

Winner of *The Philosophical Quarterly Essay Prize* 2013

## THE IDENTITY OF A MATERIAL THING AND ITS MATTER\*

BY MAHRAD ALMOTAHARI

*I have both a smaller and a larger aim. The smaller aim is polemical. Kit Fine believes that a material thing—a Romanesque statue, for example, or an open door—can be distinguished from its constituent matter—a piece of alloy, say, or a hunk of plastic—without recourse to modal or temporal considerations. The statue is Romanesque; the piece of alloy is not Romanesque. The door is open; the hunk of plastic is not open. I argue that these considerations, when combined with a proper understanding of how the use of ‘not’ is functioning, entail that the statue is the piece of alloy, and that the door is the hunk of plastic. Far from challenging the doctrine that a material thing is its matter, Fine’s observations confirm the view. My larger aim is methodological. I will show that natural language semantics can guide inquiry in certain areas of metaphysics by helping us to advance lingering debates.*

**Keywords:** negation, truth-functionality, identity, coincidence, monism, category mistakes.

‘Philosophy is the strangest of subjects: [. . .] it attempts to deal with the most profound questions and yet constantly finds itself preoccupied with the trivialities of language [. . .]’—Kit Fine in Steve Pyke’s *Philosophers*.

A quick survey of twentieth-century philosophy in the English-speaking world reveals an interesting relationship between various achievements in the study of language and certain developments in metaphysics, epistemology, and the philosophy of mind. For example, research about the semantics of proper names and natural kind terms provided a firmer ground for Aristotelian essentialism and mind–body dualism (Putnam 1973; Kripke 1980; Chalmers 1996). The analysis of counterfactual conditionals in terms of possible worlds quickly led to counterfactual theories of causation and knowledge, not to mention a

\* For immensely helpful conversations about this material, either in person or via e-mail, I’m indebted to Elizabeth Barnes, Corine Besson, Ross Cameron, Bill D’Alessandro, Aidan Gray, Adam Hosein, Nick Huggett, Bernhard Nickel, and Marya Schechtman. Special thanks to Corine for her quick and detailed written comments at the eleventh hour, and to Aidan for his illuminating criticism of the first draft. I’m also grateful to the anonymous referees who reviewed this paper; their feedback was thought-provoking.

more serious attitude towards modal metaphysics generally (Stalnaker 1968; Lewis 1971; Nozick 1981). The study of context sensitivity motivated (and was in turn partly motivated by) contextualist responses to Cartesian scepticism (DeRose 1995; Lewis 1996). Our ever-expanding knowledge of the logical properties of language encouraged some authors to treat it as a ‘basic tenet’ or ‘fundamental axiom of all analytical philosophy’ that ‘the philosophy of thought can be approached only through the philosophy of language. That is to say, there can be no account of what thought is, independently of its [linguistic] means of expression’ (Dummett 1991, p. 3). More recently—in the twenty-first century—the analysis of knowledge-how ascriptions has revitalized intellectualism in the philosophy of mind (Stanley and Williamson 2001; Stanley 2011). The list goes on, of course, and in each case we encounter a new variation on a recurring theme: a proper understanding of some fragment of language is supposed to move a substantive philosophical debate forward.

It may well be that in some or all of the cases I’ve enumerated, authors were simply overreaching. Perhaps they were mistaken to treat the results of their linguistic investigations as evidence for substantive philosophical conclusions. Maybe their claims about language, too, were objectionable. In any event, the examples are suggestive. They indicate, I think, an abiding methodological hope for long-standing disagreements to be advanced by considerations that all sides regard as both reliable and authoritative in virtue of a shared competence with natural language.

I sympathize with the methodology. Brute intuitions often vary, since authors may have different orientations and goals. And even if authors share intuitions about the subject of disagreement, they may assign the intuitions different degrees of importance. So, debates about the subject matter may be accompanied by lower-level, and difficult-to-identify, disagreements about what counts as evidence. In such cases, parties to the dispute are likely to produce question-begging arguments. The hope is that by investigating the language in which the disagreement is carried out, we’ll be able to appeal to considerations that ultimately issue from a shared competence with English, and thus constitute a body of mutually acceptable and genuinely reliable evidence. Admittedly, it’s hard to see how this could work in the abstract, but in the discussion to follow I present an extended argument that, I believe, realizes the hope in a specific and unexpected way. The argument is a new route to the familiar metaphysical doctrine that a material thing—such as a statue, or a door—is identical with its constituent matter—such as a piece of alloy, or a hunk of plastic.

Kit Fine (2003, 2006) believes that a Romanesque statue, or an open door, can be distinguished from its constituent matter without recourse to modal or temporal considerations. The statue is Romanesque; the piece of alloy from which it’s made is not Romanesque. The door is open; the hunk of plastic

that constitutes it is not open. I will argue that these considerations, when combined with a proper understanding of how the use of ‘not’ is functioning, *entail* that the statue *is* the piece of alloy, and that the door *is* the plastic. Far from challenging the doctrine that a material thing is its matter, Fine’s observations *confirm* the view.

An attractive feature of my argument is that it doesn’t rely at all on appeals to brute intuition, but only on a set of empirically verifiable—and, I hope to show, *verified*—semantic judgments issuing from our knowledge of English. Further advantages will be discussed once the argument has been fully spelled out.

There are two parts to this paper. In the first, I provide empirical support for a descriptive generalization about the availability of so-called ‘irregular’ interpretations of the word ‘not’. The sort of interpretation I have in mind is typically called ‘metalinguistic negation’, and is distinguished from ‘descriptive’ or truth-functional negation. Recent discussions of the phenomenon, along with the more familiar ways of labelling it, trace back to Laurence Horn (1985, 1989, ch. 6). But in light of objections due to van der Sandt (1991, 1998) and Geurts (1998), I take it to be an open question whether irregular occurrences of ‘not’ are genuinely non-truth-functional metalinguistic operators. It makes no difference for my purpose here whether van der Sandt and Geurts are right, but I want to avoid question-begging descriptions of the phenomenon. As it happens, I do think that irregular occurrences of ‘not’ express a non-truth-functional metalinguistic operator, but I won’t explain why on this occasion; the explanation would take us too far off track.<sup>1</sup>

In the second part of this paper, I exploit the descriptive generalization about irregular uses of ‘not’ to formulate my argument for the thesis that the statue is the constituent piece of alloy. At this stage, of course, it’s probably unclear how theorizing about the meaning of ‘not’ will enable us to formulate the sort of argument I’ve advertised. Some readers might wonder how any of this differs from the existing literature on negation, Leibniz’s law, and the identity of a material thing and its matter.<sup>2</sup> The answer is that the existing literature implements a defensive strategy in response to various objections targeting the identity thesis. Here I take an offensive stance, the aim of which is to convince the unconvinced that the statue *is* the piece of alloy.

## I. IRREGULAR NEGATION

Many of us are by now familiar with irregular negation. It’s received some attention in philosophy. Consider

<sup>1</sup> See Almotahari (unpublished).

<sup>2</sup> I have in mind Almotahari (2014) and Schnieder (2006).

(IMPLICATURE DENIAL)	Louis C. K. isn't funny; he's hilarious.
(PRESUPPOSITION DENIAL)	The king of France isn't bald; he doesn't exist!
(FORM DENIAL)	Our radar didn't detect several aircraft; it detected several aircraft.

There's a strong feeling that, in each of the examples above, 'not' is being used to somehow object to either an implicature associated with, or a presupposition carried by, or a syntactic element in, the embedded sentence. The use of 'not' isn't necessarily voicing an objection to the truth of the sentence it operates on. But is there a principled, reliable way to identify irregular occurrences of 'not'? Horn (1989, ch. 6) proposed three tests, all of which have come under fire. I want to look at two of the tests and consider the apparent counterexamples targeting them. I believe that they have certain properties that strongly suggest they aren't really instances of irregular negation after all. *A fortiori*, they aren't instances of irregular negation that fail to satisfy Horn's tests.

According to Horn (1989, p. 392), irregular negation can't be expressed with the use of a negative prefix, like 'im'-, 'un'-, 'non'-, and 'dis'-. For example:

- (1) (a) The king of France is {\*unhappy/not happy}—there isn't any king of France.<sup>3</sup>  
 (b) It {\*is impossible/isn't possible} for you to leave now—it's necessary.

Furthermore, irregular negation prohibits the use of negative polarity items (NPIs) but licenses positive polarity items (PPIs). Horn (1989, pp. 370, 374, and 396) invites us to consider the examples below.

- (2) (a) Chris didn't manage to solve {\*ANY/SOME} of the problems—he managed to solve ALL of them.  
 (b) Chlamydia is not {\*EVER/SOMETIMES} misdiagnosed, it is FREQUENTLY misdiagnosed.

<sup>3</sup> A referee objects to this example. Cersei might be neither happy nor unhappy, according to the referee, and yet 'Cersei is not happy' may well involve descriptive negation. More generally, the referee says that infelicitous negative prefixing can't be a *test* to identify irregular negation, even if it's a *hallmark* of irregular negation, because negative prefixing doesn't distinguish irregular uses of 'not' from descriptive uses when 'is not<sub>descriptive</sub> *F*' and 'is un-*F*' are non-equivalent in meaning. The referee acknowledges, however, that this complication doesn't undermine my use of infelicitous negative prefixing in Section II; 'is not<sub>descriptive</sub> open'/'is unopen' and 'is not<sub>descriptive</sub> shut'/'is unshut' are plausibly equivalent. In any case, I now doubt that PRESUPPOSITION DENIAL is an instance of irregular 'not'. I've decided to leave the main body of this article unaltered because the issues are complicated and largely irrelevant to my central line of thought. Given the use of '... there isn't any king of France' as the follow-up to 'The king of France isn't happy', the infelicity of negative prefixing in (1a) *does* test whether 'not' is being used irregularly. Secondly, it may be that 'unhappy' is sometimes understood to mean *sad*, and certainly one can be neither happy nor sad, but the relevant sense of 'unhappy' is a *state of mind that occupies a position on the spectrum between happiness and sadness which falls anywhere outside of the happiness region*. In the relevant sense, then, 'is not<sub>descriptive</sub> happy' and 'is unhappy' are equivalent in meaning.

Both the distribution of negative prefixes and the behaviour of polarity items can be explained on the assumption that irregular negation is a metalinguistic device whose contribution can be represented as follows:

- (H) Sentences of the form  $[\text{not}_{\text{irregular}}-S]$  mean that uttering  $[S]$  in the present (or some salient) context is objectionable.

I label the thesis here ‘(H)’ because it encapsulates the central claims in Horn’s treatment of negation, according to which irregular ‘not’ is a metalinguistic non-truth-functional operator.

(H) captures the metalinguistic character of irregular negation in terms of its role in communicating that the use of a specific piece of language is objectionable. It represents non-truth-functionality less directly. Assuming that a sentence is true iff what it means is true, the truth of  $[\text{not}_{\text{irregular}}-S]$  doesn’t require that  $[S]$  be false.  $[S]$  may well be true, but uttering it would, let’s suppose, be objectionable for any number of reasons. In that case, what  $[\text{not}_{\text{irregular}}-S]$  means would be true, and so would  $[\text{not}_{\text{irregular}}-S]$  itself. Thus, unlike classical ‘ $\neg$ ’, the use of ‘ $\text{not}_{\text{irregular}}$ ’ can be neutral with respect to the truth value of the sentence it embeds.

Why can’t irregular negation be expressed with a negative prefix? Consider (1a) again. ‘The king of France is unhappy’ implies the falsity of ‘The king of France is happy’. One can’t use the negative prefix and remain neutral with respect to whether the un-prefixed sentence is true or false. But as our discussion of (H) makes clear, the use of irregular negation does involve such neutrality. It’s not surprising, then, that negative prefixing is incapable of voicing irregular negation.

What about polarity? Well, given (H), the quantifiers in (2)—‘any’, ‘some’, ‘ever’, and ‘sometimes’—don’t occur within a negative environment. They’re screened off by quotation:

- (2) (c) Uttering ‘Chris managed to solve  $\{*\text{ANY/SOME}\}$  of the problems’ is objectionable in this (or some salient) context—he managed to solve ALL of them.  
 (d) Uttering ‘Chlamydia is  $\{*\text{EVER/SOMETIMES}\}$  misdiagnosed’ is objectionable in this (or some salient) context, it is FREQUENTLY misdiagnosed.<sup>4</sup>

<sup>4</sup> One might worry that this analysis of 2(a) and (b) is problematic because it screens off the use of ‘he’ and ‘it’ in the rectification clauses—‘... he managed to solve ALL of them’, ‘... it is FREQUENTLY misdiagnosed’—from their antecedents—‘Chris’ and ‘Chlamydia’, respectively—which occur within the scope of quotation marks. But we needn’t think of ‘he’ and ‘it’ as anaphoric pronouns. Instead, think of them as deictic, taking as their values the subject made salient in the initial clause.

But it's only within a negative environment that NPIs are permitted and PPIs prohibited. So, we should expect polarity items to pattern with irregular negation in precisely the way they do.

The tests here provide a nice piece of evidence in favour of (H). But Geurts (1998, p. 280) isn't convinced. He challenges the diagnostic value of the tests by drawing our attention to the examples below.

- (3) It is {impossible/not possible} that you have met the king of France, because there is no such person.
- (4) (a) Walter didn't give his ukulele to {\*SOMEBODY/ANYBODY}: he never owned a ukulele.  
 (b) Walter didn't {\*SOMETIMES regret/regret at ANY time} that he betrayed his wife: he has always been faithful to her.

Geurts claims that 'there are cases that Horn would classify as metalinguistic in which the negation is incorporated' (*ibid.*). For example, consider (3). Presumably, Geurts thinks that the negation is 'incorporated'—i.e. combines with the use of 'possible'—because one could just as easily (and without infelicity) say, 'It is impossible that you have met the king of France'. In other words, negative prefixing is licensed. Furthermore, Geurts claims, 'Horn's polarity test gives rise to similar problems. [. . .] In [(4)] a negative polarity item may occur, whereas a positive polarity item produces an awkward effect'. Geurts concludes that the tests are 'a poor diagnostic' and unreliable.

Geurts is assuming, of course, that (3) and (4) are examples of irregular negation. He writes, 'Horn would classify [them] as metalinguistic. . . ' (*ibid.*). But although they instantiate the superficial form of irregularity—namely, [not-*X*; *Y*] (Horn 1989, pp. 404 and 405)—there's an interesting difference between (1) and (2), on the one hand, and (3) and (4), on the other. The difference justifies treating (3) and (4) as cleverly disguised instances of descriptive negation.

In (1), the truth of the rectification clauses—'. . . there isn't any king of France' and '. . . it's necessary'—renders the un-prefixed initial clauses—'The king of France is not happy' and 'It isn't possible for you to leave now'—untrue, assuming that the use of 'not' is understood descriptively. But the rectification clause in (3)—'. . . there is no such person'—is compatible with and (assuming the appropriate restriction on the modal expression) *verifies* the descriptive interpretation of the initial clause—'It is not possible that you have met the king of France'. Similarly, in (2), the truth of the rectification clauses—'. . . he managed to solve ALL of them' and '. . . it is FREQUENTLY misdiagnosed'—*falsifies* the NPI-containing initial clauses—'Chris didn't manage to solve ANY of the problems' and 'Chlamydia is not EVER misdiagnosed'—if the use of 'not' is understood descriptively. In contrast, the rectification clauses in (4) *verify* the NPI-containing initial clauses, assuming that the use of 'not' is interpreted descriptively. These logical relationships between rectification clause and initial

clause are precisely what we should expect if we suppose that (1) and (2) involve irregular negation whereas (3) and (4) involve descriptive negation. Why is that? Well, Horn (1989, p. 391) hypothesizes on independent grounds that an occurrence of ‘not’ is assigned the irregular interpretation as a way of repairing the initial clause—immunizing it from the inconsistency that a descriptive interpretation would be committed to in light of the rectification clause. So, given their apparent descriptive inconsistency, (1) and (2) ‘must in effect be sent back through [the language processor], whence the marked, metalinguistic quality of the negation’ (Horn 1989, p. 388). Presumably, it’s because Horn is attracted to some such picture of the interpretive process that he regards ‘not’ as pragmatically, not semantically, ambiguous. Just as conversational implicature is generally taken to be a pragmatic phenomenon in virtue of being a post-semantic repair strategy—the purpose of which is to eliminate the appearance that the speaker is violating norms of communication which she hasn’t explicitly opted out of—so too, I gather, Horn takes the irregular interpretation of ‘not’ to result from a post-semantic repair strategy—the purpose of which is to vindicate the interpreter’s charitable assumption that the speaker wouldn’t so blatantly contradict herself. But there isn’t even the appearance of inconsistency between a descriptive interpretation of the initial clauses in (3) and (4) and their rectification clauses. So, we should expect that they wouldn’t be ‘sent back through’ to receive an irregular interpretation. Thus, Geurts is simply wrong to assume that Horn is committed to treating (3) and (4) metalinguistically.

The upshot is twofold. First, the diagnostic tests are reliable—but, I hasten to add, of limited applicability, since not all terms have a naturally occurring negative prefix, nor are all or even most instances of irregular ‘not’ voicing objections to an antecedently used (or merely salient though unused) polarity item. Fortunately, the limitations here aren’t as troubling as one might initially think, since we can reliably identify the sort of feature that the tests are supposed to detect by a closely related series of considerations.

Consider another example:

(5) Cersei is {\*unhappy/not happy}; she’s elated.

If one took ‘Cersei is happy’ to be objectionable because one believed that Cersei was in a state of mind that was *incompatible* with happiness, then there wouldn’t be anything awkward at all about expressing oneself by using ‘unhappy’, as long as one didn’t go on to say that Cersei is elated. So, when we imagine someone uttering or accepting (5), we’re imagining that she believes that Cersei is in a state of mind *compatible* with happiness. What the awkwardness of negative prefixing makes explicit, then, is that *negation isn’t functioning to communicate the speaker’s belief that the subject of her speech exhibits some property incompatible with the one actually predicated*. When *that*—call it the *diagnostic*

*feature*—is made clear, then we can be confident that ‘not’ occurs irregularly and thus isn’t being used to voice the speaker’s belief in the falsity of the embedded sentence; it’s being used to mark some other kind of infelicity, as in (5). The diagnostic feature is what the negative prefixing test is supposed to detect.

We can reliably identify the presence of the diagnostic feature even when the objectionable term doesn’t have a naturally occurring negative prefix. To illustrate, consider

(FORM DENIAL) Our radar didn’t detect several aircrafts; it detected several aircraft.

(6) Cersei doesn’t hate Tyrion; she loathes him.

Neither of the objectionable terms in the examples immediately above (‘aircrafts’, ‘hate’) have negative prefixes, but it’s clear that the diagnostic feature is present. Someone who utters or accepts FORM DENIAL isn’t objecting to the use of ‘aircrafts’ because she thinks our radar detected several watercraft, say, or several cars or trains instead. Nor is someone who utters or accepts (6) objecting to the use of ‘hate’ because she takes Cersei to have some favourable attitude towards Tyrion. In both cases, I believe, it’s clear that the use of negation isn’t functioning to communicate the speaker’s belief that some property incompatible with the one actually predicated is called for. When the negative prefixing test can’t be directly applied, we can determine whether the use of ‘not’ is irregular by asking whether the diagnostic feature is present.

I said that the upshot of our discussion is twofold. The first has been discussed. The second is that it’s not always transparent in virtue of surface form alone whether a given use of ‘not’ expresses irregular negation. The availability of an irregular interpretation of ‘not’ requires more. The non-transparency of irregular negation is actually relevant for the point I will go on to make in the second part of this paper. If an expert linguist such as Geurts can be misled to treat descriptive negations as irregular, then surely we philosophers can be misled to treat irregular negations as descriptive, and perhaps draw unwarranted conclusions.

In light of the foregoing considerations, I want to recommend

(IRREGULARITY) A sentence of the form [not- $X$ ;  $T$ ] has an irregular interpretation only if [not<sub>descriptive</sub>- $X$ ] and [ $T$ ] are incompatible.

The proposal here deserves qualification. I want to make five points about how to properly understand it.

The first is that IRREGULARITY doesn’t rule out the possibility that irregular occurrences of ‘not’ are sometimes unaccompanied by a rectification clause, [ $T$ ]. (Consider a modified example originally due to Horn: ‘For a



pessimist like him the cup isn't half full.'). IRREGULARITY merely imposes a constraint on the availability of an irregular interpretation when an initial negation *is* rectified. I do think that it is highly plausible, however, that any sentence of the form [not<sub>irregular</sub>-*X*] can *in principle* be supplemented with a rectification. So even in cases where there isn't anything like a rectification clause that explicitly accompanies the use of 'not', IRREGULARITY can be applied by just temporarily supposing that the appropriate rectification is present. (Consider: 'For a pessimist like him the cup isn't half full; it's half empty.' Note that the example now satisfies IRREGULARITY.)

Secondly, I'm assuming that [*T*] is a *canonical* rectification; it explains the mistake in uttering [*X*], but not by explicitly using quotation. In other words—to borrow a bit of useful terminology from Carnap's *Logical Syntax*—a canonical rectification explains the infelicity in the embedded sentence, but does so in the *material* mode of speech, not the *formal* mode. So, sentences like

(7) Jack doesn't love aircraft; the word you should use is 'aircraft'.

aren't, as one might have assumed, instances of the relevant form, which I will now represent as follows: [not-*X*; *T*<sub>canonical</sub>]. The feeling that (7) involves irregular negation is easily explained, however, by pointing out that if the rectification were reformulated in the material mode, the resulting construction *would* satisfy IRREGULARITY. Consider: 'Jack doesn't love aircraft; he loves aircraft'.

Thirdly, [*T*] needn't *state* the explanation in order to *provide* it; it's enough, I assume, if [*T*] employs the correct way of speaking and thereby *displays* why [*X*] is objectionable. In cases of scalar implicature denial, the explanation is that a logically stronger claim was appropriate. In cases of presupposition denial, the explanation is that a presupposition of the utterance is false. In other cases, the explanation will identify a morphological, syntactical, phonological, or stylistic infelicity.

Fourthly, the sort of incompatibility between [not<sub>descriptive</sub>-*X*] and [*T*<sub>canonical</sub>] relevant for the assessment of IRREGULARITY is *semantic* incompatibility. The truth conditions for the two clauses can't be mutually satisfied, as any number of examples—including (1) and (2) above—illustrate. I will rely on this observation a little later.

Last, I introduced IRREGULARITY on the heels of a brief description summarizing Horn's speculative psycholinguistic model of how interpreters assign irregular interpretations to 'not'. As one might recall, the suggestion was that interpreters assign the reading as a post-semantic repair strategy to avoid the attribution of inconsistency to the speaker. This may give the false impression that IRREGULARITY itself is committed to that model, or that I'm endorsing the model. I want to explicitly discourage readers from forming any such impression. Further investigation may reveal that Horn's model is

inaccurate. I defer to the expertise of researchers in cognitive science—as does Horn (1989, p. 391). For my purpose—and, I think, for Horn’s as well—the model is useful insofar as it serves to make explicit a logical or semantic asymmetry between genuine irregular negations on the one hand and cleverly disguised descriptive negations on the other. The relevant asymmetry is that irregular negations satisfy IRREGULARITY; descriptive negations, such as (3) and (4) above, don’t. One should think of IRREGULARITY as merely a proposal about the conditions for the *availability* of an irregular interpretation, which may or may not be assigned on the basis of interpretive repair.

The empirical support for IRREGULARITY derives from its correctly predicting the *unavailability* of irregular uses of ‘not’ to deny relevance-based implicatures (i.e. implicatures that are triggered by apparent violations of the maxim *Be relevant!*). For example, one can utter ‘He was able to solve the problem’ in response to ‘Did he solve the problem?’ and thus implicate that he *did* solve the problem. But the relevance-based implicature—that he did solve the problem—can’t be denied as follows: ‘He wasn’t able to solve the problem; he didn’t solve it’. Infelicity results.<sup>5</sup> The explanation, according to Horn (1989, pp. 387–92), is that a descriptive interpretation of the initial clause—‘He wasn’t able to solve the problem’—is perfectly compatible with what is supposed to be the rectification clause—‘... he didn’t solve it’. Given IRREGULARITY, it’s not at all surprising that infelicity should result, since an irregular interpretation of ‘not’ is unavailable. So, the use of ‘not’ is required to take a descriptive reading, which is incapable of functioning as a mechanism for implicature denial. Furthermore, the explanation in terms of IRREGULARITY is robust across a wide range of such examples. (I encourage interested readers to look at Section 6.3.2 of Horn 1989.) My point is simply that the very same explanation applies equally well to (3) and (4), and thus undermines the assumption—vital to the argument challenging the reliability of the prefixing and polarity item tests—that ‘Horn would classify [(3) and (4)] as metalinguistic’ (Geurts, *ibid.*).

## II. ARGUING FOR IDENTITY

Philosophers are beginning to recognize that irregular negation bears on certain philosophical debates in which the use of ‘not’ plays an important role.

<sup>5</sup> A referee informs me that these ‘was able to’ constructions are more complicated than they appear to be, and that their communicative effect may not be due to conversational implicature. In any case, the same point could be made with other examples. Plausibly, ‘to assert that someone was clever enough to do something will generally implicate that she did it’, but one can’t selectively deny this implicature by saying, ‘She wasn’t clever enough to figure out the solution; she didn’t’ (Horn 1989, p. 388). A typical utterance of ‘I broke a finger yesterday’ ‘R-implicates’ that the broken finger was mine, but I can’t deny this implicature by saying, ‘I didn’t break a finger yesterday; the broken finger wasn’t mine’.

In particular, it bears on debates involving the application of Leibniz's law. I'd like to connect my discussion in the previous section with one such debate.

A number of metaphysical doctrines have been called 'monism'. According to the thesis that interests me, spatially coincident objects are identical. Advocates of the doctrine are committed to thinking that an artefact (a statue, say, or a door) and its constituent matter (a piece of alloy, for example, or a hunk of plastic) are one and the same entity. Pluralism, then, is the doctrine that numerically distinct entities can occupy the same region of space at the same time.

Now the modal and temporal objections to monism are well known, and so are the counterpart-theoretic rejoinders. In a series of recent articles, however, Kit Fine (2000, 2003, 2006) sets out to end the dialectical standoff in favour of pluralism by presenting arguments that are immune to familiar defensive strategies.

Consider two of his arguments.

- (F<sub>1</sub>) The statue is Romanesque.
- (F<sub>2</sub>) The piece of alloy isn't Romanesque.
- (F<sub>3</sub>) Therefore, the statue isn't the piece of alloy.
- (F<sub>4</sub>) The door is {open/shut}.
- (F<sub>5</sub>) The hunk of plastic isn't {open/shut}.
- (F<sub>6</sub>) Therefore, the door isn't the hunk of plastic.

I want to make two critical observations about them.

The first observation is that Horn's negative prefixing test applied to (F<sub>5</sub>) yields an affirmative diagnosis. 'The hunk of plastic is {unopen/unshut}' is just as awkward as 'The hunk of plastic is {open/shut}', which is precisely what one would expect given the hypothesis that irregular negation is in use.<sup>6</sup> Unfortunately, the prefixing test can't be straightforwardly applied to (F<sub>2</sub>) because 'Romanesque' doesn't have a naturally occurring negative prefix. But if one thinks about the kind of thing that the prefixing test is supposed to make explicit (*thereby* indicating the irregularity of 'not'), one will appreciate that (F<sub>2</sub>) is relevantly like the cases that issue in a positive diagnosis of irregularity.

It's clear—both from Fine's discussion and from a moment's reflection—that those of us who are disposed to object to 'The piece of alloy is Romanesque' are so disposed not because we think the piece of alloy is Gothic, or Arabesque, or pre-Romanesque, or post-Romanesque, or an instance of any other aesthetic style incompatible with being Romanesque. We are disposed to object to the sentence for some other reason, which may well be compatible with the truth

<sup>6</sup> 'Unopen' and 'unshut' are relatively uncommon, but they both appear in the Unabridged Merriam-Webster Online Dictionary.

of the sentence.<sup>7</sup> One naturally wonders why the sentence is objectionable. For the moment, let me postpone answering this question. I'll return to it shortly. For now, I simply want to suggest that the diagnostic feature, which the prefixing test is supposed to make explicit, is present in (F<sub>2</sub>). As for Horn's second test, which relies on the distribution of polarity items—it's inapplicable here, since the infelicity in both 'The hunk of plastic is {open/shut}' and 'The piece of alloy is Romanesque' is due to something other than an antecedently used (or salient though unused) polarity item.

I said that I wanted to make two critical observations. The first was that we have reliable evidence for thinking that, insofar as (F<sub>2</sub>) and (F<sub>5</sub>) are acceptable, they involve irregular uses of 'not'.<sup>8</sup> My second observation is that we can exploit this evidence to formulate a new argument for the conclusion that the statue *is* the piece of alloy, and that the door *is* the hunk of plastic.

My argument will rely crucially on an earlier proposal.

(IRREGULARITY) A sentence of the form [not-*X*; *Y*<sub>canonical</sub>] has an irregular interpretation only if [not<sub>descriptive</sub>-*X*] and [*Y*<sub>canonical</sub>] are incompatible.

But for this thesis to be applicable to (F<sub>2</sub>) and (F<sub>5</sub>), we will have to reformulate certain sentences so as to render them instances of the relevant form.

<sup>7</sup> A referee suggests that we might find 'The piece of alloy is Romanesque' objectionable because the piece of alloy has some property incompatible with being Romanesque, even if the property isn't being Gothic, being Arabesque, etc. Perhaps, the referee suggests, the piece of alloy has the property of lacking any aesthetic style. But, as I observe momentarily, 'The piece of alloy is Romanesque' occurs felicitously in some discourses. The anomalous character of the sentence had better cohere with its felicity in those contexts. Furthermore, one can imagine archeologists doubting whether a certain piece of alloy was part of some ancient artefact but drawing attention to the aesthetic style of the piece of alloy as evidence for thinking that it was.

<sup>8</sup> Is there a way of criticizing monism without raising worries about how to properly interpret 'not'? Perhaps. Consider: the piece of alloy is *uninteresting*; the statue is interesting. It may seem as though we're ascribing one property to the statue and an incompatible property to the piece of alloy. So they mustn't be identical. But the monist can reasonably deny that incompatible properties really are being attributed to the statue and the piece of alloy. The context '*... is (un)interesting*' is non-transparent. Suppose your birthday is coming up and you expect one really amazing gift from your parents. When you think about your potential gift, it elicits certain questions (What will it be? Will I like it? Will others like it? Will it get me a date with so-and-so?) and it disposes you to entertain certain possibilities (If it's the car I asked for, then I'll take so-and-so to the drive-in). So, naturally, the gift is interesting. Now, as it happens, a beat-up car appeared on your street a few days ago. It's totally unremarkable. It doesn't elicit any questions or dispose you to think about future possibilities. So, naturally, it's uninteresting. Can we infer that the beat-up car is distinct from your gift? No, for your parents did indeed purchase the car as your birthday gift, and they chose to hide it in plain sight. The opacity of '*... is (un)interesting*' isn't that surprising, since whether something is interesting is a psychological matter that partly depends on the way it's presented to us. I suspect that arguments like the one above—the statue is *F*; the piece of alloy is NEGATIVE PREFIX+ *F*—will have to be dealt with on a case-by-case basis. My aim here isn't to provide a generally applicable defensive strategy; it's to present a new offensive strategy.

- (F<sub>7</sub>) The piece of ALLOY isn't Romanesque, but the STATUE is.  
 (F<sub>8</sub>) The hunk of PLASTIC isn't {open/shut}, but the DOOR is.

Now, as I said a moment ago, we have reliable evidence that the initial clauses in (F<sub>7</sub>) and (F<sub>8</sub>) are irregular negations. I'll assume from here on that the evidence is accurate: (F<sub>7</sub>) and (F<sub>8</sub>) really are examples of irregular negation. It follows, then, that both (F<sub>7</sub>) and (F<sub>8</sub>) have irregular interpretations. So, by IRREGULARITY, it follows that the initial negation-involving clauses in both (F<sub>7</sub>) and (F<sub>8</sub>), if interpreted descriptively, are semantically incompatible with their respective rectification clauses. But semantic incompatibility between, say, 'The piece of alloy is not<sub>descriptive</sub> Romanesque' and 'The statue is Romanesque' requires that the definite descriptions pick out the very same object.<sup>9</sup> It's that object that's said to be both Romanesque and not Romanesque. Now consider the additional, and highly plausible, assumption that if [the *F*] and [the *G*] both pick out the same thing, then the *F* is the *G*. An instance of this schematic principle is that if 'the statue' and 'the piece of alloy' pick out the same thing, then the statue is the piece of alloy. By appealing to this instance of the principle, we're entitled to infer that the piece of alloy is the statue. *Mutatis mutandis*, the door is the hunk of plastic.<sup>10</sup>

I can think of one response on behalf of my interlocutor. I'll formulate the response as a challenge to my argument about the statue and the piece of alloy, but with very minor adjustments one can raise parallel doubts about the argument concerning the door and the hunk of plastic.

Semantic incompatibility, recall, is a relation between sentences that have conflicting truth conditions. Therefore, semantically incompatible sentences *have* truth conditions. Consequently, a sentence lacking in truth conditions is incapable of being semantically incompatible with another sentence. Now, one

<sup>9</sup> In general, semantic incompatibility doesn't require that two sentences be about the same thing. For example, 'There is a largest prime number' and 'Barack Obama is the 44<sup>th</sup> President of the United States' are semantically incompatible; their truth conditions can't be mutually satisfied. But that's because the former is necessarily false. Both of the sentences I'm concerned with above are, we may suppose, contingently true. For contingently true sentences to be semantically incompatible, they must be about the same thing.

<sup>10</sup> A referee asks how we can tell whether (F<sub>7</sub>) and (F<sub>8</sub>) are instances of the relevant form: [not-*X*; *Y*<sub>canonical</sub>]. Even if the initial clauses in (F<sub>7</sub>) and (F<sub>8</sub>) are irregular negations, we're entitled to conclude that (F<sub>7</sub>) and (F<sub>8</sub>) suffer from semantic incompatibility only if the second conjuncts in each are canonical rectifications. But, the referee wonders, why think so? Now, as Horn noted early on, irregular negations are often accompanied by intonational focus: 'Lovers don't COPULATE!' Why should that be? What sort of communicative function does the focus effect serve? Changing one's intonation is a good way of drawing attention to the element that ought to be rectified by the rectification clause. So, given that the initial clauses in (F<sub>7</sub>) and (F<sub>8</sub>) are irregular negations, and that one can naturally focus on 'alloy' and 'plastic', we have a good reason to expect that the canonical rectifications will rectify the phrases in which these objectionable terms occur. Treating 'the {STATUE/DOOR} is' as a canonical rectification coheres nicely with this expectation. Furthermore, it suggests a satisfying explanation of the error in the initial clauses. I provide this explanation below. In short, then, we're entitled to regard (F<sub>7</sub>) and (F<sub>8</sub>) as instances of [not-*X*; *Y*<sub>canonical</sub>] on the grounds that doing so is a serviceable hypothesis.

might be inclined to say that (8) is simply a *category mistake*: it's anomalous in just the sort of way that, say, 'Colorless green ideas sleep furiously' or 'Two is green' are anomalous.

(8) The piece of alloy is Romanesque.

Sentences that exhibit this kind of anomaly appear to be worse than obviously false. To illustrate the point, consider another example:

(9) Astrology is a reliable way to predict the future.

I submit that (9) is obviously false, but notice that it doesn't suffer from the same sort of infelicity that 'Colorless green ideas sleep furiously' suffers from. That is, (9) doesn't elicit the sort of phenomenology of error that standard examples of category mistakes do. Now suppose the pluralist wants to explain the anomalous character of category mistakes by hypothesizing that they're simply meaningless. (More precisely, the hypothesis is that category mistakes lack a *literal* meaning. They can certainly be assigned a non-literal meaning when used in context.) As it happens, Kit Fine wants to say just this about our target sentence, 'The piece of alloy is Romanesque'. He writes, 'It is worth emphasizing, in this connection, that these differences lie not merely in the *correct* but also in the *meaningful* application of the predicates. [. . .] a statue can meaningfully be said to be *Romanesque* or not *Romanesque*, though not the clay or the alloy itself [. . .]' (Fine 2003, pp. 207 and 208, emphasis in original). Given that a sentence has truth conditions only if it's meaningful, and given that (8) is meaningless, it follows that (8) doesn't have truth conditions. So, the pluralist might say that (8) isn't semantically incompatible with 'The piece of alloy is not<sub>descriptive</sub> Romanesque'. Since my argument for the identity of the statue and the piece of alloy crucially relies on the claim that the two sentences *are* semantically incompatible, one might be inclined to think that it ultimately fails to establish the desired conclusion.

I'd like to make three points in response. First, as is clear from the objection we're entertaining, the hypothesis that category mistakes are meaningless is supposed to explain their anomalous character. In particular, it's supposed to explain why they appear to competent English speakers to be worse than obviously false. But there's a whole host of problems with this explanation. For a detailed presentation of many of them, I recommend that readers have a look at Ofra Magidor's (2009) article, 'Category Mistakes are Meaningful'. I'll quickly summarize one difficulty that Magidor draws our attention to and move on.

How might one translate 'Two is green', say, into French? The translation would, presumably, map the English sentence onto 'Deux est vert'. But why *that* sentence of French and not some other? The most natural answer, I think, is that 'Two is green' and 'Deux est vert' are synonymous; they have the same

meaning, which strongly suggests, of course, that they both have meaning. So, they mustn't be meaningless. One can certainly challenge the intuition that 'Deux est vert' is synonymous with 'Two is green', but, as Magidor observes,

[...] it is not entirely obvious how to explain away this synonymy intuition. For example, one might try to argue that two sentences merely seem synonymous because for each word in the English sentence there is a corresponding word in the French sentence to which it is synonymous. This, however, will not do: the Hebrew translation, for example, of 'Two is green' is 'Shtaim yarok' which contains only two words and is not word-to-word synonymous with the English sentence. Still, it is intuitively synonymous with the English sentence [...]. (Magidor 2009, p. 565)

Admittedly, there's more one can say on behalf of the hypothesis that category mistakes are meaningless. Maybe with enough time and effort, advocates of the view can develop an account of correct translation that doesn't require synonymy, and that vindicates the intuition that 'Deux est vert' correctly translates 'Two is green'. But this difficulty is one of several that place a considerable amount of pressure to expand the view in revisionary ways, the cumulative effect of which undermines its initial plausibility. The preponderance of evidence, I believe, tells in favour of an alternative theory of category mistakes.

Secondly, I'd like to sketch a more credible story about why category mistakes appear to be worse than obviously false. It doesn't require much creativity to imagine the sort of mental state one would have to occupy in order to be disposed to sincerely and literally utter or accept a sentence like (9)—that is, a sentence that merely suffers from obvious falsity. One would simply have to believe that 'things happen for a reason', and that the reasons are, so to speak, 'written in the stars'. Of course, there's very little to recommend these beliefs, but many people find them attractive. Superstition is pervasive. Though unjustified, superstitious belief is at least broadly intelligible. It's much harder to imagine an intelligible frame of mind that would dispose someone to sincerely and literally utter or accept 'Colorless green ideas sleep furiously' or 'Two is green'. With a little ingenuity, of course, it can be done; but once it's done, there's absolutely no temptation to treat these sentences as worse than obviously false.

Consider the following scenario: John is a philosopher. He recently developed a new theory in the philosophy of mathematics according to which numbers are coloured, and the colour of the number two is green. For example, John may hold some naturalist position according to which the number two is the set of all pairs of physical objects in the world. In addition, John might hold that if most such pairs have a certain colour, then the set—and therefore the corresponding number—have this colour. (Compare this to a more popular philosophical position suggested by Penelope Maddy: a set of physical objects has a spatial location and it is located wherever its members are located) (Magidor 2009, p. 567).

John's situation suggests that whether a sentence elicits the feeling of error characteristic of category mistakes—the impression that they're worse than obviously false—depends, perhaps in large part, on background beliefs that aren't often questioned but may well be. Situations like this aren't merely hypothetical. Consider the sentence 'Spacetime is curved', or the sentence 'Physical space is an algebra'. To ignorant but linguistically competent speakers of English, these sentences elicit the kind of feeling that authors often rely on to identify category mistakes; they seem worse than obviously false. But it's not at all plausible to suppose that they're meaningless, since they very likely express literal truths. At any rate, one wouldn't want to prejudge whether they express literal truths on the basis of one's theory of category mistakes. Whether 'Spacetime is curved' is literally true isn't to be settled in advance of inquiry, which may lead us, ultimately, to occupy a frame of mind that would render acceptance of 'Spacetime is curved' not only broadly intelligible but nigh on rationally mandatory. On the basis of these considerations, I want to recommend that the anomalous character of sentences we classify as category mistakes—the impression that they're worse than false—isn't due to literal meaningfulness—since 'Spacetime is curved' may well be literally true and it certainly hasn't gone from being meaningless to meaningful—but to our inability to imagine a broadly intelligible frame of mind that would rationalize accepting them.

The pluralist response to my central argument relies on the assumption that (8) is a category mistake. My third and final point is that this assumption is false, since there are contexts in which it's a natural and unobjectionable thing to say in a completely literal frame of mind. Consider

- (10) I was at the junkyard the other day and I saw a piece of alloy that had to be a work of art. I don't know whether it was a statue, or a sculpture, or a part of some building that's no longer standing, but the form was remarkable! If I had to bet, I would bet that the piece of alloy is Romanesque (Almotahari 2014, p. 511).

Set aside my reasons for doubting that category mistakes are meaningless, and suppose, along with Fine, that (8) is a meaningless category mistake. Why, then, should (8) occur felicitously in (10)? The speaker isn't using the sentence figuratively. One would like some explanation of why speakers are disposed to regard (8) as anomalous in isolation but as felicitous when embedded in (10). Treating (8) as a meaningless category mistake only makes the problem insurmountably hard.

This brings us back to a question that I set aside earlier but promised to return to. What's objectionable about (8)? The answer had better cohere with its felicity in (10). We're now in a position to consider the explanation.



A piece of alloy may or may not be a work of art. If the piece of alloy spatially coincides with a statue, then it *is* a work of art, because, we may suppose, it's a statue. However, what one communicates about the piece of alloy will depend on how one chooses to describe it. If you describe it as 'the alloy statue', or as simply 'the statue', then you communicate that you believe that it's a work of art. It's common knowledge, after all, that statues are works of art. So your conversational partner can reasonably infer on the basis of your diction that you take the subject of your speech to be an artwork. But you might choose to describe the subject of your speech in a different way. If you describe it as, say, 'the piece of alloy' instead, then your speech conversationally implicates that you don't believe that the object is a work of art. The reason is familiar: if you did take it to be a work of art, you could just as easily have communicated something stronger than you in fact did, thus conforming to the conversational maxim *Be informative!*, without violating any other conversational maxims, such as *Be relevant!* How might you have done so? Well, by using 'the alloy statue' instead, which would have conveyed the additional bit of information that you take the object to be an artwork. Given that you decided to withhold that additional information, your conversational partner can reasonably infer that you don't believe that the piece of alloy is a work of art.

Now suppose you utter (8) in isolation, fully intending to conform to principles governing cooperative discourse. What you said by uttering (8) could just as easily have been said by using 'The alloy statue is Romanesque'. (Again, I'm assuming that the piece of alloy just is the statue.) Thus, you implicate (that is, your interpreter can reasonably infer on the basis of Gricean reasoning) that you don't believe that the object is an artwork. But you also indicate that you *do* believe that the object is an artwork, since you go on to predicate being Romanesque of it. (It's common knowledge that Romanesque artefacts are works of art.) So you represent yourself as believing that the subject of your speech is an artwork and as not believing that it's an artwork. Thus, an utterance of (8) suffers from pragmatic inconsistency. But this kind of infelicity is perfectly compatible with the literal truth of the sentence. Moore's paradox provides a nice example: 'It's raining, but I don't believe that it is'. Both conjuncts may well be true, and yet by uttering the sentence one would represent oneself as suffering from a conflicted state of mind: believing that it's raining and not believing that it's raining.<sup>11</sup>

<sup>11</sup> What explains the infelicity in 'The hunk of plastic is {open/shut}'? The answer takes the same shape as the explanation above. Use of 'the hunk of plastic' triggers an implicature that generates pragmatic inconsistency. It's common knowledge that doors are designed to perform a specific function. So, if a speaker uses 'the door', she communicates that she believes the subject of her speech performs that function, or was designed to. But if the speaker uses 'the hunk of plastic' instead, she implicates that she doesn't believe the subject of her speech performs (or was designed to perform) that function. After all, she could just as easily have used 'the plastic door' instead, thereby communicating this additional information in conformity with the maxim *Be*

Unlike Moore's paradox, however, there are contexts in which a literal use of (8) is felicitous. (10) is one such context. Given my explanation of why (8) is anomalous in isolation, it's not at all mysterious why it should be felicitous in (10). The pragmatic inconsistency that makes (8) objectionable in isolation is partly due to the implicature that the speaker doesn't believe that the subject of her speech is an artwork. But this implicature is preemptively cancelled in (10). The discourse fragment begins with 'I saw a piece of alloy that had to be a work of art', which makes it unreasonable for interpreters to infer by Gricean reasoning that the speaker doesn't take the subject of her speech to be an artwork. Without this implicature, pragmatic inconsistency is avoided.

From the monist's point of view, my central argument displays a pleasing sort of irony. The considerations that were supposed to be most damaging to the commitments of her doctrine, (F<sub>1</sub>)–(F<sub>6</sub>), end up confirming them. These considerations thus constitute good abductive evidence in favour of monism itself. Another attractive feature of the argument is that it doesn't identify the artefacts in question with their constituent matter on the basis of brute intuition. It relies on widely shared linguistic judgments, ultimately grounded in our mutual competence with English, and on an empirically verified descriptive generalization. The argument thus serves as a nice example of how semantic theorizing can fruitfully interact with metaphysical inquiry, providing a new and unexpected avenue to advance an old debate.

### III. OPEN QUESTIONS AND CLOSING REMARKS

My central argument relies on the infelicity of certain sentences of English (in particular, 'The piece of alloy is Romanesque' and 'The hunk of plastic is {open/shut}'). But, an interlocutor might ask, why *should* linguistic infelicities support substantive metaphysical conclusions? I'm imagining that the interlocutor understands my central argument, so she's not asking for further clarification about any specific premise or inference, nor is she challenging any particular step. She's expressing a more general methodological puzzlement.

The puzzlement can be formulated in a slightly different way, which draws attention to both its distinguished history and its extreme difficulty. Kant famously asked how metaphysics is possible. If my strategy in the second part of this paper is successful, then we will have demonstrated the possibility of establishing substantive metaphysical conclusions from natural language semantics, since we will have actually done so. But, one might think, even if we

*informative!*, but she chose not to. It's also common knowledge that something can be open/shut only if it can perform the kind of function that doors are designed for. So, if a speaker utters 'The hunk of plastic is {open/shut}', she both implicates that she lacks a certain belief and represents herself as having that belief. See Almotahari (2014, p. 516) for an alternative explanation.

acknowledge this possibility, we will not have come any closer to understanding *how* it's possible.

I raised these questions not in order to answer them now but in order to acknowledge that they remain open and deserve serious consideration.<sup>12</sup> It also bears mentioning that this very general puzzlement over my methodology doesn't tell in favour of pluralism, since the pluralist relies on the very same strategy. All of Kit Fine's objections to monism are premised on the infelicitous use of certain predicates. He writes,

[...] the predicates in question have felicitous application to the one subject term but not to the other. Thus we can say that the door is open or shut but we cannot very well say that the plastic from which it is made is open or shut [...]. As I mentioned in the original paper [Fine 2003, p. 207], these various sorts give rise to their own 'sphere of discourse'; and predicates within one sphere will often not have felicitous application to objects belonging to other spheres. (Fine 2006, pp. 1069 and 1070)

On the basis of these observations about linguistic infelicity, Fine infers the *non*-identity of a material thing and its matter. But at no point does Fine attempt to resolve the methodological puzzle I've raised. I too make no attempt to resolve it, and take some comfort in the fact that the puzzle arises for pluralists and monists alike.

## REFERENCES

- Almotahari, M. (2014) 'Metalinguistic Negation and Metaphysical Affirmation', *Philosophical Studies*, 167: 497–517.  
 ——— Unpublished. 'Negation and Truth-Functionality'  
 Chalmers, D. (1996) *The Conscious Mind*. Oxford: OUP.  
 DeRose, K. (1995) 'Solving the Skeptical Problem', *Philosophical Review*, 104: 1–52.  
 Dummett, M. (1991) *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.  
 Fine, K. (2000) 'A Counterexample to Locke's Thesis', *Monist*, 83: 357–61.  
 ——— (2003) 'The Non-Identity of a Material Thing and Its Matter', *Mind*, 112: 195–234.  
 ——— (2006) 'Arguing for Non-Identity: A Response to King and Frances', *Mind*, 115: 1059–82.  
 Geurts, B. (1998) 'The Mechanisms of Denial', *Language*, 74: 274–307.  
 ——— (1985) 'Metalinguistic Negation and Pragmatic Ambiguity', *Language*, 61: 121–74.  
 ——— (1989) *A Natural History of Negation*. Stanford: CSLI.  
 Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.  
 Lewis, D. (1971) *Counterfactuals*. Oxford: Wiley-Blackwell.  
 ——— (1996) 'Elusive Knowledge', *Australasian Journal of Philosophy*, 74: 549–67.  
 Magidor, O. (2009) 'Category Mistakes Are Meaningful', *Linguistics & Philosophy*, 32: 553–81.  
 Nozick, R. (1981) *Philosophical Explanations*. Cambridge, MA: Harvard University Press.  
 Putnam, H. (1973) 'Meaning and Reference', *Journal of Philosophy*, 70: 699–711.  
 Schnieder, B. (2006) "'By Leibniz's Law": Remarks on a Fallacy', *Philosophical Quarterly*, 56: 39–54.  
 Stalnaker, R. (1968) 'A Theory of Conditionals', in W. L. Harper, R. Stalnaker and G. Pearce (eds) *Ifs*, 41–55. Kluwer: Dordrecht.

<sup>12</sup> I develop an account in work that's currently in progress.

Stanley, J. (2011) *Know How*. Oxford: OUP.

Stanley, J. and Williamson, T. (2001) 'Knowing How', *Journal of Philosophy*, 98: 411–44.

Van der Sandt, R. (1991) 'Denial', in *Papers from the Chicago Linguistics Society: The Parasession on Negation*, Vol. 27, 331–44. Chicago, IL: University of Chicago.

*University of Illinois at Chicago, USA*