Figure and Ground in Counterfactuals

Dean McHugh*

Draft of October 2025 – comments welcome!

Abstract

If Socrates resembled Adonis, they would both be handsome. If Adonis resembled Socrates, neither would be handsome. And if Adonis and Socrates resembled each other, they might both be handsome and they might not. This is a puzzle for theories of conditionals that predict 'if *A*, would *C*' and 'if *B*, would *C*' to be equivalent whenever *A* and *B* have the same truth conditions. For Socrates resembles Adonis just in case Adonis resembles Socrates, which holds just in case they resemble each other. Moreover, their corresponding facts, or truthmakers, are plausibly taken to be the same: the state of Socrates resembling Adonis is identical to the state of Adonis resembling Socrates, which is identical to the state of them resembling each other. I show that this puzzle cannot be solved by appealing to *de re* readings of the comparator, and offer a solution by developing a semantics of conditionals that is sensitive to what conditional antecedents are about.

Goodman (1947) noticed that logically equivalent sentences can behave differently as conditional antecedents.

(1) a. If New York City were in Georgia, then New York City would be in the South.
b. If Georgia included New York City, then Georgia would not be entirely in the South.

*For insightful discussions of the present material, I am grateful to Arseny Anisimov, Bridget Copley, Chris Kennedy, Aybüke Özgün, Lorenzo Pinton, Benjamin Spector, Joost Zwarts, Zhuoye Zhao and the participants of the Converging on Causal Ontology Analyses (COCOA) Seminar, October 2024.

Examples like Goodman's are important because they appear to challenge Substitution, the widely-endorsed principle of counterfactual logic that whenever two sentences are logically equivalent—true in all the same worlds—they are intersubstitutable salva veritate in counterfactual antecedents.[1] New York City is in Georgia if and only if Georgia includes New York City, but the first seems true while the second seems false.

In addition, I argue that they even challenge the weaker principle of Exact Substitution, stating that whenever two sentences have

This paper makes five points. First, I show that contrasts like those in (14) are no isolated edge case. They are far more widespread than we might initially have thought, appearing with comparatives, quantifiers, identity statements—indeed, any construction whatsoever that distinguishes featuring a subject and object. They turn out to be so commonplace that any theory of counterfactuals ought to say something about them.

Second, I argue that not only do the examples like Goodman's challenge Substitution, they also challenge the weaker principle of *Exact* Substitution, stating that whenever two sentences have the same truthmakers—or, exact verifiers, in Fine's words)

Second, I argue against Lewis's (1973, 43) response to Goodman's pair. Lewis devises two principles to account for the contrast, one appealing to the availability of *de dicto* and *de re* readings, the other appealing to variability in the strictness of the counterpart relation. I argue that these principles, besides being entirely stipulative, conflict with well-established, independent observations governing *de dicto*/*de re* readings and counterpart relations.

Third, I argue that examples like Goodman's pair are best explained in terms of the *figure–ground* distinction (Talmy 1975).

This poses a challenge to Substitution, the rule that whenever two sentences are logically equivalent—true in all the same worlds—they are intersubstitutable salva veritate in conditional antecedents.[2]

Substitution. If A and B are logically equivalent, then 'if A, would C' is true just in case 'if B, would C' is.

A wide range of theories of counterfactuals validate Substitution, including

[1] Substitution is also known as Left Logical Equivalence, or LLE (Arló-Costa 2007). Chellas (1975) calls it RCEA, 'the Rule of Classically Equivalent Antecedents.'

[2] Substitution is also known as Left Logical Equivalence, or LLE (Arló-Costa 2007). Chellas (1975) calls it RCEA, 'the Rule of Classically Equivalent Antecedents.'

Stalnaker (1968) and Lewis (1973a). Other theories, such as Fine (2012), invalidate Substitution by placing stricter demands on when antecedents can be freely replaced without affecting the counterfactual's truth value. Fine proposes that whenever two sentences have the same exact verifiers, they are intersubstitutible salva veritate in conditional antecedents, where a sentence's exact verifiers are, intuitively, those states that are 'wholly relevant' to the sentence's truth Fine (2017c).

Exact Substitution. If *A* and *B* have the same exact verifiers, then 'if *A*, would *C*' is true just in case 'if *B*, would *C*' is.

Goodman's pair seems to challenge even this weaker rule. For the exact verifiers of their antecedents are plausibly taken to be the same. Presumably, both antecedents are exactly verified by the relative locations of New York and Georgia.

Socrates was not known for being handsome, possessing a snub nose and bulging eyes (Plato, *Theaetetus* 143e). Adonis, on the other hand, was famously handsome. With that in mind, consider (4).

(2) a. If Socrates resembled Adonis, they would both be handsome.
b. If Adonis resembled Socrates, they would both be handsome.
c. If Adonis and Socrates resembled each other, they would both be handsome.

The first is intuitively true, the second false, and the third indeterminate.

Counterfactuals like those in (4) are important because pose a challenge Substitution, the widely-endorsed principle that whenever two sentences are logically equivalent—true in all the same worlds—they are intersubstitutable salva veritate in conditional antecedents.[3]

Substitution. If *A* and *B* are logically equivalent, then 'if *A*, would *C*' is true just in case 'if *B*, would *C*' is.

(4) challenges Substitution because resemblance is symmetric. Socrates resembles Adonis just in case Adonis resembles Socrates, which is true just in case they resemble each other.

A wide range of theories of counterfactuals validate Substitution, including Stalnaker (1968) and Lewis (1973a). Other theories, such as Fine (2012), invalidate Substitution by placing stricter demands on when antecedents can be freely replaced

[3]Substitution is also known as Left Logical Equivalence, or LLE (Arló-Costa 2007). Chellas (1975) calls it RCEA, 'the Rule of Classically Equivalent Antecedents.'

without affecting the counterfactual's truth-value. Fine proposes that whenever two sentences have the same exact verifiers, they are intersubstitutible salva veritate in conditional antecedents, where a sentence's exact verifiers are, intuitively, those states that are 'wholly relevant' to the sentence's truth; Fine (2017c).

Exact Substitution. If A and B have the same exact verifiers, then 'if A, would C' is true just in case 'if B, would C' is.

However, examples like (4) challenge even this weaker rule. For their exact verifiers of their antecedents are plausibly taken to be the same (presumably, the state composed Adonis's appearance and Socrates's appearance).

A familiar example of the same contrast is Goodman's (1947) New York–Georgia pair.

(3) a. If New York City were in Georgia, then New York City would be in the South.
b. If Georgia included New York City, then Georgia would not be entirely in the South.

These also challenge to Substitution, since New York City is in Georgia just in case Georgia includes New York City, and to Exact Substitution, since they appear to have the same exact verifiers (namely, the relative positions of New York City and Georgia). However, Lewis (1973a, 43) argues that they do not challenge Substitution after all. He devises two principles that distinguish subjects and objects in counterfactual antecedents. (In (14a), 'New York City' is the subject and 'Georgia' the object, while in (14b) it is the reverse.) First, subjects are interpreted *de dicto* while objects are interpreted *de re*. Where L_G and L_{NYC} are the *actual* locations of Georgia and New York City, respectively, according to this principle we interpret 'if New York City were in Georgia' as 'if New York City were in L_G', and 'if Georgia included New York City' as 'if Georgia included L_{NYC}'—the antecedents are no longer equivalent.

Second, I argue against a response

Third, I argue that contrasts like those in (4) are best accounted for in terms of figure and ground.

aboutness theory of counterfactuals straightforwardly accounts for them.

which has—to my knowledge—so far gone unchallenged in the literature on counterfactuals.

Second, subjects are interpreted using a less stringent counterpart relation than objects.

I argue that these principles provide an unsatisfying resolution to the problem. They violate

'New York City' appears as a subject in (14a) and an object in (14a), while 'Georgia' appears as a subject)

(such as 'New York City' in (14a) and 'Georgia' in (14b)') are interpreted *de dicto* while object terms (such as 'Georgia' in (14a) and 'New York City' in (14b)')

This essay is about one environment where figure–ground relations show up remarkably clearly: counterfactuals. Consider (4).

(4) *Adonis is handsome, while Socrates is not.*

   a. If Socrates resembled Adonis, they would both be handsome.
   b. If Adonis resembled Socrates, they would both be handsome.
   c. If Adonis and Socrates resembled each other, they would both be handsome.

The first is intuitively true, the second false, and the third indeterminate. Their antecedents, however, are logically equivalent. Socrates resembles Adonis just in case Adonis resembles Socrates, which is true just in case they resemble each other. Moreover, their corresponding facts or truthmakers are plausibly taken to be the same: the state of Socrates resembling Adonis is identical to the state of Adonis resembling Socrates, which is identical to the state of them resembling each other.

As we will see, such figure–ground contrasts in counterfactuals appear in a wide array of environments. A number of theories of counterfactuals—including those due to Stalnaker, Lewis, and Fine—cannot account for them. In response, I develop a new theory of counterfactuals, what I call an *aboutness theory*, according to which when we interpret a counterfactual, we vary the part of the world the antecedent is about. Aboutness is hyperintensional: two sentences may be true in all the same worlds but nonetheless differ in what they are about. I proceed to argue against an alternative explanation of our data according to which comparators are read *de re*; for instance, that "if Socrates resembled Adonis" is interpreted as "if Socrates's appearance resembled Adonis' *actual* appearance." I show that this alternative view incorrectly predicts readings to be available which are intuitively not, a difficulty not faced by the aboutness theory.

In all, our data require taking seriously a ubiquitous, but hitherto neglected source of hyperintensionality, one made salient by but not limited to counterfactuals: the contrast between figure and ground.

1 Figure and Ground

The figure–ground distinction goes back to the beginnings of psychology. It was first studied in earnest by Edgar Rubin in his doctoral dissertation (Rubin 1915), in which he introduced Rubin's vase (right). The distinction was important to Gestalt psychologists such as Max Wertheimer, who argued that structured wholes or *Gestalten*, not pure sensory stimuli, are the primary units of mental life (Wertheimer 1923; Wagemans et al. 2012). Perceiving an object as a *Gestalt* requires distinguishing it from its environment.[4]

Around the same time, Russell realised that sentences exhibit an analogous distinction. In *The Principles of Mathematics*, he distinguishes the *subject* and the *assertion* of sentences: "In a relational proposition, say '*A* is greater than *B*,' we may regard *A* as the subject and 'is greater than *B*' as the assertion, or *B* as the subject and '*A* is greater than' as the assertion. There are thus, in the case proposed, two ways of analyzing the proposition into subject and assertion" (Russell 1903, §48). He then proposed that equivalent sentences, by structuring information in different ways, can express distinct propositions.

> The proposition "humanity belongs to Socrates," which is equivalent to "Socrates is human," is an assertion about humanity; but it is a distinct proposition. In "Socrates is human," the notion expressed by *human* occurs in a different way from that in which it occurs when it is called *humanity*, the difference being that in the latter case, but not in the former, the proposition is *about* this notion.
>
> (Russell 1903, §48)

Russell did not apply his observations about comparative statements—such as *A is greater than B*—to counterfactuals. When we do so we find some stark contrasts.

(5) *Alice is 25, Bob is 15. One must be over 21 to enter the bar.*

a. If Alice were younger than Bob, they could both enter the bar.

b. If Bob were older than Alice, they could both enter the bar.

[4] Benjamin Lee Whorf later proposed that the figure–ground distinction is universal, writing that "There is one thing on which all observers of the appearance of a running boy will agree [...]. They will all divide it into (1) a figure or outline having more or less of motion (the boy) and (2) some kind of background or field against which, or in which, the figure is seen" (Whorf 1940, 163).

We naturally read the first as false and the second as true. However, Alice is younger than Bob just in case Bob is older than Alice. When we interpret "If Alice were younger than Bob" we have a default tendency to imagine changing Alice's age and leaving Bob's fixed: Alice's age is the figure, Bob's age the ground. Similarly, when we interpret "If Bob were older than Alice" we tend to imagine changing Bob's age while leaving Alice's fixed. Each sentence privileges a particular division of figure and ground, just as we do when we adopt a particular interpretation of Rubin's vase.

The same contrast appears in comparatives in general. For example:

(6) *Alice is 1.5m tall, Bob is 1m tall. One must be at least 1.2m tall to ride the Ferris wheel.*

a. If Alice were shorter than Bob, they could ride the Ferris wheel together.
b. If Bob were taller than Alice, they could ride the Ferris wheel together.

Leonard Talmy (1975, 2000) has significantly advanced the discussion of figure–ground relations. He thinks of the distinction as follows.

> The FIGURE object is a moving or *conceptually* mov*able* point whose path or site is conceived as a variable the particular value of which is the salient issue. [...] The GROUND object is a reference-point, having a stationary setting within a reference-frame, with respect to which the FIGURE's path or site receives characterization.
>
> (Talmy 1975, 419)

Talmy noticed that figure–ground contrasts appear not only for comparatives, such as *A is younger than B*, but for transitive verbs in general, such as *A is near B*. Like Russell, who argued that *humanity belongs to Socrates* and *Socrates is human* express distinct propositions, Talmy proposed that *the bike is near the house* and *the house is near the bike* "in fact do not say the same thing" (Talmy 1975, 420). In the former, the house "has a set location within a framework [...] and is to be used as a reference-point by which to characterize the other object's (the bike's location), understood as a variable," whereas the latter makes all the reverse specifications. Talmy does not discuss reciprocals such as *the bike and house are near each other*, though these also appear to exhibit a different division of figure and ground compared to *the bike is near the house* and *the house is near the bike*, with the figure consisting of both objects.

Counterfactuals are sensitive to this division of figure and ground. For example:

(7) *Alice is at home and Bob is outside in the rain.*

a. If Alice were near Bob, they would both be dry.
b. If Bob were near Alice, they would both be dry.
c. If Alice and Bob were near each other, they would both be dry.

We naturally judge the first to be false, the second to be true, and the third to be indeterminate. Though Alice is near Bob just in case Bob is near Alice, which holds just in case they are near each other.[5]

All sentences whatsoever that express spatial comparisons appear to exhibit figure–ground contrasts: *near*, *above/below*, *to the left/right of*, and so on. Suppose that three blocks, A, B, and C, are placed in a grid as in Figure 1, where each block occupies exactly one square. Consider (8).

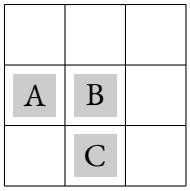
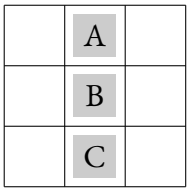


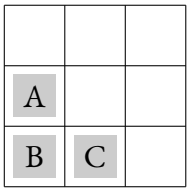
Figure 1

(8) a. If block A were on top of block B, the blocks would form a straight line.
b. If block B were beneath block A, the blocks would form a straight line.

The first is most naturally judged true and the second false. Figure 2 depicts the arrangements we tend to imagine when we interpret each antecedent.



(a) If block A were on top of block B, ...



(b) If block B were beneath block A, ...

Figure 2

Talmy observes that non-spatial relations also express a figure–ground distinction. Take resemblance: "the sentence 'she resembles him,' which might be thought to derive from something like 'she is near him in appearance, or her appearance is near his appearance,' is not understood in the same sense as 'he resembles her' " (Talmy 1975, 421). In conditionals this contrast surfaces as a difference in truth

[5] See Dowty (1991, 556) and Gleitman et al. (1996) for arguments that *near* is truth-conditionally symmetric.

conditions. We saw this in (4) with *resembles*. With *similar* we see an analagous contrast.

(9) a. If Bach were similar to Black Sabbath, Bach would be played less at church.
b. If Black Sabbath were similar to Bach, Bach would be played less at church.
c. If Bach and Black Sabbath were similar, Bach would be played less at church.

1.1 Figure–Ground Contrasts in Quantifiers

Spector and Mouly (2025) observe that Quantifiers also exhibit figure–ground contrasts.[6] As Hempel (1945) noticed, *Every raven is black* is true just in case *Everything that is not a raven is not black* is, but they seem to differ in what observations support them. We see a similar contrast in counterfactuals. Imagine a simple universe each object must be either a raven or a parrot, and either black or white. Thus in this universe there are four possible kinds of object: a black raven, a white raven, a black parrot, and a white parrot (Figure 3).

Black raven White raven Black parrot White parrot

Figure 3

Consider (10), for example, in the context of Figure 3.

(10) a. If every raven were black, there would be three black birds.
b. If every white bird were a parrot, there would be three black birds.

Intuitively, the first is true and the second false. When we imagine every raven black, we tend to change the ravens and leave the non-ravens as they are. While when we imagine every white bird a parrot, we tend to change the white birds and leave the non-white birds as they are (Figure 4).

[6] I am grateful to Benjamin Spector for bringing the case of quantifiers to my attention.

(a) If every raven were black, ... (b) If every white bird were a parrot, ...

Figure 4

Given that every object must be either a raven or a parrot, and either black or white, every raven is black just in case every white bird is a parrot. In this limited universe, then, the antecedents in (10) have the same truth conditions.[7]

The point applies broadly. *No A is B* and *No B is A* have the same truth conditions, as do *Some A is B* and *Some B is A*. But they come apart in counterfactual antecedents.

(11) a. If no raven were white, there would be three black birds.
b. If no white bird were a raven, there would be three black birds.

(12) a. If some white raven were black, there would be three black birds.
b. If some black bird were a white raven, there would be three black birds.

In the context of Figure 3, intuitively the first sentence in each pair is true and the second false.

One might object that our knowledge of the world—say, our knowledge that black ravens are more plentiful than white ravens—impacts our judgements in these cases. In reply, consider a more clinical setting. Imagine a simple universe in which each object must be either a square or circle, and either light or dark. Consider four shapes, arranged as in Figure 5.

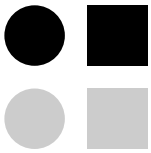
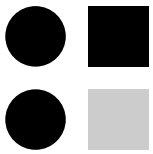


Figure 5

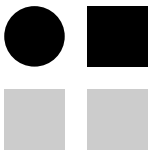
"Every circle is dark" has the same truth conditions as "every non-dark object is not a circle," which in this limited universe is equivalent to "every light shape is square." But these sentences intuitively give rise to different possibilities as conditional antecedents, illustrated in Figure 6.

Here is a final, more subtle example. Talmy claims that in a sentence such as $y = 3x^2 + 1$, y appears "figure-like" while x appears "ground-like" (Talmy 1975, 422). While this case is less clear-cut, it arguably gives rise to similar contrasts in

[7] If one is not willing to buy this restriction to such a simple universe, we may replace "Every white bird is a parrot" with "Every non-black bird is not a raven." While this formulation is more cumbersome, the intuitive judgement remains the same.



(a) If every circle were dark, ...



(b) If every light shape were square, ...

Figure 6

conditionals. Let m, n, and k be three numbers. We are told that $m = 5$, $n = 2$, and $k = 3$. Compare (13) in the following two contexts.

(13) *Context 1: we are told that $m = n + k$.*
Context 2: we are told that $n = m - k$.
If n were 4, m would be 7.

In context 1 we seem to have a default preference for changing n but not k, in which case the conditional is true. In context 2 we are less sure whether to change m, k, or both, leading to more prevalent judgements of indeterminacy. But $m = n + k$ is true just in case $n = m - k$ is.

1.2 Goodman's New York–Georgia Pair

A more familiar example of a figure–ground contrast is Goodman's (1951) New York–Georgia pair.

(14) a. If New York City were in Georgia, then New York City would be in the South.
b. If Georgia included New York City, then Georgia would not be entirely in the South.

Lewis (1973a) argues that appeal to *de re* readings and the counterpart relation can explain their simultaneous truth.

> We can explain the simultaneous truth of Goodman's sentences [...] by the hypothesis that both are *de re* both with respect to 'New York City' and with respect to 'Georgia', and that a less stringent counterpart relation is summoned up by the subject terms 'New York City' in [(14a)] and 'Georgia' in [(14b)] than by the object terms 'Georgia' in [(14a)] and 'New York City' in [(14b)]. Then in [(14a)] we are concerned with

> the closest worlds to ours where a not-too-close counterpart of our New York is in a close counterpart of our Georgia, and hence is in (a counterpart of?) the South; whereas in [(14b)] we are concerned with the closest worlds to ours where a not-too-close counterpart of our Georgia includes a close counterpart of our New York City, and hence is not entirely included in the South.
>
> (Lewis 1973a, 43)

Lewis does not offer any indication why the counterpart relations would differ between subject and object terms in this way, writing, "I cannot say in general how grammar and context control which counterpart relation is used where" (Lewis 1973b, 436). This is unsatisfying. There is prima facie no reason to expect subject and object terms to select for different counterpart relations. It is unexpected from how Lewis understands the counterpart relation, which is "based on similarity" (1973, 41). For Lewis, "something has for counterparts at a given world those things existing there that resemble it closely enough in important respects of intrinsic quality and extrinsic relations, and that resemble it no less closely than do other things existing there" (1973, 39). These are eminently plausible ideas, but nothing in them leads us to expect that subject and object terms should differ in their stringency of counterpart relations.[8] I beleive we can give a more satisfying, independently-motivated account by appeal to the figure–ground relation. This relation is previously attested in the work of the Gestalt psychologists, Russell and Talmy, and is not specific to the modal environments where counterpart relations appear.

[8]One might reply that counterpart relations for geographical regions are complex: if the location of New York City or borders of Georgia were changed, it is hard to say in what sense each place would still count as the same object that it actually is. However, other cases we have discussed are much clearer. Consider (8). In each of the resulting worlds (illustrated in Figure 2) the counterpart relations are easily discerned: the counterpart of block A is the block with 'A' written on it, similarly for the other blocks. Or when in (5), we are asked to imagine Alice younger than Bob and then Bob older than Alice, it is clear who Alice and Bob's counterparts would be in the resulting worlds. It is said that age is just a number, a change in age is not enough to render the counterpart relations hard to discern.

2 The Puzzle

The contrasts we have just observed appear to be counterexamples to the following rule of Substitution, where $A > C$ denotes "if A, would C."[9]

Substitution. If A and B are logically equivalent, then $A > C$ is true if and only if $B > C$ is true.

For Socrates resembles Adonis just in case Adonis resembles Socrates, which is true just in case they resemble each other. Alice is younger than Bob just in case Bob is older than Alice. Alice is shorter than Bob just in case Bob is taller than Alice. Alice is near Bob just in case Bob is near Alice, and so on.

Almost all semantics of *would*-conditionals validate Substitution, including those of Stalnaker (1968) and Lewis (1973a).[10] Indeed, any theory that only makes use of antecedents' truth conditions will validate Substitution: if A and B have the same truth conditions, such a theory cannot distinguish them. One might therefore propose that the fault lies with the coarse grain of truth conditions, and that by moving to a more fine-grained, hyperintensional notion of meaning, in which logically equivalent sentences can be distinguished, the overall architecture of our theory of conditionals might be maintained.

The problem, however, runs deeper than that. One way to see this is to note that Substitution follows from the following two principles.[11]

Weaker than Entailment. If A entails C, then $A > C$ is true.[12]

Reciprocity. If $A > B$ and $B > A$ are true, then $A > C$ is true if and only if $B > C$ is true.

[9] Substitution is also known as Left Logical Equivalence, or LLE (Arló-Costa 2007). Chellas (1975) calls it RCEA, 'the Rule of Classically Equivalent Antecedents.'

[10] Semantics of conditionals that invalidate Substitution include Nute (1980, 32), Fine (2012), Ciardelli, Zhang, and Champollion (2018a), and Santorio (2018). Nute and Santorio do so to validate Simplification of Disjunctive Antecedents—that $(A \vee B) > C$ implies $A > C$ and $B > C$—without validating Antecedent Strengthening—that $A > C$ implies $(A \wedge B) > C$. For A is equivalent to $A \vee (A \wedge B)$. Given Substitution, $A > C$ would imply $(A \vee (A \wedge B)) > C$, which by Simplification implies $(A \wedge B) > C$. Nute (1980, 34) claims that apart from this issue there is no evidence against Substitution.

[11] The name Reciprocity comes from Egré and Rott (2021). Nute (1980) calls it CSO, and Starr (2022) calls it Substitution of Subjunctive Equivalents.

[12] Weaker than Entailment in turn follows from plausible principles such as Identity ($A > A$) and Right Weakening (if B entails C then $A > B$ entails $A > C$). For a discussion of semantics that—surprisingly—invalidate Identity, see Mandelkern (2019, 2021).

For if A and B are logically equivalent, they entail one another, so by the Weaker than Entailment, $A > B$ and $B > A$ are both true. Then by Reciprocity, $A > C$ is true just in case $B > C$ is. Thus no matter how fine-grained we take meanings to be, provided our theory of conditionals validates Weaker than Entailment and Reciprocity, it will automatically validate Substitution too.

In what follows I show that two of the most well-known theories of conditionals, due to Stalnaker and Lewis, validate Weaker than Entailment and Reciprocity, and therefore must validate Substitution. I will then show that a theory of conditionals of a quite different character—Fine's (2012) truthmaker semantics of conditionals—validates an analogue of Substitution, given some additional assumptions which Fine himself endorses.

2.1 Stalnaker

Stalnaker introduces a selection function that takes a proposition and a possible world and returns a single world—intuitively, one in which the antecedent is true and which otherwise differs minimally from the world of evaluation. Stalnaker validates Weaker than Entailment and Reciprocity, due to two constraints he imposes on the selection function f.[13]

(i) A is true at $f(A, w)$.
(ii) If A is true in $f(B, w)$ and B is true in $f(A, w)$, then $f(A, w) = f(B, w)$.

Stalnaker regards these constraints as essential to his theory, remarking that "the selection is based on an ordering of possible worlds with respect to their resemblance to the base world. If this is correct, then [(ii)] must be imposed [...]. These conditions on the selection function are necessary in order that this account be recognizable as an explication of the conditional" (Stalnaker 1968, 36).

[13] To see this, suppose A and B are logically equivalent. By (i), A is true at $f(A, w)$. Then given that A and B are logically equivalent, B is true at $f(A, w)$. Vice versa, A is true at $f(B, w)$. Then by (ii), the selected A-world is the selected B-world. A fortiori, C is true at the selected A-world just in case it is true at the selected B-world.

2.2 Lewis

Lewis (1973a) proposes that $A > C$ is true at a world w just in case C is true at all of the most similar worlds to w where A is true.[14] Like Stalnaker, Lewis also validates Weaker than Entailment and Reciprocity, from which Substitution follows. No matter how fine-grained we take propositions to be, Substitution is part of the logic of Lewis's semantics.[15]

2.3 Fine

The problem runs deeper than merely arguing against Substitution. For a number of recent theories that invalidate Substitution in general nonetheless predict the antecedents we have considered to be equivalent.

Fine (2012) proposes a semantics of counterfactuals that does not validate Substitution. It is based on the notion of *exact verification*. The guiding idea is that a part of the world, or state, exactly verifies a sentence just in case the state's obtaining is wholly relevant to the truth of the sentence (Fine 2017c). Fine also uses a transition

[14] This formulation holds under the limit assumption (for discussion see Kaufmann 2017). As we show in note 15, Lewis validates Weaker than Entailment and Reciprocity regardless of whether we make the limit assumption.

[15] Let us show this here. For each world w, let $\leq_w$ be a reflexive and transitive binary relation over the set of possible worlds. (Lewis (1981, 220) himself begins with a irreflexive order $<_w$ and constructs a reflexive order $\leq_w$ by taking $w' \leq_w w''$ just in case $w' <_w w''$ or $w' = w''$.) Lewis (1981, 230) takes $A > C$ to be true at w just in case for every A-world x, there is an A-world $y \leq_w x$ such that for every world $z \leq_w y$, if A is true at z, C is true at z. Weaker than Entailment is immediate. We prove that Reciprocity is valid as follows.

Proof. For any world w and sentence A, let $w \models A$ denote that A is true at w. Pick any world w and suppose $A > B$, $B > A$ and $B > C$ are true at w. To show that $A > C$ is true at w, pick any $x \models A$. We have to show that there is a $y \models A$ such that $y \leq_w x$ and for all $z \leq_w y$, $z \models A \supset C$, where $\supset$ is the material conditional.

Since $w \models A > B$ and $x \models A$, there is a $v \models A$ such that $v \leq_w x$ and (i) for all $v' \leq_w v$, $v' \models A \supset B$. Since $\leq_w$ is reflexive, $v \leq_w v$, so $v \models A \supset B$. Thus $v \models B$.

Since $w \models B > A$ and $v \models B$, there is a $u \models B$ such that $u \leq_w v$ and (ii) for all $u' \leq_w u$, $u' \models B \supset A$.

Since $w \models B > C$ and $u \models B$, there is a $y \models B$ such that $y \leq_w u$ and (iii) for all $z \leq_w y$, $z \models B \supset C$. Since $y \leq_w u$, by (ii), $y \models B \supset A$. Then as $y \models B$, $y \models A$. And as $y \leq_w u \leq_w v \leq_w x$, by transitivity of $\leq_w$, $y \leq_w x$.

We show that $z \models A \supset C$ for all $z \leq_w y$. Pick any $z \leq_w y$. Then $z \leq_w y \leq_w u \leq_w v$, so by transitivity of $\leq_w$, $z \leq_w v$. Then by (i), $z \models A \supset B$. And since $z \leq_w y$, by (iii), $z \models B \supset C$. Hence $z \models A \supset C$.

relation between states, stating when one state is a possible outcome of another at a world.[16] He offers the following semantics.

> $A > C$ is true at a world w just in case for every exact verifier t of A and possible outcome u of t at w, u contains an exact verifier of C. (Fine 2012, 237)

Following Fine (2014, 576), let us call two sentences *exactly equivalent* just in case they have the same exact verifiers and falsifiers. Two sentences may be logically equivalent without being exactly equivalent; for example, A, $A \vee (A \wedge B)$, and $(A \wedge B) \vee (A \wedge \neg B)$ are logically but not exactly equivalent. Fine's semantics therefore does not validate Substitution, but it does satisfy an analogous principle.

Exact Substitution. If A and B are exactly equivalent, then $A > C$ is true if and only if $B > C$ is true.

Recall from (8) that the "Block A is on top of block B" and "Block B is beneath block A" behave differently in counterfactual antecedents. There is, however, a compelling argument to be made that these statements are exactly equivalent. Fine himself puts it like this.

> Suppose that a given block a is on top of another block b. Then there is a certain state of affairs s_1, we may describe as the state of a's being on top of b. There is also a certain state of affairs s_2 that may be described as the state of b's being beneath a. Yet surely the states s_1, and s_2 are the same. There is a single state of affairs s "out there" in reality, consisting of the blocks a and b having the relative positions that they do; and the different descriptions associated with s_1, and s_2 would merely appear to provide two different ways at getting at this single state of affairs.
>
> (Fine 2000, 3)

According to Fine, then, *Block a is on top of block b* and *Block b is beneath block a* are exactly equivalent.[17] However, in (8), repeated below, we saw these are not equivalent in counterfactual antecedents.

---

[16] We will not be concerned with the transition relation here since as we will see, our points to come apply regardless of which transition relation is in question.

[17] Geach and Williamson propose a related, though not identical view. Geach states that "a relation neither exists nor can be observed apart from its converse relation" (Geach 1957, 33), and Williamson argues that relations are identical to their converses: " '- - - stabs ...' and '- - - is stabbed by ...' (and '... stabs - - -') stand for the same relation" (Williamson 1985, 249). For discussion see MacBride (2007), Gaskin and Hill (2012), and Liebesman (2013).

(8) a. If block A were on top of block B, the blocks would form a straight line.
b. If block B were beneath block A, the blocks would form a straight line.

Assuming that their antecedents are exactly equivalent, then, (8) is a counterexample to Exact Substitution.

For a second example, consider the relation of adjacency. Fine writes that "the state of *a*'s being adjacent to *b* is surely the same as the state of *b*'s being adjacent to *a*" (2000, 17). Now consider (15).

(15) *Alice has little money. Her house is in an affordable area and her office is in an expensive area.*
a. If Alice's house were adjacent to her office, her rent would be unaffordable.
b. If Alice's office were adjacent to her house, her rent would be unaffordable.
c. If Alice's office and house were adjacent, her rent would be unaffordable.

Intuitively, the first is true, the second false, and the third indeterminate.

Here I will not attempt to settle the complex question of whether we ought to regard the states in question as identical, as Fine does, or distinct. Nonetheless, it is clearly desirable to have a theory of conditionals that makes the right predictions for sentences like (8) and (15) without controversial commitments regarding the ontology of relations. To predict the contrasts in (8) and (15) on Fine's approach to conditionals, we must assume that the state of *a* being on top of *b* is distinct from the state of *b* being below *a*, and that the state of *a* being adjacent to *b* is distinct from the state of *b* being adjacent to *a*, which is in turn distinct from the state of *a* and *b* being adjacent, *pace* Fine himself. However, one can consistently maintain that these states are identical while nonetheless acknowledging a contrast in (8) and (15). Fine's semantics of conditionals rules out this position as inconsistent.

The same point applies to the other contrasts we have observed. It is plausible to maintain that the state of Socrates resembling Adonis is identical to the state of Adonis resembling Socrates, which is in turn identical to the state of them resembling each other. The state of Alice being younger than Bob is just the state of Bob being older than Alice, and so on. Given these identities, each of the contrasts we observed in conditionals above are counterexamples not only to Substitution, but also to

Exact Substitution.[18]

2.4 Reciprocals

Quite apart from ontological considerations, there are arguments of a purely linguistic character that, augmented with some assumptions about the relationship between meanings and state spaces, speak in favour of treating these states as identical; for example, that the state of Alice being similar to Bob is identical to the state of Bob being similar to Alice. The argument comes from reciprocals.

Reciprocals are constructions that express a shared relationship, such as *A and B are similar*, *A and B are close*, *A and B are near each other*, and *A and B resemble each other*. Each reciprocal has binary forms; for example, the binary forms of *A and B are similar* are *A is similar to B* and *B is similar to A*.

In general there are two approaches one may adopt to the relation between reciprocals and their binary forms. We may treat the binary form as primary and derive the reciprocal from it (Gleitman 1965; Dimitriadis 2008; Rubinstein 2009; Siloni 2012), what we will call a *relation-first approach*, or we may treat the reciprocal as primary and derive the binary form from it (Lakoff and Peters 1969; Lasersohn 1995; Winter 2018), what we will call a *reciprocal-first approach*. The relation-first approach takes the relations *A is similar to B* and *B is similar to A* as primitive and defines that *A* and *B* are similar just in case *A* is similar to *B* and *B* to *A*. The reciprocal-first approach goes the other way, positing a primitive predicate *similar* that applies to pluralities, and defines that *A* is similar to *B*, and *B* to *A*, just in case *A* and *B*—understood as a plural individual—are similar.

Lasersohn (1995) offers the following argument for the reciprocal-first approach. As Goodman (1951, 108) observes, some things can be pairwise similar without the group as a whole counting as similar. Consider (16).[19]

[18] These considerations also show that Kratzer's semantics of counterfactuals does not account for the figure–ground contrasts we have observed. Kratzer (1981, 1989, 2012) offers a semantics of counterfactuals in terms of situation semantics. Situations are parts of possible worlds, and propositions are evaluated at situations rather than whole worlds. Without going into the details of her account here (see Kratzer 2012, 133), it suffices to note that on Kratzer's theory, propositions that are true in exactly the same situations are predicted to be intersubstitutable salva veritate in conditional antecedents. Given that, say, "A resembles B" and "B resembles A" are true in identical situations (similarly for the other contrasts we have seen), Kratzer therefore faces the same difficulties as Fine in accounting for figure–ground contrasts in conditionals.

[19] Amos Tversky presents a related example, based on one due to William James: "Jamaica is similar to Cuba (because of geographical proximity); Cuba is similar to Russia (because of their political

(16) a. John, Mary, and Bill are similar.
b. John is similar to Mary, Mary is similar to Bill, and Bill is similar to John.

The first intuitively requires John, Mary, and Bill to be similar in the same respect. The second does not: it is true, say, "if John is similar to Mary in having red hair, but similar to Bill in being seven feet tall, and Mary is similar to Bill in liking peach ice cream" (Lasersohn 1995, 29). The same observations hold for other predicates, such as *agreed* (Winter 2018) and *close*.

(17) a. John, Mary, and Bill agreed.
b. John agreed with Mary, Mary agreed with Bill, and Bill agreed with John.

(18) a. John, Mary, and Bill are close.
b. John is close to Mary, Mary is close to Bill, and Bill is close to John.

This challenges the relation-first approach, since it is not at all clear how to derive the shared-feature interpretation via compositions of binary relations between individuals. Given that John agrees with Mary, Mary with Bill, and Bill with Bob, that does not determine whether all three agree. If the pairs happen to all agree on the same point, then all three agree, while if each pair agrees on a different point, the three do not.

No such difficulties arise on the reciprocal-first approach. We assume a primitive meaning of *similar* that applies to pluralities: a plurality P is similar just in case there is a relevant feature that all members of P share; P agree just in case there is a statement that they all believe; P are close just in case there is some respect in which the group collectively exhibits sufficient companionship. Using this collective predicate, we define that A is similar to B just in case the plurality consisting of A and B is similar. This accounts for the fact that A is similar to B just in case A and B are similar, while also accounting for Goodman's observation that A, B and C can be pairwise similar without being collectively so.

Under some additional assumptions, this purely linguistic argument provides evidence for a particular view concerning the identity and distinctness of the states assumed by truthmaker semantics. There are various perspectives one may take on the ontological status of the state space. A legitimate possibility—though by no means the only one—is cognitive: the states represent the way our minds carve up

affinity); but Jamaica and Russia are not similar at all" (Tversky 1977, 329).

the world. This gets us to cognition, though not yet to natural language ontology. Moving one step further, a plausible (albeit controversial) view is that the ontology of natural language serves as a guide to the structure of our conceptual life (Fodor 1987; Pietroski 2018; Moltmann 2019). Indeed, this assumption is a key tenant of cognitive linguistics (Fodor 1975; Lakoff 1987; Croft and Cruse 2004).

I do not wish to advocate for these views here, but merely point out they are intelligible and legitimate positions one may adopt. For those who hold such views, arguments concerning which relations to adopt as primitive in one's semantics of natural language serve as evidence for which relations are represented in the state space assumed by truthmaker semantics. In this case, the above argument for the reciprocal-first theory serves as evidence that the state space contains the symmetric state of John and Mary being similar, rather than the asymmetric states of John being similar to Mary and Mary being similar to John. Given the absence of these asymmetric states in the state space, *A is similar to B* and *A is similar to B* are exactly equivalent. They are therefore equivalent as antecedents according to Fine's semantics of counterfactuals, which as we have seen in (8) and (15), is an undesirable prediction.

3 An Aboutness Theory of Counterfactuals

We have seen that conditionals are sensitive to figure–ground contrasts, and that the theories of Stalnaker, Lewis and Fine do not account for this fact. The goal of this section is to develop an alternative theory, based on a notion of *aboutness*, and show how it accounts for figure–ground contrasts in counterfactuals.

The key tenent of the aboutness theory is that when we interpret a counterfactual, we change the part of the world the antecedent is about. We split the world into *foreground* and *background*, allowing the foreground to vary while fixing the background. In other words, the background is the *ceteris* in *ceteris paribus*—the "all else" in "all else being equal."

Ciardelli, Zhang, and Champollion (2018b) make use of the same foreground–background distinction in formulating their *background semantics* of counterfactuals. They propose that "Background facts are held fixed while making a counterfactual assumption, while foreground facts are allowed to change" (2018b, 599).[20]

[20]While Ciardelli, Zhang, and Champollion (2018b) formulate their account in terms of a foreground–background distinction, their ultimate proposal does not account for the figure–ground contrasts we have observed. They propose that a fact is allowed to change when it is "called into

There is a striking parallel between the foreground–background distinction, as it appears in the semantics of counterfactuals, and the figure–ground distinction, as studied by Gestalt psychologists, Russell, and Talmy.[21] In particular, the idea that we allow the foreground to vary and fix the background parallels how Talmy understands the figure–ground distinction, where the figure is "a moving or *conceptually* mov*able* point whose path or site is conceived as a variable," the ground "a reference-point, having a stationary setting within a reference-frame" (Talmy 1975). We vary the figure and fix the ground.

It is tempting to think that these two distinctions—the figure–ground distinction on the one hand, studied by Gestalt psychologists and Talmy, and the foreground–background distinction on the other, as it applied in the semantics of counterfactuals—are in fact one and the same. Thus if we follow Talmy in saying that in "Alice is near Bob," by default Alice is the figure and Bob the ground, this gives us an immediate prediction for how counterfactuals behave. We predict that when we put this sentence in a counterfactual antecedent, "if Alice were near Bob," we vary Alice but not Bob. As we saw in (7), this is precisely what we observe.

Let us situate this within the aboutness theory to counterfactuals, which involves not only making a change to a world but also considering how the world evolves after that change. On this approach, when we interpret counterfactuals we are like film editors. We pick a frame and select a part of that frame to edit. The edit changes not only that frame itself, but has downstream consequences for how the film plays out. The edit, however, doesn't change the past. What we end up with is a Frankenstein-film of sorts, with the past from one film and the present and future from another, a film that likely never existed until we constructed it by an act of our imagination. The theory interprets counterfactuals in the following steps.

---

question" by the conditional antecedent. This occurs when the fact either (i) contributes to the falsity of the antecedent in the actual world, or (ii) is causally dependent on such a fact. They define that a fact f contributes to the falsity of a proposition p at a world w just in case there is some set of facts F at w that is consistent with p, but $F \cup \{f\}$ is not. For them consistency is a purely intensional notion: a set of propositions is consistent with a proposition just in case there is a possible world where they are jointly true (see Ciardelli, Zhang, and Champollion 2018b, note 27). This account therefore does not have the resources to distinguish between logically equivalent antecedents, such as "Alice is younger than Bob" and "Bob is older than Alice" that appear in figure–ground contrasts.

[21] As we saw in section 1, Russell did not conceive of the distinction in terms of figure and ground—terminology that only appeared later—but in terms of subject and assertion (Russell 1903). The differing terminology seems to nonetheless track the same distinction.

1. Pick a time at which to imagine the change. This is the *intervention time*.[22]
2. Allow the part of the world the antecedent is about at intervention time to vary.
3. Play forward the laws.
4. Stick on the actual past.
5. Restrict to worlds where the antecedent holds.
6. Check if the consequent holds at all of/the selected resulting world(s).

The distinction between figure and ground comes in step 2. The figure is the part of the world the antecedent is about at intervention time, the ground everything else.

When we imagine changing a part of the world in response to a conditional antecedent, we change not only the parts of the world the antecedent is directly or wholly about, but also those parts it is indirectly or partially about. To illustrate the point, consider Figure 7, of a circle on a background.

Figure 7

Imagine changing the colour of the circle, say, by interpreting the antecedent, "If the circle were green, … ." Plausibly, this is fully about the colour of the circle, but only partially about the left half of the image. But now suppose that we kept the left half of the image the same when imagining the circle a different colour. By parity or reasoning, we would also keep the right half of the image the same. But then the two halves of the image would conspire to keep the whole circle as it currently is, blocking us from achieving what we sought out to do: imagine the circle a different colour.

While each half of the image is not fully about the circle's colour, it is partially about this state. I propose to analyse partial aboutness via *overlap*. The left half of the image is partially about the state of the circle's colour since it overlaps that state: the left half of circle is part of both the left half of the whole image and the circle's colour (Figure 8).

In other words, a sentence is partially about a state just in case the state contains part of what the sentence is about. When we interpret a conditional antecedent A, we allow to vary any state that the conditional antecedent is *partially* about. This produces the right result: when we imagine the circle a different colour, we allow the colour of all parts of the circle to vary but fix the colour of the background.

[22]Other semantics of conditionals that make use of such a time include Thomason (1970), Thomason and Gupta (1980), Kaufmann (2005), Ippolito (2013), Khoo (2015), and Canavotto (2020). Condoravdi (2002) calls this the *temporal perspective*, Arregui (2005) and Ippolito (2013) call it the *accessibility time*, Khoo (2015) calls it the *modal time* and Khoo (2017) calls it the *counterfactual time*.

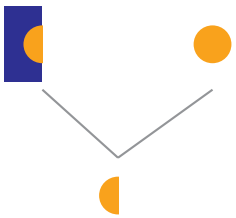


Figure 8: The left half of the image and the circle overlap at the left half of the circle.

What does it mean to fix a part of the world? It means requiring that it be part of every state we consider. For example, fixing the background of the image when we image the circle green means requiring that the background be part of the images we imagine. Allowing a part of the world to vary is simply the negation: not requiring that the background be part of what we consider.

Let us give the formal details of the account and then apply them to a worked-out example. Our model has five components: $(S, \leq, \mathcal{A}, P, V)$. I describe each in turn.

S is a *state space*, familiar from situation semantics (Barwise and Perry 1983; Kratzer 1989) and truthmaker semantics (Fine 2017c). States represent (or perhaps just are) parts of possible worlds. For example, there is the state of Alice being 25 and the state of Adonis having the appearance he has. We assume that the same state can be part of multiple possible worlds.[23] We will further take states to be instantaneous: they are snapshots representing the part of a world at an instant in time.

$\leq$ is a *parthood relation* between states. We assume it is a partial order: reflexive, transitive, and anti-symmetric. For example, the state of London is part of the state of England, the state of Alice being 25 is part of the state of her being 25 and 1.5m tall, and the state of the circle part of the state of the whole image. Two states overlap just in case they share a part.

We assume the existence of maximal states, what we call *moments*, to reflect the fact that they are maximal states at a point in time. Moments are states that are not

[23]This is in contrast to situation semantics (Barwise and Perry 1983; Kratzer 1989), where situations are taken to be particulars, each existing in no more than one world. One who wishes to use world-bound situations may introduce an exact resemblance relation between situations and recover states as equivalence classes of exactly resemblant situations.

a proper part of any other state.[24] We assume that every state is part of a moment.[25] We then take a *world* to be a set of moments extended in time.[26] A world is a linearly ordered set of moments.[27] We let M denote the set of moments and W the set of worlds.

$$M := \{t \in S : t \leq u \textit{ implies } t = u \textit{ for all } u \in S\},$$
$$W := \{(M', \preceq) : M' \subseteq M,\ \preceq \text{ is a linear order}\}.$$

$\mathcal{A}$ is an *aboutness relation* between language and the world, specifying when a sentence is about a state. For example, the aboutness relation might specify that "Alice is taller than 1 metre" is about the state of Alice's height, and "Socrates resembles Adonis" about the state of Socrates' appearance.[28]

P is a distinguished subset of worlds. Intuitively P is the set of *lawful*, or *nomically possible*, worlds: the worlds that obey the laws of nature. For example, a world where Isaac lets go of an apple and it falls will be nomically possible, while a world where the apple spontaneously turns to gold will not.

V is a valuation function, specifying for each sentence of our language and world whether the sentence is true or false at that world.

Putting the above components together formally, given a set $\mathcal{L}$ of sentences of our language, our models are of the form $(S, \leq, \mathcal{A}, P, V)$ where S is a nonempty

[24] If one wishes to countenance impossible states (such as the state of the ball being red and not red), then we must qualify this statement following Fine (2017c): introduce a distinguished set of possible states and say that moments are possible states that are not a proper part of any other possible state.

[25] In the terminology of Fine (2017a), our state spaces are *world spaces*. Note that for some applications the requirement that every state be part of a moment will not be enough, since it would permit there to be an infinite chain of states without a moment that contains the entire chain. To avoid this one must impose the stronger requirement that for every chain (i.e. totally ordered subset) of states there is a moment containing every state on the chain. For our purposes the weaker requirement that every state is part of a moment will suffice.

[26] Note that our use of linear orders is perfectly compatible with branching, familiar from models of branching time. Using linear orders, we represent branching as two worlds agreeing until some moment and then diverging.

[27] Note that a linear order allows each element to appear at most once in the order. If one wishes to allow the same moment to repeat in a world (say, if the world goes through cycles), one may index moments and take worlds to be linearly ordered sets of indexed moments.

[28] This take on aboutness, where sentences are about states, conforms to Fine's (2017; 2020) notion of subject-matter, rather than Yablo's (2014) or Berto's (2022), who represent subject matters as entities in their own right.

set, $\leq \subseteq S \times S$ a partial order where each state is part of a maximal state, $\mathcal{A} \subseteq \mathcal{L} \times S$, $P \subseteq W$, and $V : \mathcal{L} \times W \to \{0, 1\}$.[29]

Turning to truth conditions, given a sentence A and moment t, we introduce the notion of an A-variant of t, representing the moments that result from varying the part of the world the antecedent is about (Step 2). Define that moment t' is an *A-variant* of t just in case every part of t that does not overlap any state A is about is part of t': $\forall s \leq t\big((\neg\exists u \in S : \mathcal{A}(A, u) \wedge s \circ u) \Rightarrow s \leq t'\big)$, where $s \circ u$ denotes that s and u overlap (that is, $x \leq s$ and $x \leq u$ for some state x). Figure 9 illustrates the definition of A-variants.

For any world w and moment t on w, let $w_{\prec t}$ be the segment of w up to but not including t, let $w_{\succeq t}$ be the segment of w from t on, including t, and let $\frown$ denote concatenation.[30] Playing forward the laws at a moment t' amounts to considering the lawful futures of a moment: the segment $w'_{\succeq t'}$ where w' is lawful. And sticking on the past of a moment t at world w amounts to taking $w_{\prec t}$. Concatenating the actual past and possible futures gives us $w_{\prec t} \frown w'_{\succeq t'}$. The resulting set we call the *modal horizon*, following von Fintel (2001). Where A is the conditional antecedent,

[29]The reader may wonder whether how our framework compares to that of Structural Causal Models (Pearl 2000), especially since there are a number of theories of counterfactuals in terms of that framework (such as Galles and Pearl 1998; Pearl 2000; Hiddleston 2005; Schulz 2011; Briggs 2012; Kaufmann 2013; Santorio 2019). One may see Structural Causal Models as a special implementation of the framework developed here. Given a Structural Causal Model M, we may derive the state space as the set of possible assignments of values to variables, and derive the set of lawful worlds P as those worlds that follow the structural equations of M, that is, those worlds where the value of each variable at a given moment is determined by the value of its parents at the previous moment according to the structural equations, which shows that our proposal here is compatible with the Structural Causal Modelling framework.

At first blush one might think that Structural Causal Models will struggle to represent figure–ground contrasts in counterfactuals. For it is generally required that the variables in a structural causal model be logically independent (Hitchcock 2004, 145; Hitchcock 2007, 502; Woodward 2003, 128; Woodward 2016, 1053; McDonald 2025). Given this constraint, we cannot introduce separate variables representing, say, whether Socrates resembles Adonis and whether Adonis resembles Socrates, since these are logically dependent We may nonetheless introduce a variable representing Adonis' appearance and one representing Socrates' appearance. The framework of Structural Causal Models does not itself tell us on which variables to intervene when we interpret "Socrates resembles Adonis" or "Adonis resembles Socrates." To do so we require an aboutness relation between sentences and variables, telling us on which variables to intervene given a sentence. Thus my general point that we need an aboutness relation to represent figure–ground contrasts in counterfactuals still stands.

[30]Concatenation is defined as follows. Where = $(X, \preceq)$ and $(Y, \preceq')$ are linear orders, $(X, \preceq) \frown (X', \preceq') = (X \uplus Y, \preceq \frown \preceq')$ where $X \uplus Y$ is the disjoint union of X and Y, and $x \preceq \frown \preceq' y$ just in case (i) $x, y \in X$ and $x \preceq y$, or (ii) $x, y \in Y$ and $x \preceq' y$, or (iii) $x \in X$ and $y \in Y$.

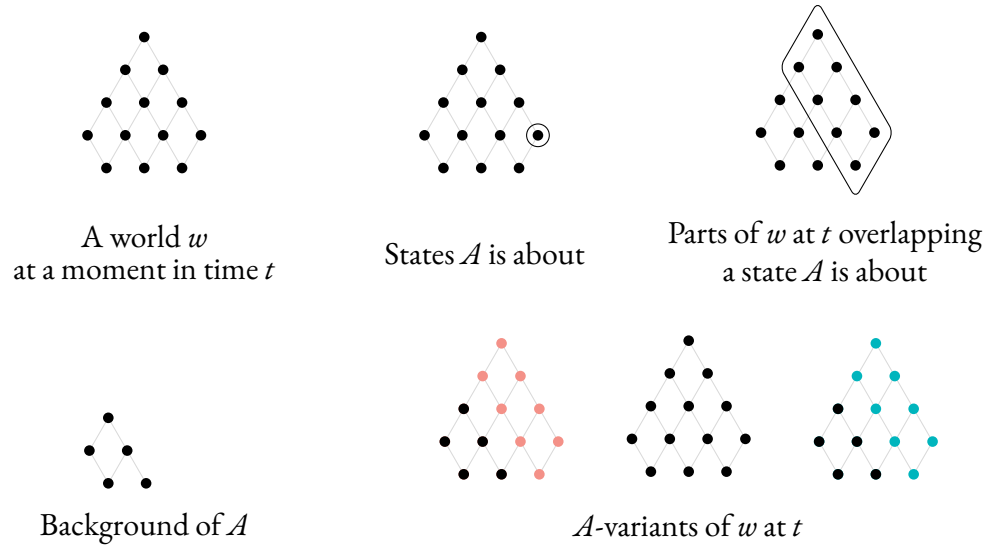


Figure 9: Steps to construct the A-variants of a world at a moment in time.

w the world of evaluation and t the intervention time, we define the modal horizon as follows, illustrated in Figure 10.

$$mh(A, w, t) = \{w_{\prec t} \frown w'_{\succeq t'} : t' \text{ is an } A\text{-variant of } t, t' \in w', \text{ and } w' \in P\}$$

We then restrict to worlds where the antecedent is true (Step 5). Where $|A| = \{w \in W : V(A, w) = 1\}$ is the set of worlds where A is true, counterfactuals are evaluated at the set $mh(A, w, t) \cap |A|$. We then state that a *would*-conditional $A > C$ is true just in case the consequent holds at every world in this set, or at a unique selected world, depending on one's favoured approach.[31]

> **The aboutness theory.** Given a model $M = (S, \leq, \frown, P, V)$, intervention time t, and selection function s, a *would*-conditional $A > C$ is true at a world w in M just in case
> $mh(A, w, t) \cap |A| \subseteq |C|$ UNIVERSAL
> $s(w, mh(A, w, t) \cap |A|) \in |C|$ SELECTIONAL

[31] A selection function $s : W \times ((W) \setminus \emptyset) \to W$ takes a world w and a nonempty set of worlds p and returns a world, such that $s(w, p) \in p$ (*success*), and if $w \in p$ then $s(w, p) = w$ (*centering*). For some work on the selectional approach to *will* and *would* conditionals, see Mandelkern (2018), Cariani and Santorio (2018), and Cariani (2021).

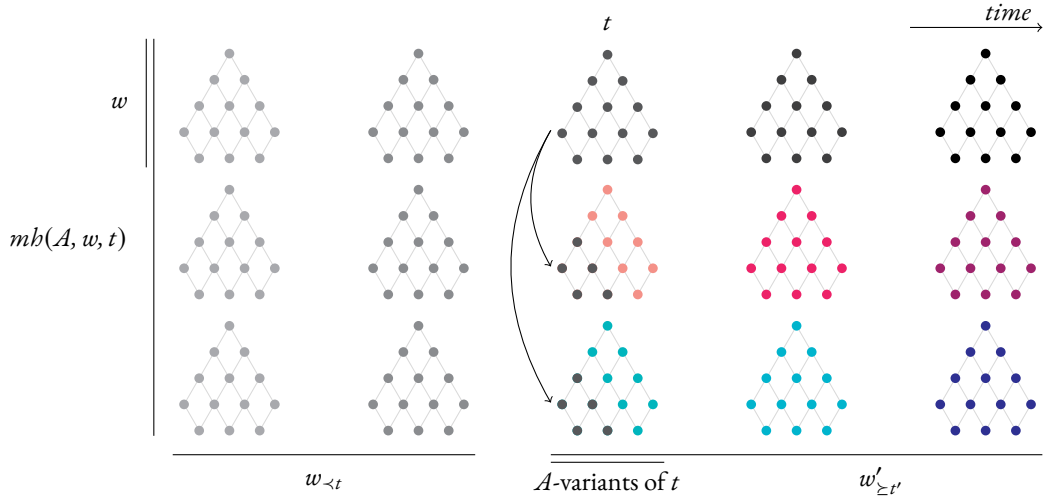


Figure 10: Constructing the modal horizon.

3.1 Illustrating the Aboutness Theory

Let's work through a simple example. Suppose two switches, A and B, are connected to a light. The light is on just in case both switches are up. Currently switch A is down and switch B is up, so the light is off. Take (19).

(19) If switch A were up, the light would be on.

(19) is straightforwardly true. Let's see how the aboutness theory predicts this. The state space consists of the possible states of each switch (up or down) and the light (on or off), and the possible combinations thereof. The current moment, together with its mereological structure, is given in Figure 11, with upward lines indicating parthood.

We assume, plausibly enough, that the sentence "Switch A is up" is about the state of switch A being up and the state of switch A being down. The sentence is therefore partially about every state that overlaps this state, such as the state of switch A being down and switch B being up. These are the figure, the other states the ground: when we interpret a conditional with antecedent "if switch A were up," we allow these states to vary and fix the rest (see Figure 12).

We can put the point in terms of state subtraction—a notion that goes back to the very beginnings of formal mereology with Stanisław Leśniewski's work in the 1920s. The background results from removing the state of switch A being down from the current moment. This is illustrated in Figure 13, where $-$ is Leśniewski's operation of *relative complement*: for any states s and t, $s - t$ is the largest part of s

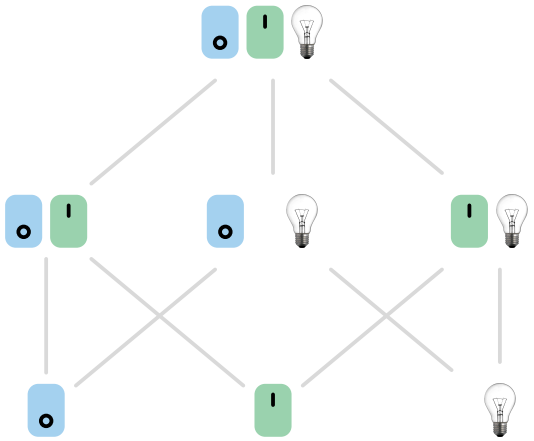


Figure 11: The current moment and its parts.

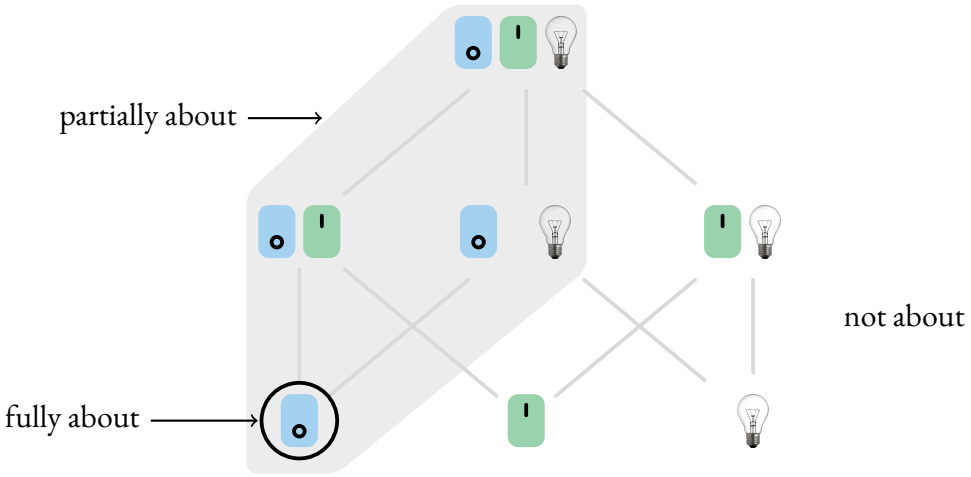


Figure 12: Aboutness relations for "Switch A is up."

that does not overlap *t* (Leśniewski 1931, see also Simons 1987, 14). The largest part of the actual moment that does not overlap switch A being down is the state of switch B being up and the light being off.

Thus when we interpret "if switch A were up," we fix the states of switch B being up and the light off. This background is part of two moments: one where switch A is up, and one where switch A is down. Step 5 ("Restrict to worlds where the antecedent holds") rules out the latter, leaving only the former moment, where switch A is up, switch B is up, and the light is off.[32]

[32]We assume that the tense information required to interpret the sentences under consideration has already been supplied. Thus the sentence "switch A is up" is evaluated with respect to a particular

Figure 13: The division of Figure and Ground, in terms of mereological subtraction.

In this way the state space and aboutness relation allow us to model the operation of editing a part of the world at a moment in time—in this case editing the position of switch A. However, this operation does not by itself predict the conditional "If switch A were up, the light would be on" to be true. A change to the state of switch A is not by itself a change to the light. For this we look to the laws. Recall that the light is on just in case both switches are up. This law, like so many dynamical laws in physics, is not assumed to hold at every point in time but only at equilibrium.[33] It is perfectly possible for the world to be in a state where both switches are up and yet the light is off, albeit only temporarily, when the light is yet to change. To keep the illustration simple let us assume discrete time and that each change takes exactly one moment. The system is then governed by the following law: *A world w is lawful just in case for any moments m and m' of w where m' is the direct successor of m, the light is on at m' just in case both switches are up at m.* This determines the set of lawful worlds P. For instance, a world where the switches are flicked up and the light turns on is lawful (Figure 14).

Figure 14: A lawful world.

Whereas a world where the light spontaneously flickers on and off, as if in a thriller film, is unlawful (Figure 15).

time, hence restricting to worlds where that sentence is true means restricting to worlds where the sentence is true at the relevant time of evaluation. Note that here we are not taking a stand on the more controversial question of whether the propositions expressed by sentences are eternal (have the same truth value at all times) or temporal (have different truth values at different times). On this latter issue see Brogaard (2011).

[33] For example, concerning the ideal gas law Kline writes, "the gas laws are intended as descriptions of gases in equilibrium. They are not to be taken as describing processes. ... Once a change in the volume occurs it is some time until the gas reaches equilibrium and is correctly described by [the ideal gas law, $PV = NrT$]" (Kline 1980, 296–97).

Figure 15: An unlawful world.

Our operation of editing the current moment returned one where both switches are up and the light is off. Step 3 of the aboutness theory then tells us to play forward the laws. That is, we consider the lawful futures of this moment. Given the laws of the current scenario, every such future is one where the light turns on. In this way the aboutness theory correctly predicts (19), "If switch A were up, the light would be on" to be true. There is much more to explore about the aboutness theory beyond this simple example, but this discussion should help the reader appreciate how it makes predictions for run-of-the-mill counterfactuals.

3.2 The Subject Constraint

The aboutness theory developed here is ideally suited to account for figure–ground contrasts in counterfactuals. Recall, for example, (4).

(4) a. If Socrates resembled Adonis, they would both be handsome.
b. If Adonis resembled Socrates, they would both be handsome.
c. If Adonis and Socrates resembled each other, they would both be handsome.

To predict the intuitive judgements in this case, where the first is true, the second false, and the third indeterminate, the aboutness relation must specify that "Socrates resembles Adonis" is about Socrates' appearance and not Adonis', "Adonis resembles Socrates" is about Adonis' appearance and not Socrates', and "Adonis and Socrates resemble each other" is about both.

There is a tantalising generalisation underlying many of the figure–ground contrasts we have observed, including those discussed by Russell and Talmy: sentences are about their subjects.[34] We will qualify this as a default, rather than a universal generalisation, for reasons discussed below.

The Subject Constraint. By default, sentences are only about their

[34] By 'subject' here I mean the grammatical subject. For a discussion of the notion of grammatical subject in English, see Conner (1968, 43ff.) For a list of criteria to identify subjects, see Biber et al. (2021, 127).

subjects.
If x is the subject of sentence A, then A is about a state s only if s is part of the state of x.

In other words, the Subject Constraint says that by default, two notions of subject—subject matter and grammatical subject—coincide.[35]

The Subject Constraint forces the aboutness relation to be hyperintensional. Logically equivalent sentences may have different subjects, so by the Subject Constraint will be about different states by default.[36] In this way counterfactual antecedents are hyperintensional environments according to the aboutness theory—Substitution comes out invalid.

The subject sets a bound on what the sentence is about. We implement this in terms of parthood: if a sentence A is about a state s, s is part of the state of A's subject. In "Socrates resembles Adonis," say, Socrates is the subject and Adonis the object.[37] The Subject Constraint then tells us that, by default, every state the sentence is about is part of the state of Socrates. Naturally, the sentence is not about every part of Socrates whatsoever—the Subject Constraint is an *only if* condition. Many parts (such as his place of birth or taste in wine) will remain unchanged.

---

[35]The grammatical subject of a sentence is distinct from what is often called a sentence's *topic* (Reinhart 1981) or, in the Prague School of functional linguistics, its *theme* (Firbas 1964; Mathesius 1975). We can isolate the topic using terms such as *speaking of*, *as for*, and *talking about* (Reinhart 1981), as in (i) and (ii).

(i) a. Talking about Socrates, if he resembled Adonis, they would both be handsome.
b. Talking about Adonis, if Socrates resembled him, they would both be handsome.

(ii) a. Talking about Socrates, if Adonis resembled him, they would both be handsome.
b. Talking about Adonis, if he resembled Socrates, they would both be handsome.

Forcing the topic in this way does not alter the sentences' truth value. Intuitively, both sentences in (i) are true, while both sentences in (ii) are false. While the topic and grammatical subject are distinct notions, it is nonetheless true that the grammatical subject will often happen to be the topic. According to Reinhart, "There is a strong preference in discourse to interpret the grammatical subject of the sentence as its topic, or to place the topic in subject position. [...] this preference is only a matter of tendency and we can use sentences with non-subjects as topics" (Reinhart 1981, 62).

[36]This contrasts with intensional analyses of aboutness (such as Putman 1958 and Goodman 1972), according to which logically equivalent sentences are about the same things.

[37]There is an ambiguity in whether the subject of a sentence is the linguistic material itself or its denotation; for instance, whether the subject of "Socrates resembles Adonis" is the name 'Socrates' or Socrates himself. We adopt the latter.

"Adonis and Socrates resemble each other" has a compound subject: Adonis and Socrates. Following Link (1983), we assume that plural terms denote the mereological sum (or fusion) or their constituents. Thus "Adonis and Socrates" denotes the sum of Adonis and Socrates. The Subject Constraint requires that by default, any state "Adonis and Socrates resemble each other" is about is part of the state of this sum. The Subject Constraint therefore permits that this sentence be about both of their appearances.

Turning to quantifiers, such as "All ravens are black," Yablo (2014, 2, 19–22) proposes that this sentence is about the ravens, while "All non-black things are non-ravens" is not. More generally, Matthewson (2001) argues that quantified expressions such as "all ravens" denote the plurality of objects with the property. And Ben-Yami (2012) writes that "In a sentence of the form '*q S* are *P*' the subject term *S* is used as a plural referring expression, determining the plurality over which we quantify." We follow this work, proposing that the subject of a quantified expression is the plurality of things that satisfy the restrictor. Then according to the Subject Constraint, by default any state that *All/most/some/no ravens are black* is about is part of the state of the ravens.

Paired with the aboutness theory, the Subject Constraint predicts the figure–ground contrasts we have observed. Recall (4), repeated below.

(4) a. If Socrates resembled Adonis, they would both be handsome.
b. If Adonis resembled Socrates, they would both be handsome.
c. If Adonis and Socrates resembled each other, they would both be handsome.

The Subject Constraint tells us that by default, the first antecedent is not about Adonis, the second not about Socrates. We further assume, plausibly enough, that the first is about Socrates, the second about Adonis, and the third about both. Then on the aboutness theory, when we interpret the first we fix the state of Adonis (including his appearance), when we interpret the second we fix the state of Socrates, and when we interpret the third we fix neither. Thus given the first we imagine both handsome, given the second we imagine both unattractive, and given the third we leave open whether they are both attractive or both handsome. These are intuitively the correct results.

3.3 The Default Status of the Subject Contraint

There is a potential worry one may have about the Subject Constraint. In many cases, imagining a counterfactual antecedent true requires changing more than just the subject. For instance:

(20) If Sarah had opened the window, the room would have become cooler.

Here Sarah is the subject, but we also imagine a change in the state of the window. A plausible response is that "Sarah opened the window" is indeed only about Sarah, and the change in the window comes about via the laws rather than the aboutness relation. For the intervention time need not be the exact moment when Sarah opens the window. It could be some prior moment, with the change at intervention time being to Sarah's mental state (say, her desire or intention to open the window). This is a general fact about the intervention time: it is often not the event time.

This issue has come up before. Bennett (1984) once proposed—and eventually rejected, see Bennett (2003)—the view that counterfactuals require a departure from actuality at the exact time when the antecedent holds.[38] Goggans (1992, 145–46) replied with the following counterexample.

(21) If Bush were in London right now, his staff would still be in Washington.

When we interpret (21), we tend not to imagine Bush spontaneously appearing in London, but a small change beforehand (the beginning of the 'ramp,' in Bennett's 2003 phrase), such as a change in his schedule, travel plans, and staff arrangments. We set the intervention time to before Bush is supposed to be in London. Similarly in (20), we may set the intervention time to some time before Sarah would have opened the window, in which case we can preserve the idea that the antecedent is purely about its subject at intervention time.

Nonetheless, there are other cases where this move is less plausible. Recall the blocks from Figure 1, where no two blocks can occupy the same place. Consider (22).

(22) If block A were in block B's place, block B would not be where it currently

[38] Bennett was clear that when we only evaluate similarity at the time when the antecedent holds, writing, "What must not be done is to bring in facts about the actual world at times other than T [the time to which the antecedent pertains]. If T is the present, then we must put away our history books and crystal balls, using only our eyes and our capacity for causal inference in both temporal directions" (Bennett 1984, 74).

is.

The subject is block A, but moving block A into B's place also requires moving B. One may argue that even in this case there is an earlier intervention time, whereby we only imagine changing the position of block A at intervention time, and then the laws require block B to move thereafter. However, it is also plausible to model this as a change to both blocks A and B at intervention time. In that case (22) would have to be about both blocks, overriding the Subject Constraint's default.

A further case illustrating that the Subject Constraint is a mere default comes from dummy subjects (such as *it* and *there*).

(23) a. If it were necessary to call in advance, I would do so.
b. If there were a coin in my pocket, I would spend it.

It is often assumed that dummy subjects are semantically empty: they do not denote anything in particular (see Haider 2019). Strictly enforcing the rule that sentences are only about their subjects would imply that (23) are not about anything, and therefore on the aboutness theory that we do not change anything when we interpret them—an undesirable result.[39]

A further case is the passive.[40]

(24) a. If Alice hugged Bob, he would feel butterflies in his stomach.
b. If Bob were hugged by Alice, he would feel butterflies in his stomach.

These are intuitively equivalent. In "Alice hugged Bob," Alice is the subject, while in "Bob was hugged by Alice," Bob is the subject. If sentences were always only about their subjects, then on the aboutness theory the second conditional would only allow us to vary Bob's state, keeping Alice's state the same. This is an incorrect result, since imagining Bob being hugged by Alice also requires changing the state of Alice (her position, movements, and so on).

A final case showing the default status of the Subject Constraint is when the subject is something we are disinclined to imagine changing, such as a constant. Suppose Alice is 17 and consider (25).

---

[39] An alternative stance is that dummy subjects refer to a general, loosely specified situation. Given this response, dummy subjects would no longer challenge the idea that sentences are only about their subjects.

[40] I am grateful to Ivano Ciardelli for raising this point.

(25) If 16 were the legal drinking age, Alice could order wine.

If we strictly required that sentences be only about their subjects, "16 is the legal drinking age" would be about the number 16. The aboutness theory would then say that when we interpret (25) we should vary the state of the number 16 and fix the legal drinking age. However, the most plausible interpretation in this case is to modify the legal drinking age—changing the object rather than the subject.

4 The *De Re* Reply

In this section we develop, and ultimately argue against, an alternative response to the figure–ground contrasts we have seen in this essay, that the comparator is interpreted *de re*, with its value anchored to the actual world. We call this response the *de re* reply.

Aristotle observes that "It is possible to walk while sitting" has both a consistent and a contradictory reading (*Sophistical Refutations* 166a24–166a30; see Dutilh Novaes 2003). Today we would label the consistent reading as *de re*, and analyse it as saying that there is a possible world w where someone who is sitting in the actual world w_0 is walking in w. The contradictory reading is the *de dicto* one: there is a possible world w where someone who is sitting in w is walking in w.

De re/*de dicto* ambiguities also appear in comparatives. In 'On Denoting' Russell recalls the following anecdote.

> I have heard of a touchy owner of a yacht to whom a guest, on first seeing it, remarked, "I thought your yacht was larger than it is"; and the owner replied, "No, my yacht is not larger than it is." What the guest meant was, "The size that I thought your yacht was is greater than the size your yacht is"; the meaning attributed to him is, "I thought the size of your yacht was greater than the size of your yacht."
>
> (Russell 1905, 489)

The same ambiguity appears in conditional antecedents (Lewis 1973a, 37).

(26) If your yacht were larger than it is, you would boast about it more.

This has a consistent reading, in which we imagine the yacht being larger in the hypothetical worlds than it is in the actual world.

We observe the same behaviour in quantifier restrictors (Percus 2000).

(27) a. Alice thinks that everyone inside the room is outside the room.
b. If everyone inside the room were outside the room, it would be empty.

The belief claim and antecedent have a consistent reading, saying that everyone who is inside the room *in the actual world* is outside the room in the worlds, respectively, compatible with Alice's beliefs and raised by the conditional antecedent.

Lewis (1973a, 37) represents the different readings via quantifier scope. He translates *If your yacht were larger than it is, ...* as *The size of your yacht is an x such that, if the size of your yacht exceeded x, ...*. This results in the right reading for Russell's example but not for (27b). This sentence does not mean that *Everyone in the room is an x such that, if x were outside the room, it would be empty*. On this reading we remove one person at a time and the sentence is false, whereas when we interpret (27b) we remove them all at once and the sentence is true.

We can instead represent *de re–de dicto* readings using world indices (Percus 2000; Keshet 2010). Assuming a simple, Hintikka (1962) style semantics of belief for expository purposes, we can represent the two readings as follows, where w_0 is the world of evaluation.

(28) I thought your yacht was larger than it is.
a. **De dicto.** For every world w compatible with my beliefs at w_0, the size of the yacht at w is greater than the size of the yacht at w.
b. **De re.** For every world w compatible with my beliefs at w_0, the size of the yacht at w is greater than the size of the yacht at w_0.

(29) Alice thinks that everyone inside the room is outside the room.
a. **De dicto.** For every world w compatible with Alice's beliefs at w_0 and for all x such that x is inside the room at w, x is outside the room at w.
b. **De re.** For every world w compatible with Alice's beliefs at w_0 and for all x such that x is inside the room at w_0, x is outside the room at w.

One may argue that the figure–ground contrasts we have considered here exhibit a similar ambiguity: "Socrates resembles Adonis" means that Socrates' appearance at w is sufficiently close to Adonis' appearance at w (*de dicto*) / w_0 (*de re*), that "Alice is younger than Bob" means that Alice's age at w is less than Bob's age at w (*de dicto*) / w_0 (*de re*), that "Alice is near Bob" means that Alice's location at w is sufficiently close to Bob's location at w (*de dicto*) / w_0 (*de re*), and so on.

This *de re* reading appears to be available. Suppose Alice believes that Bob resembles the actor Remi Malek. Alice isn't aware that Remi has an identical twin

brother Sami Malek—she has no idea that Sami exists.

(30) Alice believes that Bob resembles Sami Malek.

I submit that, in the right context, (30) may have a true reading.

Or suppose Suzy can see the cat but not the dog (Figure 16). She has no idea that the dog is there. Actually, the cat is near the dog. Consider:

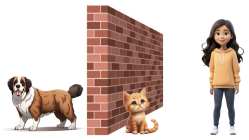


Figure 16

(31) Suzy believes that the cat is near the dog.

This is most naturally read as false—after all, Suzy has no beliefs whatsoever about the dog. But I submit that it may also have a true reading in this context, saying that Suzy believes that the cat is near the location where the dog actually happens to be. On this *de re* reading, (31) is true.

The *de re* reply asserts that *de re* readings are responsible for the contrasts we have observed. Recall, for example, (5).

(5) *Alice is 25, Bob is 15. One must be over 21 to enter the bar.*

a. If Alice were younger than Bob, they could both enter the bar.
b. If Bob were older than Alice, they could both enter the bar.

On the *de re* reading these are equivalent, respectively, to (32).

(32) a. If Alice were younger than 15, they could both enter the bar.
b. If Bob were older than 25, they could both enter the bar.

The antecedents in (5) are equivalent *de dicto* but not *de re*. Read *de re*, they are not counterexamples to Substitution and can easily be handled by extant semantics of counterfactuals, such as Stalnaker's or Lewis's.

In what follows I explain why I believe the *de re* reply to be implausible, and cannot account for the full range of data under consideration.

4.1 Against the *De Re* Reply

It is generally assumed that *de dicto* and *de re* readings are always in principle available (Percus 2000; Keshet and Schwarz 2019). We use pragmatic reasoning, contextual features and our general world knowledge to disambiguate. This has the attractive feature of making the resolution of *de re*–*de dicto* ambiguity of a kind with the resolution of any other ambiguity, such as referential ambiguity (for instance, determining

who 'they' refers to in a given context) and polysemy (disambiguating riverbanks from financial banks). Grammar supplies both as interpretative possibilities, and context disambiguates.

A pervasive feature of ambiguity resolution is that we prefer resolutions on which the sentence is true to those on which it is false. In Russell's yacht example, "I thought your yacht was larger than it is," we opt for the *de re* reading. It is easy to understand why: the *de dicto* reading is clearly false. We can account for this behaviour by a principle of interpretative charity.[41]

Principle of Charity. When multiple readings of a sentence are available, some of which are true, we opt for one of the true readings.

To illustrate, suppose you see someone you know searching through a crowd. When you ask what they are doing here they reply, "I'm looking for a friend." Here we easily interpret them as saying that there is a particular person who is their friend whom they seek. This is the *de re* reading.

Now suppose you meet someone at a bar who tells you they are new to the neighbourhood and don't know anyone yet. When you ask what they are doing here they reply, "I'm looking for a friend." Now we naturally interpret them as saying that they are looking to have a friend. There is no particular person for whom they are looking. For all we know they might not even have any friends. This is the *de dicto* reading.

We easily manage to pick the reading that makes the sentence true. If a *de re*–*de dicto* ambiguity were at play in the counterexamples to Substitution we have considered, we would likewise expect each reading to be available, and that when the sentence is true on one reading and false on the other, we opt for the true reading.

To test this, consider the following scenario. Suzy is playing a game in which she is given two cards, A and B, at random. Each card shows either 1, 2, or 3. The draws are independent (for example, we may imagine that they are drawn from separate decks). Suzy wins just in case card A is 3 and card B is 2. On this particular occasion, card A is 2 and card B is 2, so Suzy lost (see Figure 17). Consider (33).

---

[41] The Principle of Charity I have in mind here is of a piece with the principle that appears in the work of Quine and Davidson. Wilson, who coined the term, wrote that "we act on what might be called the Principle of Charity. We select as designatum that individual which will make the largest possible number of [...] statements true" (Wilson 1959, 532). As Davidson puts it, the task of solving for meaning is aided by "assigning truth conditions to alien sentences that make native speakers right when plausibly possible" (Davidson 1973, 324; for discussion see Glüer 2011).

Figure 17

(33) If card B were lower than card A, there is a 50% chance that Suzy would have won.

(33) is clearly false on the *de re* reading. On this reading "B is lower than A" means that B is lower than the actual value of A, which is 2. And if card B were lower than 2, there is a 0% chance that Suzy would have won.

What about the *de dicto* reading? The predictions here depend on one's semantics. Let us test this on Lewis's. It is plausible to assume that changing one card requires less of a departure from reality than changing two. Evidence for this comes from sentences such as (34).

(34) If card A were 3, Suzy would have won.

This is intuitively true, showing that when we change A we leave B as is. (We can construct similar sentences under other winning conditions to show this for all possible values of cards A and B.) This generates the following similarity order: the closest world to the actual world is the actual world itself—(2, 2)—the next closest worlds, all equally close to the actual world, are those where one card changes and the other is still 2—(1, 2), (3, 2), (2, 1), (2, 3)—and the next closest worlds, again all equally close to the actual world, are those where both cards change—(1, 1), (1, 3), (3, 1), (3, 3). Given this similarity order, the most similar worlds to the actual world where card B is lower than card A are (3, 2) and (2, 1), given in bold.

$$(2, 2) \;<_w\; (1, 2), (\mathbf{3}, \mathbf{2}), (\mathbf{2}, \mathbf{1}), (2, 3) \;<_w\; (1, 1), (1, 3), (3, 1), (3, 3)$$

One of these, (3, 2), is the winning hand. Since the cards were given randomly, each outcome has an equal chance of occurring, so among these possibilities there is a 50% chance that Suzy would have won. Thus on Lewis's semantics (33) is true on its *de dicto* reading.

There is nothing especially complicated about entertaining this similarity order, nor evaluating conditionals as complex as (33). Take, for instance, (35), in the same context given in Figure 17.

(35) If card A were odd, there is a 50% chance that Suzy would have won.

(35) is readily judged true, exactly as this similarity order predicts.

For a second example, consider (36).

(36) If card A were higher than card B, there's a 50% chance that Suzy would have won.

This is also false on the *de re* reading, on which "A is higher than B" means that A is higher than 2. Intuitively, if A were higher than 2, there is a 100% chance that Suzy would have won.[42] On the *de dicto* reading, however, (36) is equivalent to (33) according to theories that validate Substitution. Lewis therefore predicts (36), like (33), to be true on its *de dicto* reading.

The trouble for Lewis is that (33) and (36) are clearly false. They do not have any true reading. Furthermore, there is no reading on which they are equivalent.

No such difficulties arise on the aboutness theory. It predicts that (33) and (36) are false on both their *de re* and *de dicto* readings (see Table 1).

Predictions for (33) and (36)	*De re*	*De dicto*
Similarity theory	✗	✓
Aboutness theory	✗	✗

Table 1

Let's first derive these predictions on the *de re* readings. The antecedent of (33) on its *de re* reading is "if card B were lower than 2." This is about card B and not about card A, so we vary the value of card B and fix the value of card A. When we restrict to those cases where B is lower than 2, we have that A is 2 and B is 1, in which case Suzy has no chance of winning. Similarly, the antecedent of (36) on its *de re* reading is "if card A were higher than 2," which is about A and not B. When we allow A to vary but fix B, and restrict to worlds where A is higher than 2, we find that A is 3 and B is 2, in which case Suzy certainly wins.

The aboutness theory makes the same predictions for the *de dicto* readings. The antecedent "if card B were lower than card A" is about its subject, card B, and not about its object, card A. So we allow the value of card B to vary and fix the value of

[42] One might also consider the possibility that we vary B simultaneously here. In that case, if A were higher than 2 and B were allowed to vary, there would be three possibilities—(3, 1), (3, 2), and (3, 3)—and therefore a 1 in 3 chance that Suzy would have won. On this reading (36) still comes out false.

card A. When we restrict to the resulting worlds where card B is lower than card A (read *de dicto*), we find that A is 2 and B is 1, in which Suzy has zero chance of winning. Similarly, the antecedent "if card A were higher than card B" is about card A and not about card B, so we allow A to vary and fix B. When we restrict to the resulting worlds where A is higher than B (read *de dicto*), we find that A is 3 and B is 2, in which case Suzy certainly wins.

To summarise this section, the received view of the *de re–de dicto* ambiguity is that, just as with any ambiguity, both readings are in principle available. By a principle of interpretative charity, when one reading is false and the other true, we opt for the true reading. As we have seen the similarity approach predicts that (33) and (36) are true on their *de dicto* reading. However, they do not have any true reading.[43]

The similarity approach has a possible response: assert that the *de re* reading is obligatory. If we rule out the *de dicto* reading entirely, we account for the intuitive falsity of (33) and (36). This response goes against what we know in general about the availability of *de dicto* readings. However, let us be charitable and entertain it. In what follows we show that the *de dicto* readings are in fact available. Thus theories that validate Substitution face an uncomfortable dilemma. On the one hand, to account for (33) and (36) the *de dicto* readings must not be available, while in other contexts they must be.

4.2 The Availability of *de Dicto* Readings

In this section we present three environments which clearly show that the *de dicto* readings are available for the antecedents we have considered.

Conjunctive antecedents.

(37) If Adonis resembled Socrates and Socrates were handsome, they would both be handsome.

(37) is most naturally judged true. It is true on the *de dicto* reading of "Adonis resembles Socrates" but false on the *de re* reading. Read *de dicto*, the antecedent is true at a world w just in case Adonis' appearance at w is sufficiently like Socrates' appearance at w and Socrates is handsome at w. In this case both are handsome

[43] While we have focused our objection on comparatives such as *lower than*, one can make the same argument for the other environments we have considered: spatial comparisons (*near*, *on top of*), non-spatial relations (*resembles* and *is similar to*), and quantifiers (*every*, *no*, *some*).

and we therefore judge the conditional true. Read *de re*, the antecedent is true at a world w just in case Adonis' appearance at w is sufficiently like Socrates' appearance at the *actual world* w_0—in which Socrates is not handsome—and Socrates is handsome at w. When we read (37) in this way, we imagine Adonis unattractive and Socrates handsome, and so judge the conditional false. The fact that (37) has a true reading shows that "Adonis resembles Socrates" has a *de dicto* reading, even inside a conditional antecedent. Proposing a general prohibition on the *de dicto* reading of "Adonis resembles Socrates" in conditional antecedents would wrongly predict (37) to be false.

The same goes for the cards example from the previous section, where Suzy wins just in case A is 3 and B is 2. Suppose that this time cards A and B are both 1 (Figure 18).

Figure 18

(38) If card A were higher than card B and card B were 2, Suzy would have won.

This is intuitively true. It is true *de dicto* but false *de re*. If card A were higher than the actual value of card B (which is 1), and card B were 2, Suzy would have lost.

Recall from the previous section that the *de re* reply required the *de re* readings to be obligatory to make the correct predictions—even when the *de dicto* reading would rescue the sentence from falsity. Yet here the *de dicto* readings are clearly available. The same observations apply to the other contrasts we have considered; for example:

(39) *Alice and Bob are both 15. One must be 21 or over to enter the bar.*
If Bob were older than Alice and Alice were 21, they could both enter the bar.

(40) *Alice is at home and Bob is outside in the rain.*
If Bob were near Alice and Alice were outside, they would both be wet.

We can also test the sentences from the previous section that the *de re* reply required to be read *de re*.

(41) *Cards A and B are both 2. Each card can be 1, 2, or 3. Suzy wins just in case card A is* 3 *and card B is 2.*
If card B were lower than card A and card A were 3, there is a 50% chance that Suzy would have won.

This has a salient true reading. For the same reasons as in (37), this is true *de dicto* and false *de re*. On the latter, the antecedent says that card B is lower than the actual value of card A (which is 2), and card A is three. In that case card B would be 1 and Suzy would have zero chance of winning.

The *de re* reply is required to say that when we interpret "If card B were lower than card A" by itself, the *de dicto* reading is unavailable—even when it would save the conditional from falsity—while when we add the conjunct "...and card A were 3," it suddenly becomes available again. Such moves are extremely stipulative and unsatisfying. We should be deeply suspicious of any approach that requires us to perform such dazzling feats of *de re–de dicto* gymnastics.

One might try to rescue a general prohibition on *de dicto* readings in conditional antecedents as follows: "Adonis resembles Socrates" is indeed read *de re* in (37), but it is not anchored to the actual world. Instead, it is anchored to the world where we imagine Socrates handsome. Let w_S be the world we consider when we imagine Socrates handsome. (For simplicity we assume this is a single world. Since there are many ways for Socrates' to be handsome, plausibly this should be a set of worlds, and we would represent *de re* readings using *sets* of world indices rather than single world indices.) The thought is that we interpret (37) as follows.

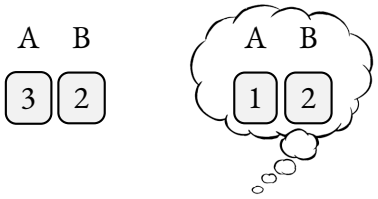
(42) If Adonis' appearance were similar to Socrates' appearance at w_S, and Socrates were handsome, they would both be handsome.

While this makes the right prediction for (37), it is extremely *ad hoc*. It is unclear what principles govern which world indices are available. It is also unclear how we should replace world indices with sets of world indices to account for cases where the antecedent raises multiple possibilities, and how we should rewrite our general understanding of the *de re–de dicto* ambiguity to accommodate this change.

Moreover, this reading does not respect the linear order of information given by the antecedent. The information that Adonis resembles Socrates comes before the information that Socrates is handsome. One would need to say that the meaning of "Adonis resembles Socrates" shifts, cataphorically, in just the right way once we encounter "Socrates is handsome"; namely, by shifting the *de re* anchor. While this is in principle a possible move, it is highly implausible.

The lesson I draw from this is that theories validating Substitution must introduce convoluted bells and whistles to account for these data, raising unanswered (and perhaps unanswerable) questions about what rules govern when *de re* readings are permitted and when they are obligatory. In contrast, the aboutness theory handles these complex antecedents straightforwardly. Unlike the similarity theory, the aboutness theory never needed to rule out the *de dicto* readings to make the right predictions. For the aboutness theory predicts (37) to be true on its *de dicto* reading. According to it, "Socrates resembles Adonis" and "Adonis resembles Socrates" behave differently as conditional antecedents on both their *de re* and *de dicto* readings. Read *de dicto*, they are logically equivalent, but on to the aboutness theory counterfactuals are also sensitive to what their antecedents are about.

Belief contexts. The *de dicto* readings that the *de re* reply claims to be unavailable are also readily available in belief contexts. Take the antecedent of (33), "Card B is lower than card A." Suppose that Suzy has mistaken beliefs about what cards she has (Figure 19). In reality card A is 3 and card B is 3, but Suzy believes that card A is 1 and card B is 2. Consider (43).



Reality Suzy's beliefs

Figure 19

(43) Suzy believes that card B is lower than card A.

This is most naturally read as false. Suzy believes that card B is 2 and card A is 1, and so believes that card B is higher than card A. One the *de dicto* reading, (43) says that Suzy believes that card B is lower than the actual value of card A, which is 3. And indeed, Suzy does believe that card B is lower than 3.

The fact that we by default judge (43) false shows the false *de dicto* reading to be clearly accessible; in fact, it is the most salient reading. This is despite the fact that opting for the *de re* reading, which is true, would rescue the sentence from falsity.

Recall that for the similarity approach to make the correct prediction for (33) it requires the *de dicto* reading to be unavailable in the conditional antecedent, "if card B were lower than card A." Yet here we see the *de dicto* reading clearly on show in a belief report. The challenge for the similarity approach is to state a principled reason why the *de re* reading is seemingly impossible in some intensional environments (conditional antecedents) but the most salient reading in others (such as belief reports).

Or suppose blocks A, B, and C are arranged with A above B and B above C. Suzy believes that the blocks are at the same location, but mistakenly believes that C is above A and A above B (Figure 20). Consider (44).

Figure 20

(44) Suzy believes that block A is above block C.

(44) easily reads as false. It is indeed false on a *de dicto* reading. However, it is true on the *de re* reading, on which it means that Suzy believes that block A is above where block C actually happens to be; namely, the bottom position.

This presents a challenge to the *de re* reply, which proposed that comparators must be interpreted *de re*. The judgements we have observed point in exactly the opposite direction: in belief reports comparators are by default interpreted *de dicto*.[44] In many cases the *de re* interpretation is extremely hard to access, even when it would rescue the sentence from falsity.

In contrast, these judgements are perfectly compatible with the aboutness theory. That approach does not rely on *de re* readings to make the right predictions for conditional antecedents, instead making use of aboutness, an independent component of meaning. Its prediction that, say, *A is above B* and *B is below A* behave differently in conditional antecedents is more robust, holding on both the *de re* and *de dicto* readings (Table 1).

5 Conclusion

The figure–ground distinction, studied by the Gestalt psychologists, Russell and Talmy, is fundamental to how we structure information around us. A hitherto unnoticed fact about the distinction is that it gives rise to striking contrasts in counterfactuals. The contrasts show up in a wide range of environments, including comparatives (*taller than*), locatives (*near*, *adjacent to*, *on top of*), transitive verbs (*resembles*, *is similar to*), and quantifiers (*every*, *some*, *no*). As Talmy noticed, there are logically equivalent sentences that nonetheless differ in their division of figure

[44]The difficulty in accessing *de re* readings here aligns with recent experimental work on the topic. Zhang and Davidson (2021, 2024) tested the availability of *de re* and *de dicto* readings in belief reports. They found that *de dicto* readings were robustly accepted by nearly all speakers, while speakers disagreed over the availability of *de re* readings: "While judgments for *de dicto* readings overwhelmingly aggregate toward the 'highly agree' end, judgments for *de re* readings are bimodal—although more than half of the judgments are agreed with, another sizable proportion goes to the 'highly disagree' edge" (Zhang and Davidson 2021, 6).

and ground: it is a widespread source of hyperintensionality in language, operating along a dimension of meaning beyond truth conditions.

This is a problem for theories of counterfactuals that validate Substitution, such as Stalnakers' and Lewis's, which are committed to the idea that logically equivalent sentences are intersubstitutable salva veritate in conditional antecedents. Indeed, it is also a challenge to the weaker principle of Exact Substitution and Fine's semantics of counterfactuals, despite the remarkable fine grain of that approach. For it is plausible that the sentences exhibiting the figure–ground not only have the same truth conditions, but also the same truthmakers. This is supported by Fine's own arguments, as well as independent arguments from the theory of reciprocals.

The aboutness theory is ideally placed to account for figure–ground contrasts in counterfactuals. It holds that when we interpret a counterfactual antecedent we distinguish foreground from background: the facts we allow to vary from those we hold fixed. This lines up neatly with how Talmy think of the distinction, with the figure "a moving or conceptually moveable point" and the ground having "a stationary setting." Generalising from Russell and Talmy's examples, we proposed the Subject Constraint: by default, sentences are only about their subjects. Paired with the aboutness theory, this constraint immediately makes the right predictions for figure–ground contrasts in conditionals. For example, it predicts that when we interpret "if Socrates resembled Adonis," we allow Socrates' appearance to vary and fix Adonis', while this is reversed with "if Adonis resembled Socrates," and given "if Socrates and Adonis resembled each other" we allow the appearance of both to vary.

There is a response available to proponents of (Exact) Substitution: the *de re* reply. It alleges that comparators are anchored to their values in the actual world; for instance, that "if Socrates resembled Adonis" is read as "if Socrates' appearance resembled Adonis' appearance in the actual world." While this results in plausible truth conditions, it fails to predict the correct range of attested readings. Specifically, it predicts there to be true *de dicto* readings where in fact we observe none. Ruling out the *de dicto* readings entirely does not account for the behaviour of complex conditional antecedents and belief reports. It moreover violates what we know about the *de re*/*de dicto* ambiguity in general, and the principle of charity in particular. In contrast, on the aboutness theory these difficulties do not arise, since it accounts for figure–ground contrasts regardless of whether the antecedents in question are read *de dicto* or *de re*.

There are many questions to ask about the aboutness theory. How we decide in general what parts of the world sentences are about? What other factors, beyond

the sentence's subject, determine the aboutness relation? One may appeal to truthmakers here (Fine 2017c), arguing that a sentence A is about state s just in case s exactly verifies or falsifies A, with the additional default constraint that s be part of the subject of A to comply with the Subject Constraint. These are rich open questions, though we hope to have shown that they are questions worth asking: that aboutness is a worthy primitive in the theory of counterfactuals, one superior to the notion of similarity in its account of figure–ground contrasts. For the last half-century, theories that limit themselves to the truth conditions of counterfactual antecedents have been at the foreground of work on counterfactuals. It is time for these theories to move to the background, and for the hyperintensional notion of aboutness to take their place.

References

Arló-Costa, Horacio (2007). The Logic of Conditionals. *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2007. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/fall2007/entries/logic-conditionals/.

Arregui, Ana (2005). On the accessibility of possible worlds: The role of tense and aspect. PhD thesis. University of Massachusetts Amherst. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.456.7358&rep=rep1&type=pdf.

Barwise, Jon and John Perry (1983). *Situations and Attitudes*. MIT Press.

Ben-Yami, Hanoch (2012). Response to Westerståhl. *Logique et Analyse* 55.217, pp. 47–55.

Bennett, Jonathan (1984). Counterfactuals and temporal direction. *The Philosophical Review* 93.1, pp. 57–91. DOI: 10.2307/2184413.

— (2003). *A Philosophical Guide to Conditionals*. Oxford University Press. DOI: 10.1093/0199258872.001.0001.

Berto, Francesco (2022). *Topics of thought: The logic of knowledge, belief, imagination*. Oxford University Press. DOI: 10.1093/oso/9780192857491.001.0001.

Biber, Douglas et al. (2021). *Grammar of spoken and written English*. John Benjamins. DOI: 10.1075/z.232.

Briggs, Ray (2012). Interventionist counterfactuals. *Philosophical Studies* 160.1, pp. 139–166. DOI: 10.1007/s11098-012-9908-5.

Brogaard, Berit (2011). *Transient truths: An essay in the metaphysics of propositions*. Oxford University Press.

Canavotto, Ilaria (2020). Where Responsibility Takes You: Logics of agency, counterfactuals and norms. PhD thesis. University of Amsterdam. DOI: 10.1007/978-3-031-17111-6.

Cariani, Fabrizio (2021). *The Modal Future: A Theory of Future-Directed Thought and Talk*. Cambridge University Press.

Cariani, Fabrizio and Paolo Santorio (2018). *Will* done better: Selection semantics, future credence, and indeterminacy. *Mind* 127.505, pp. 129–165. DOI: 10.1093/mind/fzw004.

Chellas, Brian F. (1975). Basic conditional logic. *Journal of Philosophical Logic* 4.2, pp. 133–153. DOI: 10.1007/BF00693270.

Ciardelli, Ivano, Linmin Zhang, and Lucas Champollion (2018a). Two switches in the theory of counterfactuals. *Linguistics and Philosophy* 41.6, pp. 577–621. DOI: 10.1007/s10988-018-9232-4.

— (2018b). Two switches in the theory of counterfactuals: A study of truth conditionality and minimal change. *Linguistics and Philosophy* 41.6, pp. 577–621. DOI: 10.1007/s10988-018-9232-4.

Condoravdi, Cleo (2002). Temporal Interpretation of Modals: Modals for the Present and for the Past. *The Construction of Meaning*. CSLI Publications, pp. 59–88. URL: https://semanticsarchive.net/Archive/2JmZTIw0/temp-modals.ps.

Conner, Jack Edward (1968). *A grammar of standard English*. Boston: Houghton Mifflin Company.

Croft, William and D. Alan Cruse (2004). *Cognitive Linguistics*. Cambridge University Press.

Davidson, Donald (1973). Radical interpretation. *Dialectica*, pp. 313–328.

Dimitriadis, Alexis (2008). The event structure of irreducibly symmetric reciprocals. *Event structures in linguistic form and interpretation*. Ed. by Martin Schäfer Johannes Dölling Tatjana Heyde-Zybatow, pp. 327–354. DOI: 10.1515/9783110925449.327.

Dowty, David (1991). Thematic proto-roles and argument selection. *Language* 67.3, pp. 547–619. DOI: 10.1353/lan.1991.0021.

Dutilh Novaes, Catarina (2003). A medieval reformulation of the *de Dicto*/*de Re* distinction. *Libor Behounek, Logica Yearbook*, pp. 111–124.

Egré, Paul and Hans Rott (2021). The Logic of Conditionals. *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University.

Fine, Kit (2000). Neutral relations. *The Philosophical Review* 109.1, pp. 1–33. DOI: 10.2307/2693553.

— (2012). Counterfactuals without possible worlds. *Journal of Philosophy* 109.3, pp. 221–246. DOI: 10.5840/jphil201210938.

— (2014). Truth-Maker Semantics for Intuitionistic Logic. *Journal of Philosophical Logic* 43.2, pp. 549–577. DOI: 10.1007/s10992-013-9281-7.

— (2017a). A theory of truthmaker content I: Conjunction, disjunction and negation. *Journal of Philosophical Logic* 46.6, pp. 625–674. DOI: 10.1007/s10992-016-9413-y.

— (2017b). A theory of truthmaker content II: Subject-matter, common content, remainder and ground. *Journal of Philosophical Logic* 46.6, pp. 675–702. DOI: 10.1007/s10992-016-9419-5.

— (2017c). Truthmaker Semantics. *A Companion to the Philosophy of Language*. Wiley Blackwell. Chap. 22, pp. 556–577. DOI: 10.1002/9781118972090.ch22.

— (2020). Yablo on subject-matter. *Philosophical Studies* 177, pp. 129–171. DOI: 10.1007/s11098-018-1183-7.

von Fintel, Kai (2001). Counterfactuals in a dynamic context. *Ken Hale: A life in language*. Ed. by Michael Kenstowicz. Vol. 36. MIT Press, pp. 123–152. URL: http://web.mit.edu/fintel/fintel-2001-counterfactuals.pdf.

Firbas, Jan (1964). On defining the theme in functional sentence analysis. *Trauaux Linguistiques de Prague* 1, pp. 267–280.

Fodor, Jerry (1975). *The Language of Thought*. New York: Thomas Y. Crowell.

— (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press.

Galles, David and Judea Pearl (1998). An axiomatic characterization of causal counterfactuals. *Foundations of Science* 3.1, pp. 151–182. DOI: 10.1023/A:1009602825894.

Gaskin, Richard and Daniel J Hill (2012). On neutral relations. *Dialectica* 66.1, pp. 167–186. DOI: 10.1111/j.1746-8361.2012.01294.x.

Geach, Peter (1957). *Mental Acts*. Routledge.

Gleitman, Lila R (1965). Coordinating conjunctions in English. *Language* 41.2, pp. 260–293. DOI: 10.2307/411878..

Gleitman, Lila R et al. (1996). Similar, and similar concepts. *Cognition* 58.3, pp. 321–376. DOI: 10.1016/0010-0277(95)00686-9.

Glüer, Kathrin (2011). The Principle of Charity. *Donald Davidson: A Short Introduction*. Oxford. Chap. 3, pp. 112–152. DOI: 10.1093/acprof:osobl/9780195382976.003.0003.

Goggans, Phillip (1992). Do the closest counterfactual worlds contain miracles? *Pacific Philosophical Quarterly* 73.2, pp. 137–149. DOI: 10.1111/j.1468-0114.1992.tb00332.x.

Goodman, Nelson (1947). The Problem of Counterfactual Conditionals. *Journal of Philosophy* 44.5, pp. 113–128. DOI: 10.2307/2019988.

— (1951). *The Structure of Appearance (third edition)*. Springer. DOI: 10.1007/978-94-010-1184-6.

— (1972). *Problems and Projects*. Bobbs-Merrill.

Haider, Hubert (2019). On expletive, semantically void, and absent subjects. *Semantic and syntactic aspects of impersonality* 26, p. 11.

Hempel, Carl G (1945). Studies in the Logic of Confirmation (I.) *Mind* 54.213, pp. 1–26.

Hiddleston, Eric (2005). A Causal Theory of Counterfactuals. *Noûs* 39.4, pp. 632–657. DOI: 10.1111/j.0029-4624.2005.00542.x.

Hintikka, Jaakko (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press.

Hitchcock, Christopher (2004). Routes, processes and chance-lowering causes. *Cause and chance*. Ed. by Phil Dowe and Paul Noordhof. Routledge, pp. 138–151.

— (2007). Prevention, preemption, and the principle of sufficient reason. *The Philosophical Review* 116.4, pp. 495–532. DOI: 10.1215/00318108-2007-01.

Ippolito, Michela (2013). *Subjunctive Conditionals: A Linguistic Analysis*. MIT Press.

Kaufmann, Stefan (2005). Conditional predictions. *Linguistics and Philosophy* 28.2, pp. 181–231. DOI: 10.1007/s10988-005-3731-9.

— (2013). Causal premise semantics. *Cognitive Science* 37.6, pp. 1136–1170. DOI: 10.1111/cogs.12063.

— (2017). The limit assumption. *Semantics and Pragmatics* 10, pp. 1–29. DOI: 10.3765/sp.10.18.

Keshet, Ezra (2010). Situation economy. *Natural Language Semantics* 18, pp. 385–434. DOI: 10.1007/s11050-010-9059-1.

Keshet, Ezra and Florian Schwarz (2019). *De re/de dicto. The Oxford Handbook of Reference*. Ed. by Jeanette Gundel and Barbara Abbott. Oxford. Chap. 10, pp. 168–202. DOI: 10.1093/oxfordhb/9780199687305.013.10.

Khoo, Justin (2015). On indicative and subjunctive conditionals. *Philosophers' Imprint* 15.32, pp. 1–40. URL: http://hdl.handle.net/2027/spo.3521354.0015.032.

— (2017). Backtracking counterfactuals revisited. *Mind* 126.503, pp. 841–910. DOI: 10.1093/mind/fzw005.

Kline, A. David (1980). Are There Cases of Simultaneous Causation? *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1, pp. 292–301. DOI: 10.1086/psaprocbienmeetp.1980.1.192573.

Kratzer, Angelika (1981). Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic* 10.2, pp. 201–216. DOI: 10.1007/BF00248849.

— (1989). An investigation of the lumps of thought. *Linguistics and Philosophy*, pp. 607–653. DOI: 10.1007/BF00627775.

— (2012). *Modals and conditionals*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199234684.001.0001.

Lakoff, George (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Lakoff, George and Stanley Peters (1969). Phrasal conjunction and symmetric predicates. *Modern studies in English*. Ed. by Sanford A. Schane David A. Reibel. Pretince-Hall, pp. 113–142.

Lasersohn, Peter (1995). *Plurality, conjunction and events*. Kluwer. DOI: 10.1007/978-94-015-8581-1.

Leśniewski, Stanisław (1931). O podstawach matematyki [On the Foundations of Mathematics], I–V. *Przeglad Filozoficzny*. Volume 30 (1927), 164–206; 31 (1928), 261–291; 32 (1929), 60–101; 33 (1930), 77–105; 34 (1931), 142–170. Abridged English translation by Vito Sinisi 1983.

Lewis, David (1973a). *Counterfactuals*. Wiley-Blackwell.

— (1973b). Counterfactuals and comparative possibility. *Ifs: Conditionals, Belief, Decision, Chance and Time*. Springer, pp. 57–85. DOI: 10.1007/978-94-009-9117-0_3.

— (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of philosophical logic*, pp. 217–234. DOI: 10.1007/BF00248850.

Liebesman, David (2013). Converse and identity. *Dialéctica* 67.2, pp. 137–155. DOI: 10.1111/1746-8361.12020.

Link, Godehard (1983). The Logical Analysis of Plurals and Mass Terms: A Lattice Theoretical Approach. *Meaning, Use, and Interpretation of Language/*. Walter de Gruyter. DOI: 10.1515/9783110852820.302.

MacBride, Fraser (2007). Neutral relations revisited. *Dialectica* 61.1, pp. 25–56. DOI: 10.1111/j.1746-8361.2007.01092.x.

Mandelkern, Matthew (2018). Talking about worlds. *Philosophical Perspectives* 32.1, pp. 298–325. DOI: 10.1111/phpe.12112.

— (2019). Crises of *Identity*. *Proceedings of the 22nd Amsterdam Colloquium*, pp. 279–288.

— (2021). If p, then p! *Journal of Philosophy* 118.12, pp. 645–679. DOI: 10.5840/jphil20211181242.

Mathesius, Vilém (1975). *A functional analysis of present day English on a general linguistic basis*. Walter de Gruyter.

Matthewson, Lisa (2001). Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9.2, pp. 145–189. DOI: 10.1023/A:1012492911285.

McDonald, Jenn (2025). Essential structure for causal models. *Australasian Journal of Philosophy*, pp. 1–23.

Moltmann, Friederike (2019). Natural language and its ontology. *Metaphysics and cognitive science*. Ed. by Alvin I. Goldman and Brian P. McLaughlin, pp. 206–232. DOI: 10.1093/oso/9780190639679.003.0009.

Nute, Donald (1980). *Topics in conditional logic*. Springer. DOI: 10.1007/978-94-009-8966-5.

Pearl, Judea (2000). *Causality*. Cambridge University Press.

Percus, Orin (2000). Constraints on some other variables in syntax. *Natural Language Semantics* 8.3, pp. 173–229. DOI: 10.1023/A:1011298526791.

Pietroski, Paul M (2018). *Conjoining meanings: Semantics without truth values*. Oxford University Press.

Putman, Hilary (1958). Formalization of the concept "About". *Philosophy of Science* 25.2, pp. 125–130.

Reinhart, Tanya (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27.

Rubin, Edgar (1915). Synsoplevede Figurer: Studier i psykologisk Analyse [Visually experienced figures: Studies in psychological analysis]. PhD thesis. University of Copenhagen.

Rubinstein, Aynat (2009). Groups in the semantics of reciprocal verbs. *Proceedings of the North East Linguistic Society, NELS38*. Ed. by Muhammad Abdurrahman Anisa Schardl Martin Walkow, pp. 269–282.

Russell, Bertrand (1903). *The Principles of Mathematics*. Cambridge University Press.

Russell, Bertrand (1905). On denoting. *Mind* 14.56, pp. 479–493. DOI: 10.1093/mind/XIV.4.479.

Santorio, Paolo (2018). Alternatives and truthmakers in conditional semantics. *The Journal of Philosophy* 115.10, pp. 513–549. DOI: 10.5840/jphil20181151030.

— (2019). Interventions in premise semantics. *Philosophers' Imprint*. URL: http://hdl.handle.net/2027/spo.3521354.0019.001.

Schulz, Katrin (2011). "If you'd wiggled A, then B would've changed". *Synthese* 179.2, pp. 239–251. DOI: 10.1007/s11229-010-9780-9.

Siloni, Tal (2012). Reciprocal verbs and symmetry. *Natural Language & Linguistic Theory* 30, pp. 261–320. DOI: 10.1007/s11049-011-9144-2.

Simons, Peter (1987). *Parts: A study in ontology*. Oxford University Press.

Sinisi, Vito (1983). Leśniewski's foundations of mathematics. *Topoi* 3.1, pp. 3–52. DOI: 10.1007/BF00139700.

Spector, Benjamin and Vincent Mouly (2025). Restrictor Constancy. *Manuscript*.

Stalnaker, Robert (1968). A theory of conditionals. *Ifs*. Ed. by William L. Harper, Robert Stalnaker, and Glenn Pearce. Springer, pp. 41–55. DOI: 10.1007/978-94-009-9117-0_2.

Starr, W. (2022). Counterfactuals. *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Winter 2022. Metaphysics Research Lab, Stanford University.

Talmy, Leonard (1975). Figure and ground in complex sentences. *Annual Meeting of the Berkeley Linguistics Society*, pp. 419–430. DOI: 10.3765/bls.v1i0.2322.

— (2000). *Toward a Cognitive Semantics: Concept Structuring Systems*. Vol. 1. MIT Press.

Thomason, Richmond (1970). Indeterminist time and truth-value gaps. *Theoria* 36.3, pp. 264–281. DOI: 10.1111/j.1755-2567.1970.tb00427.x.

Thomason, Richmond and Anil Gupta (1980). A theory of conditionals in the context of branching time. *Ifs*. Springer, pp. 299–322. DOI: 10.1007/978-94-009-9117-0_15.

Tversky, Amos (1977). Features of similarity. *Psychological review* 84.4, p. 327. DOI: 10.1037/0033-295X.84.4.327.

Wagemans, Johan et al. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological bulletin* 138.6, pp. 1172–1217. DOI: 10.1037/a0029333.

Wertheimer, Max (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung* 4. (Translated extract reprinted as "Laws of organization in

perceptual forms." In W. D. Ellis (Ed.), (1938). *A source book of Gestalt psychology* (pp. 71–94). Routledge, pp. 301–350.

Whorf, Benjamin Lee (1940). Gestalt technique of stem composition in Shawnee. Appendix to *Shawnee Stems and the Jacob P. Dunn Miami Dictionary* by C. F. Voegelin. Reprinted in *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*, 1956, MIT Press, pp. 393–406.

Williamson, Timothy (1985). Converse relations. *The Philosophical Review* 94.2, pp. 249–262. DOI: 10.2307/2185430.

Wilson, Neil L (1959). Substances without substrata. *The Review of Metaphysics* 12 (4), pp. 521–539.

Winter, Yoad (2018). Symmetric predicates and the semantics of reciprocal alternations. *Semantics and pragmatics* 11, pp. 1–52. DOI: 10.3765/sp.11.1.

Woodward, James (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press. DOI: 10.1093/0195155270.001.0001.

— (2016). The problem of variable choice. *Synthese* 193, pp. 1047–1072. DOI: 10.1007/s11229-015-0810-5.

Yablo, Stephen (2014). *Aboutness*. Princeton University Press. DOI: 10.1515/9781400845989.

Zhang, Yuhan and Kathryn Davidson (2021). *De re* interpretation in belief reports: An experimental investigation. *Experiments in Linguistic Meaning* 1, pp. 310–321. DOI: 10.3765/elm.1.4874.

— (2024). Interpreting referential noun phrases in belief reports–the *de re*/*de dicto* competition. *Glossa Psycholinguistics* 3.1. DOI: 10.5070/G60111258.