, by Strawson-validating Antecedent Strengthening, Dynamic SA Strawson-validates Strengthening with a Possibility. But we can show something stronger: By Strawson-validating Antecedent Strengthening, Dynamic SA *classically* validates Strengthening with a Possibility. The proof is straightforward. Suppose that for a given context $c$, (i) $A \boxto C$ is true in $c$, and (ii) $A \diamondto B$ is true in $c$. It follows from (ii) and Dynamic SA that the domain in $c$ contains worlds where $A$ and $B$ are both true. But that's just to say (iii) that $(A \wedge B) \boxto C$ is *defined* in $c$. Since Antecedent Strengthening is Strawson-valid, (i)—that $A \boxto C$ is true in $c$—and (iii)—that $(A \wedge B) \boxto C$ is defined in $c$—together entail that $(A \wedge B) \boxto C$ is true in $c$.

This fact is important. If Strengthening with a Possibility is classically valid, we can't appeal to the dynamic resources of Dynamic SA to account for apparent counterexamples. To see this, return to the match example:

(5) a. If I had struck the match, it would have lit.
b. But of course, if I had struck the match *and* it had been soaked in water last night, it wouldn't have lit.

Earlier we noted that Dynamic SA Strawson-validates, but does not classically validate, Antecedent Strengthening. Dynamic SA does not say that (5-b) is true in any context in which (5-a) is true—it allows contexts where (5-a) is true yet (5-b) is *undefined*. That's how it accounts for the felicity of Sobel sequences like (5). If (5-b) is undefined when (5-a) is uttered, asserting (5-b) shifts the context so that it comes out true.

If Strengthening with a Possibility is classically valid, Dynamic SA cannot account for apparent counterexamples to Strengthening with a Possibility in the same way. By the definition of classical validity, $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ is *defined* (and true) in any context in which $A \mathrel{\Box\!\!\rightarrow} C$ and $A \mathrel{\Diamond\!\!\rightarrow} B$ are true. But if $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ is defined, then *asserting* $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ won't change the context. The domain will not expand to make $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ false, as we would hope; $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$ will simply come out true in the original context in which $A \mathrel{\Box\!\!\rightarrow} C$ and $A \mathrel{\Diamond\!\!\rightarrow} B$ are uttered.

Dynamic SA is banking on the (classical) validity of Strengthening with a Possibility, then. It wagers that $A \mathrel{\Box\!\!\rightarrow} C$ and $A \mathrel{\Diamond\!\!\rightarrow} B$ together entail $(A \wedge B) \mathrel{\Box\!\!\rightarrow} C$. In the next section, we show that Strengthening with a Possibility is not valid—it has counterexamples. Dynamic SA rests on a mistake.

## 5 A Counterexample to Dynamic SA

Consider this case:

> *Dice*: Alice, Billy, and Carol are playing a simple game of dice. Anyone who gets an odd number wins \$10; anyone who gets even loses \$10. The die rolls are, of course, independent. What Alice rolls has no effect on what Billy rolls, and *vice versa*. Likewise for Alice and Carol, as well as for Billy and Carol.
>
> Each player throws their dice. Alice gets odd; Billy gets even; Carol

gets odd.

Now consider this sequence of counterfactuals:

(14) a. If Alice and Billy had thrown the same type of number, then at least one person would still have won $10.
b. If Alice and Billy had thrown the same type of number, then Alice, Billy and Carol could have *all* thrown the same type of number. (Because they could have all thrown odd.)
c. If Alice, Billy and Carol had all thrown the same type of number, then at least one person would still have won $10.

(14-a) and (14-b) seem true, but (14-c) is dubious. (14-a) seems right because if Alice and Billy had thrown the same type of number, nothing would have changed with respect to *Carol*—she'd still have rolled odd. So someone would still have won $10.

(14-b) seems right, too. If Alice and Billy had thrown the same type of number, either Alice or Billy would have gotten a different number from the one they actually got. But there's no reason to think it would have been Alice rather than Billy: Billy might have thrown odd, along with Alice and Carol.

But (14-c) seems wrong. There are two ways for Alice, Billy, and Carol to throw the same type of number. They could all roll odd or they could all roll even. And we can't just rule out the latter. If Alice, Billy, and Carol had thrown the same type of number, they might have thrown even, so there might have been no winner: (14-c) is false.

Dynamic SA wrongly predicts that (14-c) follows from (14-a) and (14-b). For (14-a), (14-b), and (14-c) are respectively equivalent to:[12]

[12] In assuming that (14-b) is equivalent to (14-b)*, we assume Duality. However, as we noted, the counterexample does not ultimately rely upon it. We can state the dual of the *would*-counterfactual using wide-scope negation:

(15) a. If Alice and Billy had thrown the same type of number, then at least one person would still have won $10.
b. It's not true that if Alice and Billy had thrown the same type of number, then Alice, Billy and Carol *wouldn't* have all thrown the same type of number.
c. If Alice, Billy and Carol had all thrown the same type of number, then at least one

(16) a. *Alice Billy same* $\Box\!\!\rightarrow$ *someone wins* $*10*

b. *Alice Billy same* $\Diamond\!\!\rightarrow$ (*Alice Billy same* $\wedge$ *Billy Carol same*)

c. (*Alice Billy same* $\wedge$ *Billy Carol same*) $\Box\!\!\rightarrow$ *someone wins* $*10*

Suppose (16-a) and (16-b) are true. Since (16-b) is true, some worlds in the domain are ones where its antecedent and consequent are true—that is, where Alice, Billy, and Carol all throw the same type of number. But that's just to say that (16-c) is *defined*. Dynamic SA Strawson-validates Antecedent Strengthening. So, if (16-a) is true, and (16-c) is defined, then (16-c) must be true, too.

That's wrong. (16-a) and (16-b) are true, and (16-c) is not.

## 6 Dynamic 'Mights' to the Rescue?

In its current form, Dynamic SA cannot account for our judgments about these sentences. Is there a way to modify Dynamic SA so that it can? We don't think so. Let us explain.

We know that someone wins if and only if someone rolls odd. To predict that (14-a) is true, then, there must be someone who rolls odd in all domain-worlds where Alice and Billy roll the same type of number. And to predict that (14-c) is false, some domain-worlds where Alice and Billy (and Carol) roll the same type of number must be ones where everyone rolls *even*. The domain must expand between utterances of (14-a) and (14-c). It must start out containing no worlds where Alice, Billy, and Carol all throw even, and acquire some by the time we evaluate (14-c).

There are only two ways for the domain to expand between (14-a) and (14-c). Either asserting (14-b) expands the domain, or asserting (14-c) does. We've already seen that asserting (14-c) does not expand the domain. It remains to be seen whether asserting (14-b) can.

We've been assuming that *might*-counterfactuals update the domain just like *would*-counterfactuals do, presupposing that their antecedents are possible with respect to the counterfactual domain. But this account will not allow (14-b) to shift

---

person would still won $10.

We notice no difference in our judgements here.

the domain. (14-a) and (14-b) have the same antecedent, so there can be no shifting that is triggered by the latter that isn't already triggered by the former. We know that (14-a) doesn't expand the domain to include worlds where Alice, Billy, and Carol throw even—if it did, (14-a) would come out *false*, but we hear it as true. So asserting (14-b) can't trigger a domain expansion, either.

A different idea can be found in Gillies (2007). He says that the dynamic effect of *might*-counterfactuals is distinct from that of *would*-counterfactuals. The *might*-counterfactual $A \diamondto B$ presupposes that the domain contains worlds where $A$ and $B$ are both true.[13] For example, (14-b) presupposes that the domain contain worlds where Alice, Billy, and Carol all roll the same.[14]

Here's how things would have to go if Gillies' theory is to help with our case. For simplicity, suppose that the initial context in *Dice* is such that, in all domain-worlds, everyone rolls as they actually do—Alice and Carol roll odd, and Billy rolls even. (14-a)'s presupposition isn't met in this context, so asserting (14-a) expands the domain, adding worlds where Alice and Billy roll the same. Suppose we include worlds where Alice and Billy roll even, but none where they roll odd. We can't include any worlds where *Carol* rolls even, lest we render (14-a) false. But if we add no worlds where Alice and Billy roll odd, and no worlds where Carol

[13]Note that at the end of the day, Gilies does not think this assertability condition is a genuine presupposition. While he calls it an entertainability presupposition, he rightly points out that the presupposition of the 'might'-counterfactual is not plausibly an existence presupposition on a quantifier domain. We don't rely on any particular way of cashing out entertainability presuppositions in our arguments.

[14]As Gillies himself notes, this has an unintuitive consequence in a static framework: it makes all utterances of $A \diamondto B$ true. Nonetheless, as Gillies points out, it looks like the Dynamic SA needs to help itself to something like this kind of effect. For *might*-counterfactuals seem to give rise to sequences similar to Sobel sequences. Compare:

(17) a. If Sophie went to the parade, she would have seen Pedro.
b. But if she had gone and been stuck behind someone tall, she might not have seen him.

(18) a. If Sophie had gone to the parade and been stuck behind someone tall, she might not have seen Pedro.
b. #But if she had gone to the parade, she would have seen Pedro.

Without attributing some dynamic effects to *might*-counterfactuals, these sequences would be counterexamples to Dynamic SA, just as our original Sobel sequences were counterexamples to the simplest form of SA. Gillies (2007) chooses to retain this dynamic effect and discusses some ways of addressing the problems it gives rise to.

rolls even, the domain will not contain any worlds where Alice, Billy, and Carol all roll the same type of number, and so the presupposition of (14-b) will not be met. (The consequent of (14-b) is true in a world only if Alice, Billy, and Carol roll the same there.) This means that asserting (14-b) will add worlds where Alice, Billy, and Carol all throw the same type of number. If we include worlds where they all throw even, (14-c) comes out false.

So far things are looking better for Dynamic SA. But trouble is near. One problem is that our Gillies-inspired story cannot be told when (14-b) is uttered *before* (14-a), as in the following sequence:

(14-b) If Alice and Billy had thrown the same type of number, then Alice, Billy and Carol could have *all* thrown the same type of number (because Alice and Billy might have thrown odd along with Carol).

(14-a) If Alice and Billy had thrown the same type of number, then someone would still have won $10.

(14-c) If Alice, Billy and Carol had all thrown the same type of number, then at least one person would still have won $10.

Reversing the order of (14-a) and (14-b) doesn't change our judgments. (14-a) still seems true: If Alice and Billy had rolled the same type of number, *Carol* would still have rolled odd, so someone would still have won $10. And (14-c) still seems false: If Alice, Billy, *and* Carol had rolled the same type of number, they might have all rolled even, in which case *nobody* would have won. The problem is that if (14-b) introduces worlds where Alice, Billy, and Carol all throw even, (14-a) will come out false. That's wrong. Even when uttered after (14-b), (14-a) seems true.

There are also problems when (14) is uttered in its original order. Suppose (14-b) introduces worlds where Alice, Billy, and Carol all throw even. This will indeed make (14-c) false. But it has other, less welcome consequences. Consider the sequence:

(14-b) If Alice and Billy had rolled the same type of number, Alice, Billy, and Carol might have *all* rolled the same type of number (because Alice and Billy might have thrown odd along with Carol).

(19) If Alice and Billy had rolled the same type of number, Carol might not have rolled odd.

(14-b) is true, but (19) is false. Indeed, (19) is false for the same reason that (14-a) is true—there's no reason to suppose that, if Alice and Billy had rolled the same type of number, things might have changed with respect to *Carol*. She would have still rolled odd. But if (14-b) adds to the domain worlds where Alice, Billy, and Carol throw even, (19) will come out true, contrary to intuition.[15]

We don't want (14-b) to add worlds where Alice, Billy, and Carol all roll even. When we evaluate (14-b), we're still holding fixed that Carol rolls odd—that's why we judge (19) false. (We judge (14-b) true not because we think they might have all thrown even, but because we think they might have all thrown odd.) To be sure, things change by the time we get to (14-c). At that point, we are considering worlds where they all throw even—we judge (14-c) false because they might have all thrown even and lost. But it isn't (14-b) that makes those worlds relevant. It is only when we hear (14-c) that we consider worlds where Carol rolls even.

Let's take stock. Our goal was to predict that (14-a) and (14-b) are true, and that (14-c) is false. We said that Dynamic SA classically validates Strengthening with a Possibility. Any context in which (14-a) and (14-b) are true is one in which (14-c) is true. This was an important observation. It meant that we couldn't appeal to the dynamic effects of (14-c) to account for our false judgement of this sentence: If (14-c) is true, it's also defined, so asserting (14-c) won't expand the domain. We then asked whether asserting (14-b) could expand the domain to make (14-c) false. We found that it could, but this offered little comfort. For one, if we reverse the order of (14-a) and (14-b), we wrongly predict that (14-a) is false. For another, if asserting (14-b) adds worlds where Alice, Billy, and Carol roll even, we will indeed predict that (14-c) is false, but we will *also* predict that (19) is true. But (19) is heard as false.

We've seen that asserting (14-b) doesn't expand the domain, and neither does

[15] We can make this same point with the following *would*-counterfactual:

(20) If Alice and Billy had rolled the same type of number, Carol would still have rolled odd.

(20) is intuitively true. But if (14-b) introduces worlds where Alice, Billy, and Carol roll even, (20) will come out false.

asserting (14-c). But if there's no domain expansion between (14-a) and (14-c), Dynamic SA cannot predict a false reading of (14-c).

# 7 Variably Strict Semantics

By its very structure, SA is committed to Strengthening with a Possibility. No assumptions about its underlying closeness relation were needed to prove this. Not so for VSA. Strengthening with a Possibility is not written into the semantics of VSA; rather, it corresponds to a certain formal constraint on the closeness ordering $\preceq_{c,w}$, which has been enforced by many of VSA's proponents, including Stalnaker (1968) and Lewis (1973).[16] We show that by removing that constraint, we can render Strengthening with a Possibility invalid. The resulting logic is that of Kratzer (1981a). Finally, we show how to use a Kratzerian *ordering source* to generate an ordering with the right structure in a principled way.

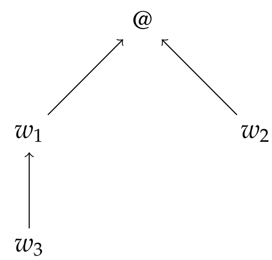
## 7.1 Where we want to go

Here's a model that has the structure we need. Suppose we have four worlds:

- @ is the actual world. (Alice and Carol throw odd; Billy throws even.)
- In $w_1$, Alice and Billy throw even; Carol throws odd.
- In $w_2$, Alice, Billy, and Carol all throw odd.
- In $w_3$, Alice, Billy, and Carol all throw even.

Now suppose we have an ordering that looks like this:

[16]Both Stalnaker (1968) and Lewis (1973) say that whatever else is true about the ordering on worlds, it is *total*: for every $w_1$ and $w_2$ either $w_1 \preceq_w w_2$ or $w_2 \preceq_w w_1$. Total orderings rule out incomparabilities that are essential to the model we give below.




The arrows represent similarity to @. For example, the arrow going from $w_3$ to $w_1$ indicates that $w_1$ is more similar to @ than $w_3$ is. (We omit transitivity and reflexivity arrows.) If there is no arrow between two worlds, then they are *incomparable* with respect to their similarity to @: neither is more similar to @ than the other, and they are not equally similar to @.

In this model, $@ \prec_@ w_1$, $@ \prec_@ w_2$ and $w_1 \prec_@ w_3$. $w_1$ and $w_2$ are incomparable, and $w_2$ and $w_3$ are incomparable.

The closest worlds to @ where Alice and Billy throw the same type of number are $w_1$ and $w_2$: $f(\textit{Alice and Billy same}, @) = \{w_1, w_2\}$. In both worlds, Carol throws odd and wins \$10. Moreover, in one of these worlds—specifically, $w_2$—Alice, Billy, and Carol *all* throw the same type of number. This means that (14-a) and (14-b), repeated below, are true.

(14-a) If Alice and Billy had thrown the same type of number, someone would still have won \$10.

(14-b) If Alice and Billy had thrown the same type of number, Alice, Billy and Carol might have all thrown the same type of number.

Finally, the closest worlds to @ where Alice, Billy, and Carol all throw the same are $w_2$ and $w_3$: $f(\textit{Alice, Billy, Carol same}, @) = \{w_2, w_3\}$. (Why is $w_3$ included? Because the only worlds that are strictly closer to @ than $w_3$ are $w_1$ and @ itself; but Alice, Billy, and Carol do *not* roll the same in either of these.) In $w_2$, Alice, Billy, and Carol throw odd, so they all win. But in $w_3$, they throw even, so *nobody* wins. (14-c), repeated below, is false:

(14-c) If Alice, Billy, and Carol had all thrown the same type of number, someone would still have won \$10.

We've successfully modeled a failure of Strengthening with a Possibility: (14-a) and (14-b) are true, yet (14-c) is false.

The incomparabilities in our model are crucial. Say that the closeness ordering $\preceq_w$ is *almost-connected* just in case $\forall w_1, w_2, w_3 : (w_1 \prec_w w_2 \rightarrow (w_1 \prec_w w_3) \vee (w_3 \prec_w w_2))$. If $w_1$ is closer to $w$ than $w_2$ is, then for any third world $w_3$, either $w_1$ is closer to $w$ than $w_3$ is, or $w_3$ is closer to $w$ than $w_2$ is. Simplifying, if $w_1$ beats $w_2$, then either $w_1$ beats $w_3$, or $w_3$ beats $w_2$. Where $\preceq_w$ is a partial order, Strengthening with a Possibility is valid just in case $\preceq_w$ is almost-connected.[17] Our model predicts that Strengthening with a Possibility fails precisely because it fails to be almost-connected.

## 7.2 Adding a Kratzerian Ordering Source

We've given an ordering that allows VSA to predict the right judgments in *Dice*. But how do we ensure that that the context *actually supplies* an ordering with this structure?

Our suggestion is to follow Kratzer (1981a) and Kratzer (1981b) and posit an extra contextual parameter—an *ordering source*, a function $g$ that takes a world $w$

[17]**Proof**: $\Rightarrow$: Our model that follows demonstrates that if Strengthening with a Possibility is valid, then $\preceq$ is almost-connected. If a frame is not almost-connected, then we can build a model on it like the one in the text.

$\Leftarrow$: Suppose that $\preceq$ is almost-connected and suppose that, for contradiction, that Strengthening with a Possibility is not valid. Then there is some world $w_1$ such that $A \mathrel{\Box\!\!\rightarrow} C$ and $A \mathrel{\Diamond\!\!\rightarrow} B$ are true there but $A \wedge B \mathrel{\Box\!\!\rightarrow} C$ is not. This means that $f(A, w_1) \subseteq C$, there is a world $w_2 \in f(A, w_1)$ such that B is true at $w_2$ and there is a world $w_3 \in f(A \wedge B, w)$ such that $\neg C$ is true there. $w_3$ cannot be in $f(A, w)$: unlike $w_3$ all worlds in $f(A, w)$ are $C$ worlds. By definition of $f$, this means that there must be some world $w_4$ in $f(A, w)$ such that $w_4 \prec_{w_1} w_3$.

Now consider whether either $w_4 \prec_{w_1} w_2$ or $w_2 \prec_{w_1} w_3$. In fact, the first disjunct cannot hold: by definition of $f$, if it did then $w_2$ would not be in $f(A, w_1)$ after all. But the second disjunct cannot be true either. Again by definition of $f$, if it were then $w_3$ would not be in $f(A \wedge B, w_1)$. But now we have proved that, contrary to our supposition that $\preceq$ is not almost connected: $w_4 \prec_{w_1} w_3$ but neither $w_4 \prec_{w_1} w_2$ nor $w_2 \prec_{w_1} w_3$. So if $\preceq$ is almost-connected Strengthening with a Possibility must be valid.

(To the best of our knowledge, this result was first shown by Veltman (1985).)

and returns a set of propositions.[18,19] This set of propositions represents the facts about $w$ that the speakers judge relevant to determining similarity. We then define our ordering in terms of those propositions. $w_1$ is at least as close to $w$ as $w_2$ is if and only it makes true all the same ordering source propositions as $w_2$, and possibly more. Formally:

(21) $w_1 \preceq_w w_2$ iff $\{p \in g(w) : w_1 \in p\} \supseteq \{p \in g(w) : w_2 \in p\}$

$w_1$ is at least as close to $w$ as $w_2$ is just in case every proposition in $g(w)$ that is true in $w_2$ is also true in $w_1$. $w_1$ is strictly closer to $w$ than $w_2$ is just in case, every proposition in $g(w)$ that's true in $w_2$ is true in $w_1$, and some proposition in $g(w)$ that's true in $w_1$ is false in $w_2$.

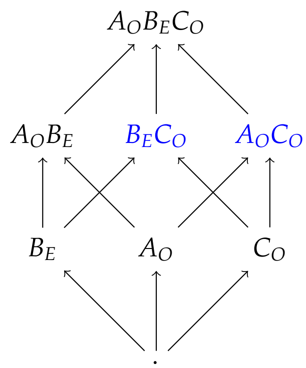
In our example, the relevant facts are those that concern who got what type of number. We assume, then, that the ordering source looks like this:

(22) $g(w) = \{\textit{Alice gets odd}, \textit{Billy gets even}, \textit{Carol gets odd}\}$

Let $A_O$, $B_E$, and $C_O$ be the propositions that Alice rolls odd, Billy rolls even, and Carol rolls odd, respectively. The ordering source in (22) gives rise to the following ordering:

[18] If we were to spell out the semantics in full detail, we would also need to add a Kratzerian *modal base* to our semantics, a function from worlds to sets of propositions. For us, the modal base will hold fixed certain key facts about the scenario, namely the rules of the particular game being played. On our way of thinking of things, the modal base represents the background facts which we hold fixed in evaluating counterfactuals; and the ordering source represents the facts we allow to vary and that contribute to closeness.

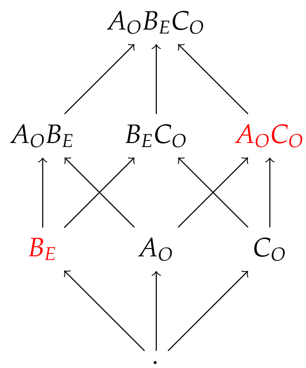
[19] While we add an ordering source to our semantics, we put it to work in a different way to Kratzer. Kratzer's ordering sources for counterfactuals are *totally realistic*: the intersection of the ordering source propositions is the set containing just the world of evaluation. Her modal bases on the other hand are empty. We disagree with Kratzer on both points. We treat modal bases and ordering sources differently in order to predict the failure of Strengthening with a Possibility but we suspect there are other important advantages. Among other things, we appear to avoid the problems Kratzer attempts to solve in Kratzer (1981b) and Kratzer (1989).




In the top-ranked worlds, things are just as they actually are—Alice and Carol roll odd, and Billy rolls even. Next we have worlds where things differ in *one* respect—worlds where either Alice or Carol rolls even instead of odd, or Billy rolls odd instead of even. Then we have worlds differing in *two* respects, and finally, worlds where *everything* is different—Alice and Carol roll even, and Billy rolls odd.

Let's see how VSA predicts the right judgments in *Dice* using this ordering. The closest worlds where Alice and Billy throw the same type of number are in blue. In both worlds, Carol rolls odd and wins $10, so (14-a) is true: If Alice and Billy had rolled the same, one person would still have won $10. Moreover, in one of the closest worlds where Alice and Billy roll the same, Alice, Carol, and Billy all roll odd. So (14-b) is also true: If Alice and Billy had rolled the same, Alice, Billy, and Carol might have all rolled the same.

Finally, turn to (14-c), which says that if Alice, Billy, and Carol had rolled the same, someone would still have won $10. We find the worlds where Alice, Billy, and Carol roll the same type of number. They are highlighted in red:




We have worlds where Alice, Billy, and Carol all throw odd (top right) and worlds where they all throw even (bottom left). These are incomparable—neither is closer to actuality than the other the other is. (The reason they are incomparable is that the sets of ordering source propositions true at each are disjoint.) Both are among the closest worlds where Alice, Billy, and Carol roll the same. So (14-c) is false: In some of the closest worlds where they throw the same, they throw even, and nobody wins.[20]

## 7.3 Why does Strengthening with a Possibility seem valid?

We've argued that Strengthening with a Possibility has counterexamples, and we've offered a new variably strict account that doesn't validate it. Still, the inference often *seems* valid. Suppose I confidently say that if I had taken modal logic last semester, I would have passed. You reply that if I had taken the course, it might have been taught by Joe, who's notorious for his difficult problem sets and harsh grading. If I accept your response, I seem to have two options. I could stand firm, insisting that I would have passed even Joe's challenging course, or I could retreat, rescinding my earlier claim that I would have passed the class. What I *can't* do is maintain that I would have passed the course, even though I might not have passed a course taught by Joe. That I don't have this option is only explained if Strengthening with a Possibility does not fail in this particular case. We must place

[20]We aren't committed to using an ordering source that simply records who gets what type of number in every version of this case. It seems to be the natural one here, but a case with a more complicated structure may call for a more complicated ordering source.

certain constraints on when the inference can fail. We want it to fail in *Dice*, but not here.

We give a pragmatic explanation of this difference. Our idea is that Strengthening with a Possibility fails only when there is more than one relevant way of making the antecedent true. What are the relevant ways of making the antecedent true? We will answer this question by linking ordering sources to salient questions in context and prove that, in the resulting technical sense, Strengthening with a Possibility does indeed only fail when there is more than one way of making the antecedent true.

First, a word about questions. It is standard to let the semantic value of a question be the set of its *answers*. Following Groenendijk and Stokhof (1985), we will assume the set of answers is a set of propositions that are mutually exclusive and exhaustive.[21] That is, questions are a *partition* of the set of worlds. So, for instance, the value of the question *Which of Alice and Bob ate cookies?* would be the set containing the propositions *Alice and Bob ate cookies*, *Alice but not Bob ate cookies*, *Bob but not Alice ate cookies* and *Neither ate cookies*.

Some questions make more distinctions than others. For instance, the question *Which of Alice and Bob ate cookies?* does not distinguish between worlds where *Carol* ate cookies and worlds where she didn't; worlds of both kinds will be contained in every answer to this question. The question *Which of Alice, Bob and Carol ate cookies? does* make such distinctions: all the worlds in any given cell agree on whether Carol ate cookies. In other words, the question *Which of Alice, Bob and Carol ate cookies?* divides up the set of worlds more finely than the question *Which of Alice and Bob ate cookies?*. Representing questions with partitions captures this: every cell in the value of *Which of Alice, Bob and Carol ate cookies?* is a subset of some cell in the value of *Which of Alice, Bob and Carol ate cookies?*, but not *vice versa*. The former makes all the distinctions of the latter, and more besides. As is standard, we will say that $Q'$ *refines* $Q$ iff every element of $Q'$ is a subset of some element of $Q$.

Because questions can distinguish more or less finely between worlds, they have proven to be useful in representing issues that are live in the context. We as-

[21] Note that we do not really mean to take a stand on the semantics of questions in natural language. We simply use this formalism to represent a way that an issue can be live.

sume that in contexts in which we utter counterfactuals, certain *non-counterfactual* questions are relevant. So, for instance, in the Joe case above, a question like *Did I study? And did Joe teach?* is relevant. These questions plausibly represent the kinds of distinctions we are interested in making. We will assume that there is a most refined such question that captures exactly the level of detail that is relevant in a context. It is to these questions that ordering sources will be linked.[22]

To do this, we first define from the ordering source a function $g^{\neg}$. For any given world, $g^{\neg}(w)$ collects together the negations of the propositions in $g(w)$. More formally, $g^{\neg}(w) = \{p : \neg p \in g(w)\}$. From $g$ together with $g^{\neg}$ we can build a partition on worlds. Construct the set of maximal consistent propositions built out of $g(w) \cup g^{\neg}(w)$; call it $G_w$. Since $G_w$ is a partition, it represents a question that we might be concerned with in our context.

We connect this question, $G_w$, to the relevant questions in context. Say that $Q_c$ is the most refined salient (non-counterfactual) question in $c$. We propose that the partition we can construct from our ordering source, $G_{w_c}$ just is the most refined salient question $Q_c$. This has the effect that the ordering source cannot distinguish between worlds in ways that are not already present in the most refined salient question. This does not pin $g$ down uniquely, but does impose a substantial constraint on it: whatever $g_c$ is, it must be the case that $G_{w_c} = Q_c$. We call this constraint *Question Sensitivity*.

Given *Question Sensitivity*, we can prove that we get failures of Strengthening with a Possibility only if there are two distinct answers to $Q_c$ that realise the antecedent of the final strengthened conditional.

> There are $A, B, C$ such that $[\![A \Box\!\!\rightarrow C]\!]^{c,w,g} = 1$, $[\![A \Diamond\!\!\rightarrow B]\!]^{c,w,g} = 1$ and $[\![A \wedge B \Box\!\!\rightarrow C]\!]^{c,w,g} = 0$ only if $\exists p, q \in Q_c : p \vDash A \wedge B$ and $q \vDash A \wedge B$ and $p \neq q$.[23]

[22] Note that these questions must be distinguished from *questions under discussion* in the sense of Roberts (1996). Among other things, the current question under discussion must be *unanswered* in the context, a feature our questions do not share. (If this is felt to be unintuitive, our partitions can instead, following Lewis (1988), be thought of as relevant subject matters.)

[23] **Proof.** Suppose that $A \Box\!\!\rightarrow C$, $A \Diamond\!\!\rightarrow B$, but $A \wedge B \Box\!\!\rightarrow C$ is false. For contradiction, suppose there's just one cell that makes $A \wedge B$ true. Call it $Q$. We appeal to three facts:

1. All worlds in a partition cell are equally good. This is because they all make the same ordering source propositions true.

Put informally, the reason for this is as follows: if there is only one partition cell, call it $p$, that entails $A \wedge B$, then either $p$ is a subset of the closest $A$-worlds or not. If it is, then, if $A \mathrel{\Box\!\!\rightarrow} C$ is true, all the closest $A \wedge B$-worlds will have to be $C$-worlds. If it isn't, then, since all worlds in $p$ are equally close, no $B$-worlds will be among the closest $A$-worlds and so $A \mathrel{\Diamond\!\!\rightarrow} B$ will be false.

To see how this helps, let us return to the case of Joe. Here, quite plausibly the most salient question is *Did I take logic? And did Joe teach?*, which gives us the following partition:

{*I take logic and Joe teaches*, *I take logic and Joe doesn't teach*,
*I don't take logic and Joe teaches*, *I don't take logic and Joe doesn't teach*}

There is only one cell of the partition which makes true our strengthened antecedent, namely, *I take logic and Joe teaches*. We showed that, when there is only one cell entailing the strengthened antecedent, Strengthening with a Possibility cannot fail. So we rightly predict it should seem good here.

We also allow for it to fail in **Dice**. There $G_{w_c}$ is the following partition:

*Only Alice gets odd*, *Only Billy gets odd*, *Only Carol gets odd*, *Only Alice and Billy get odd*, *Only Billy and Carol get odd*, *Only Alice and Carol get odd*, *No one gets odd*, *Everyone gets odd*

Here we do have more than one way of making the strengthened antecedent true: *Only Carol gets odd* and *Everyone gets odd*. So our constraint predicts that Strengthening with a Possibility can fail here, just as we want.

This is an interesting, and previously untapped, advantage of using an ordering source. Not only can it offer an account of when Strengthening with a Possibility fails, it can also be used to explain why the inference often seems to go through.

2. $Q = Q \cap A = Q \cap (A \wedge B)$ This is because $Q$ already contains only $A \wedge B$ worlds.
3. $Q \cap (A \wedge B) = f(A \wedge B, w)$ This follows from the definition of $f$ plus the fact that $Q$ is the only $A \wedge B$ cell.

Either $Q \cap A \subseteq f(A, w)$ or it isn't. Suppose it is. Then, by facts 2 and 3, $f(A \wedge B, w) \subseteq f(A, w)$. And since $f(A, w) \subseteq C$, $A \wedge B \mathrel{\Box\!\!\rightarrow} C$ is true, contrary to our supposition. Suppose it isn't. Then $f(A, w)$ contains *no* $B$-worlds: $Q$ contains the only $A \wedge B$ worlds and, by fact 1, they are all equally good. So $A \mathrel{\Diamond\!\!\rightarrow} B$ is false, contrary to our supposition.

In cases like ours where Strengthening with a Possibility fails, we are interested in *different* ways in which the antecedent could be true. In simple cases, we do not make such fine distinctions and so the inference seems valid.

## 8 Disjunctive Antecedents

The last order of business is to address a potential concern about our case. We say that (14-c), repeated below, is false in *Dice*:

(14-c) If Alice, Billy, and Carol had all rolled the same type of number, someone would still have won \$10.

Why do we judge (14-c) false? Here's a natural explanation. There are two ways for Alice, Billy, and Carol to roll the same type of number—by all rolling odd and by all rolling even. If they had all rolled odd, someone—indeed, all of them—would have won \$10. But not if they all rolled even. If they had all rolled even, nobody would have won. We reject (14-c) because one way of rolling the same is not a way of winning \$10.

Many will worry about this reasoning. It seems to rely on an inference pattern that's perilously similar to a controversial inference pattern known as *Simplification of Disjunctive Antecedents* (Simplification). It's far from obvious that Simplification is valid. And so if the judgment that (14-c) is false does indeed rely on Simplification, that's reason enough to question it. We will argue that despite appearances, we are not relying on Simplification, but a close relative of that principle, which we call *Weak Simplification with a Possibility*. And we will show that, unlike Simplification *itself*, Weak Simplification with a Possibility is validated by VSA.

First let's articulate the worry in more detail. Simplification says that from a counterfactual $(A \vee B) \mathrel{\Box\!\!\rightarrow} C$ one can infer $A \mathrel{\Box\!\!\rightarrow} C$ and $B \mathrel{\Box\!\!\rightarrow} C$.[24] Formally:

(23) **Simplification**. $A \vee B \mathrel{\Box\!\!\rightarrow} C \vDash A \mathrel{\Box\!\!\rightarrow} C \wedge B \mathrel{\Box\!\!\rightarrow} C$

Let's look at an example. Consider (24):

[24] Simplification was discovered by Fine (1975) and independently by Nute (1975).

(24) If I'd taken the bus or the subway to work, I would have arrived at work time.

(24) strongly suggests that both (25) and (26) are true:

(25) If I'd taken the bus to work, I would have arrived on time.

(26) If I'd taken the subway to work, I would have arrived on time.

Why think that our judgment that (14-c) is false relies on something like Simplification? There's a way of thinking about Simplification that doesn't make essential reference to disjunction. There are two salient ways of making the antecedent true. I could take the bus to work or I could take the subway to work. We can think of Simplification as saying that the counterfactual is true only if *every salient way of making the antecedent true* is also a way of making the consequent true. This explains why we judge (24) false when, say, (25) is false. To say that (25) is false is to say that one salient way of making the antecedent true is not a way of making the consequent true.

But *that* reasoning looks very much like the reasoning that lead us to reject (14-c). We said that we judge (14-c) false because there are two ways for Alice, Billy, and Carol to roll the same type of number—they could all roll odd or they could all roll even—and only *one* of them is a way of winning $10.

Why should we be concerned about all of this? Because VSA doesn't validate Simplification. To see this, suppose that in all of the closest worlds where I take the bus, I arrive on time, but in some of the closest worlds where I take the subway, I arrive late. Moreover, the worlds where I take the bus are closer than those where I take the subway. In this scenario, (24) is true. In all of the closest worlds where I take the subway or the bus, I take the bus, so arrive on time. But (26) is false. In some of the closest worlds where I take the subway, I arrive late.

So Simplification is not valid, according to VSA. But, as we said, our judgment that (14-c) is false seems to rely on something like Simplification. But surely there's something wrong with relying on an inference pattern that our own theory doesn't validate.

There's reason to doubt that Simplification really is the principle we're relying on, though. It's well known that the Simplification inference is sometimes *cancelled*. Suppose you're asked which side Spain fought for in World War II, and you reply: 'Spain didn't fight in the war. But if she'd joined either the Axis or the Allies, she would have joined the Axis.'[25] No sane interlocutor would object on the grounds that if Spain had joined the Axis, she wouldn't have joined the Allies.

But (14-c) is structurally similar to the Spain counterfactual. The antecedent, *Alice, Billy, and Carol roll the same type of number*, is equivalent to the disjunction, *Alice, Billy, and Carol roll odd or Alice, Billy, and Carol roll even*. Given knowledge of the game, the consequent, *someone wins* is equivalent to *someone rolls odd*. Together with what's given in the antecedent, *someone rolls odd* is equivalent to the proposition, *Alice, Billy, and Carol all rolled odd*. This means that (14-c) is contextually equivalent to:[26]

(27) If Alice, Billy, and Carol had all rolled odd or all rolled even, they would have all rolled odd.

But (27) is just like the Spain counterfactual: 'If Spain had joined the Axis or the Allies, she would have joined the Axis.' If we aren't tempted by Simplification in the Spain case, we shouldn't be tempted by it in our case, either.[27]

There's another reason to doubt that Simplification is the principle we're relying on. As Alonso-Ovalle (2009) has noted, Simplification inferences are also tempting for *might*-counterfactuals. Take (29):

(29) If I had taken the bus or the subway, I might have arrived on time.

[25]This example is due to McKay and van Inwagen (1977).

[26]We rely on the following notion of contextual equivalence. Where $C$ is a Stalnakerian context (that is, the set of worlds representing the information compatible with what's known by the conversational participants), $[\![A]\!]^c$ and $[\![B]\!]^c$ are contextually equivalent in $C$ just in case, for all $w \in C$: $[\![A]\!]^{c,w} = [\![B]\!]^{c,w}$. Informally, $A$ and $B$ are contextually equivalent just in case they are equivalent given what's known in the context.

[27]This is precisely why dynamic SA predicts that (14-c) is true, even though it (Strawson-)validates Simplification. The domain at this point will contain only worlds where they all get odd. This means the counterfactual

(28) If Alice, Billy and Carol had all rolled even, someone would have won $10.

is *undefined*; and in cases where the conclusion is undefined, a Strawson-valid inference will not seem compelling.

(29) strongly suggests that both (30) and (31) are true:

(30) If I had taken the bus, I might have arrived on time.

(31) If I had taken the subway, I might have arrived on time.

But Simplification is sometimes cancelled with *might*-counterfactuals, just as it is with *would*-counterfactuals. Take (32), for example.

(32) If I had taken the bus or the subway, I might have taken the subway.

When you hear (32), we take it, you are not inclined to infer that if I had taken the bus, I might have taken the subway. Simplification is cancelled in (32).

Return now to *Dice*, and consider the following *might*-counterfactual:

(33) If Alice, Billy, and Carol had all rolled the same, someone might have won \$10.

(33) seems true. If Alice, Billy, and Carol had all rolled the same, they might have all rolled odd, and so they might have all won. This is another example in which the Simplification inference is cancelled. For if it weren't, the sentence would sound absurd. It's plainly not the case that if they had all rolled *even*, someone might have won \$10. Since (33) sounds true, Simplification is not in effect.

(33) is the *might*-counterfactual analogue of (14-c). (14-c) says that if Alice, Billy, and Carol had all rolled the same type of number, someone would still have won \$10; (33) says that if Alice, Billy, and Carol had all rolled the same type of number, someone *might* have won \$10. If Simplification were in effect for (14-c), we would also expect it be in effect for (33). But it isn't. Since (33) sounds true, Simplification must be cancelled.

It's unlikely, then, that our intuitions about (14-c) rely on Simplification. But we think we are relying on a close relative of that principle. And unlike Simplification itself, the close relative *is* validated by VSA.[28] We call the principle *Weak*

[28]**Proof**: Suppose $(A \vee B) \Box\!\!\rightarrow C$ and $(A \vee B) \Diamond\!\!\rightarrow B$ are true at $w_1$. Then $f(A \vee B, w_1) \subseteq C$ and $f(A \vee B, w_1) \cap A \neq \emptyset$.

We can see show $(f(A \vee B, w_1) \cap A) \subseteq f(A, w_1)$. Take an arbitrary $w_2 \in f(A \vee B, w_1) \cap A$. Suppose it were not in $f(A, w_1)$. Then there would have to be a $w_3 \in f(A, w_1)$ such that $w_3 <_{w_1} w_2$. But if

*Simplification with a Possibility*.[29] We divide the principle into two inference patterns:

(35) $(A \vee B) \,\Box\!\!\rightarrow C, (A \vee B) \,\Diamond\!\!\rightarrow A \vDash A \,\Diamond\!\!\rightarrow C$

(36) $(A \vee B) \,\Box\!\!\rightarrow C, (A \vee B) \,\Diamond\!\!\rightarrow B \vDash B \,\Diamond\!\!\rightarrow C$

There are two differences between these inference patterns and Simplification. First, they have weaker conclusions—that $C$ might have occurred, had $A$ occurred (in the case of (35)), and that $C$ might have occurred, had $B$ occurred (in the case of (36)). Second, they have more premises. To infer $A \,\Diamond\!\!\rightarrow C$ from $(A \vee B) \,\Box\!\!\rightarrow C$,

---

$w_3 \in A$ then $w_3 \in A \vee B$. So $w_2$ would not be in $f(A \vee B, w_1) \cap A$ as $w_3 \in A \vee B$ and $w_3 < w_1 w_2$. So $w_2 \in f(A, w_1)$ after all.

But if $(f(A \vee B, w_1) \cap A) \subseteq f(A, w_1)$, then since $(f(A \vee B, w_1) \cap A) \subseteq C$, $f(A, w_1) \cap C \neq \emptyset$. So $A \,\Diamond\!\!\rightarrow C$ is true at $w_1$. If we suppose further that $A \vee B \,\Diamond\!\!\rightarrow B$ is true at $w_1$ we can prove in a similar manner that $B \,\Diamond\!\!\rightarrow C$ is true at $w_1$. So $A \,\Diamond\!\!\rightarrow C \wedge B \,\Diamond\!\!\rightarrow C$ is true at $w_1$.

[29] You might be surprised that we do not rely on a more familiar inference pattern, which we will call Simplification with a Possibility:

(34) $(A \vee B) \,\Box\!\!\rightarrow C, (A \vee B) \,\Diamond\!\!\rightarrow B, (A \vee B) \,\Diamond\!\!\rightarrow A \vDash A \,\Box\!\!\rightarrow C \wedge B \,\Box\!\!\rightarrow C$

This inference pattern has been used in attempts to derive Simplification as an implicature. For instance, Bennett (2003) claims that disjunctive antecedent counterfactuals implicate both of the ‘might’-counterfactuals and so that what look like instances of Simplification are in fact instances of Simplification with a Possibility.

The reason we do not rely on Simplification with a Possibility is simple: it is equivalent to Strengthening with a Possibility, given standard classical assumptions about counterfactuals.

To show this we will assume Substitution of Logical Equivalents holds in the antecedent of counterfactuals; and the further principle that $A \,\Diamond\!\!\rightarrow B$ entails $A \,\Diamond\!\!\rightarrow A \wedge B$, which we will call Conservativity. (Both of these are validated by standard theories of counterfactuals.)

Simplification with a Possibility ⇒ Strengthening with a Possibility:

| | |
|---|---|
| 1. $A \,\Box\!\!\rightarrow C$ | *Assumption* |
| 2. $A \,\Diamond\!\!\rightarrow B$ | *Assumption* |
| 3. $(A \wedge B) \vee (A \wedge \neg B) \,\Box\!\!\rightarrow C$ | 1, *Substitution of Logical Equivalents* |
| 4. $A \,\Diamond\!\!\rightarrow (A \wedge B)$ | 2, *Conservativity* |
| 5. $(A \wedge B) \vee (A \wedge \neg B) \,\Diamond\!\!\rightarrow (A \wedge B)$ | 4, *Substitution of Logical Equivalents* |
| 6. $(A \wedge B) \,\Box\!\!\rightarrow C$ | 3,5, *Simplification with a Possibility* |

Strengthening with a Possibility ⇒ Simplification with a Possibility:

| | |
|---|---|
| 1. $(A \vee B) \,\Box\!\!\rightarrow C$ | *Assumption* |
| 2. $(A \vee B) \,\Diamond\!\!\rightarrow A$ | *Assumption* |
| 3. $(A \vee B) \wedge A \,\Box\!\!\rightarrow C$ | 1,2, *Strengthening with a Possibility* |
| 4. $A \,\Box\!\!\rightarrow C$ | 3, *Substitution of Logical Equivalents* |

we must that know $A$ might have been true, if $(A \vee B)$ had been true. Similarly, to infer $B \diamondto C$ from $(A \vee B) \boxto C$, we must know that $B$ might have been true, had $(A \vee B)$ been true.

Unlike Simplification itself, (35) and (36) are validated by VSA. To see this, take (35). If the *might*-counterfactual premise is true, then some of the closest $(A \vee B)$-worlds are $A$-worlds. Furthermore, those $A$-worlds must constitute at least some of the closest $A$-worlds. So, if $(A \vee B) \boxto C$ is true, some of the closest $A$-worlds are $C$-worlds.

Let's apply this to *Dice*. Recall (14-c), repeated below:

(11-c) If Alice, Billy, and Carol had all rolled the same type of number, someone would still have won $10.

Consider the closest worlds where Alice, Billy, and Carol all roll the same. Some of these worlds are worlds where Alice, Billy, and Carol all roll odd, since (37) is true:

(37) If they had all thrown the same number, they might have rolled even.

By Weak Simplification with a Possibility, (14-c) and (37) entail (38):

(38) If Alice, Billy, and Carol had all rolled even, someone might have won $10.

But (38) is clearly false. By the setup of the case, if they'd all rolled even, nobody would have won $10. But if (38) is false and (37) is true, then, by Weak Simplification of a Possibility, it follows that (14-c) is false.

We've seen that if Weak Simplification with a Possibility is valid, (14-c) must be false. We can also see how Weak Simplification helps make sense of our original judgements about the case. We said that (14-c) seems false because it is too committal—it seems to ignore the possibility that Alice, Billy, and Carol roll the same by all rolling even. Weak Simplification with a Possibility explains why (14-c) seems committal in just this sense. We know, by the setup of the case, that (38) is false. So if (14-c) is to be true, then (37) must be false. But, intuitively, (37) is *true*—if Alice, Billy, and Carol had all rolled the same type of number, it could have easily been that they all rolled even.

The Simplification objection gets *something* right: Our intuitions about (14-c) rely on something *like* Simplification. But it isn't Simplification itself. We've argued that we rely on a close relative of that principle, Weak Simplification with a Possibility. And this principle is validated by VSA.

## 9 Conclusion

We suggested that the debate between SA and VSA could be clarified by looking at a wider range of strengthening principles. This suggestion has been borne out. Dynamic SA validates Strengthening with a Possibility. But this inference is not valid. Counterexamples to Strengthening with a Possibility pose a much more serious problem for Dynamic SA than counterexamples to Antecedent Strengthening itself. While Antecedent Strengthening is merely Strawson-valid, Strengthening with a Possibility is *classically* valid. Counterexamples to it do not involve presupposition failure, so the dynamic principles that drive context change do not apply. But if that's right, Dynamic SA has no way to account for counterexamples to Strengthening with a Possibility. VSA, on the other hand, can easily model failures of Strengthening with a Possibility. We conclude that the failure of Strengthening with a Possibility tells strongly against Dynamic SA and in favor of an ordering source-based version of VSA.

## References


Luis Alonso-Ovalle. Counterfactuals, correlatives, and disjunction. *Linguistics and Philosophy*, 32(2):207–244, 2009.

Jonathan Bennett. *A Philosophical Guide to Conditionals*. Oxford University Press, 2003.

Kit Fine. Review of Lewis's *Counterfactuals*. *Mind*, 84:451, 1975.

Anthony Gillies. Counterfactual scorekeeping. *Linguistics and Philosophy*, 30(3): 329–360, 2007.

Jeroen Groenendijk and Martin Stokhof. On the semantics of questions and the pragmatics of answers. *Semantics: Critical Concepts in Linguistics*, 288, 1985.

Angelika Kratzer. The notional category of modality. *Words, worlds, and contexts*, pages 38–74, 1981a.

Angelika Kratzer. Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic*, 10(2):201–216, 1981b.

Angelika Kratzer. Conditionals. *Chicago Linguistics Society*, 22(2):1–15, 1986.

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

David Lewis. *Counterfactuals*. Blackwell, 1973.

David Lewis. Statements partly about observation. *Philosophical Papers*, 17(1): 1–31, 1988.

Karen Lewis. Counterfactual discourse in context. *Noûs*, 50(4), 2017.

Thomas McKay and Peter van Inwagen. Counterfactuals with disjunctive antecedents. *Philosophical Studies*, 31(5):353–356, 1977.

Sarah Moss. On the pragmatics of counterfactuals. *Noûs*, 46(3):561–586, 2012.

Cory Nichols. Strict conditional accounts of counterfactuals. *Linguistics and Philosophy*, 40(6):621–645, 2017.

Donald Nute. Counterfactuals and the similarity of words. *Journal of Philosophy*, 72(21):773–778, 1975.

Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136, 1996.

Robert Stalnaker. A theory of conditionals. *Americal Philosophical Quarterly*, pages 98–112, 1968.

Robert Stalnaker. A defense of conditional excluded middle. In William Harper, Robert C. Stalnaker, and Glenn Pearce, editors, *Ifs*, pages 87–104. Reidel, 1981.

Frank Veltman. *Logics for Conditionals*. PhD thesis, University of Amsterdam, 1985.

Kai von Fintel. Counterfactuals in a dynamic context. *Current Studies in Linguistics Series*, 36:123–152, 2001.

J. Robert G. Williams. Defending conditional excluded middle. *Noûs*, 44(4):650–668, 2010.