The Conditional Fallacy
Author(s): Daniel Bonevac, Josh Dever, David Sosa
Source: *The Philosophical Review*, Vol. 115, No. 3 (Jul., 2006), pp. 273-316
Published by: Duke University Press on behalf of Philosophical Review
Stable URL: http://www.jstor.org/stable/20446909
Accessed: 03/05/2010 01:52

The Conditional Fallacy

Daniel Bonevac, Josh Dever, and David Sosa
University of Texas at Austin

The thought that dispositional properties can be understood in terms of the holding of certain counterfactuals goes back at least to Gilbert Ryle (1949, 123):[1]

> To say that this lump of sugar is soluble is to say that it would dissolve, if submerged anywhere, at any time and in any parcel of water. To say that this sleeper knows French, is to say that if, for example, he is ever addressed in French, or shown any French newspaper, he responds pertinently in French, acts appropriately or translates correctly into his own tongue.

This project of *conditional analysis* of dispositions faces considerable difficulties in refining the rather crude conditional connections that first come to mind.[2] A more recent line of discussion, however, has suggested

1. Rudolf Carnap (1936) discusses the analysis of dispositional properties using conditional connections and conditions of manifestation and observes the shortcomings of this approach in connection with objects that are never subjected to the appropriate conditions of manifestation; but he fails to take the step to counterfactual conditionals.

2. As Ryle (1949, 123) notes:

> This is, of course, too precise. We should not withdraw our statement that he knows French on finding that he did not respond pertinently when asleep, absent-minded, drunk or in a panic; or on finding that he did not correctly translate highly technical treatises.

These problems of refinement should be kept in mind when considering the more general problem of the conditional fallacy raised below. *Sometimes* what is wrong with a conditional analysis of a disposition is that it is simply the wrong conditional. If the conditional fallacy is to serve as a generic strategy for defeating conditional analyses, its plausibility should not turn on idiosyncrasies of particular proposed conditionals.

Philosophical Review, Vol. 115, No. 3, 2006
DOI 10.1215/00318108-2006-001

that there are *general arguments* against "any attempt to account for dispositions that makes essential use of strong conditionals and counterfactuals" (Martin 1994, 7–8). This line, which begins with Shope 1978 and runs through Johnston 1992, Wright 1992, Johnston 1993, Martin 1994, Lewis 1997, Bird 1998, and Fara 2005, among others, argues that attempts to give conditional analyses of dispositional properties commit what has come to be known as the *conditional fallacy*. Very strong claims have been made for the conditional fallacy; it has been taken to show the impossibility of any conditional analysis of any dispositional property.[3] It is this strong claim that interests us. That *some* conditional analyses fail is, of course, uncontroversial. That, however, there is a generic recipe for defeating conditional analyses—that what goes wrong is their very *con-*

3. Thus:

- The attempt to render dispositional claims in terms of counterfactual claims depends upon (i) providing an account of what needs to be the case for a counterfactual claim to be true or appropriate or correct; and when that is done, then (ii) it needs to be shown that it is sufficient for dispositions. This paper has been an argument for saying that there is no hope for the success of (ii). (Martin 1994, 7)
- The simple conditional analysis has been decisively refuted by C. B. Martin. (Lewis 1997, 143)
- Thanks to Charlie Martin, the conditional analysis . . . has long been known to be incorrect. (Bird 1998, 227)
- It is now widely agreed that the simple conditional account is mistaken. (Fara 2005, 46)

Below we will argue that such claims are too strong since the argument motivating them would threaten to block the very possibility of counterfactual conditionals at all. Consider properties that are simply *stipulated* to be conditional in their analysis ("let an object be *breakish* if it would break were it struck")—how could such properties be shown not to admit of a simple conditional analysis? Thus the following claim by Wright (1992, 118) must be too strong since it would rule out the possibility of stipulatively conditional properties such as *breakish*:

> If the class of judgements in which we are interested participate—or supervene upon participants—in the causal order, and if realising the antecedent of the purported conditional analysans would perforce involve changes in states of affairs so participating, then, it seems, however unlikely interference might actually be, it at least cannot be *a priori* that the changes involved would not impinge on the truth value of *P*. Nor, therefore, can it be a priori that *P*'s truth conditions are captured by the purported analysans. So, as an attempt at a true *conceptual* claim, the analysis must fail.

Whatever the final verdict on the conditional fallacy, it cannot be a requirement on analyses that it be a priori that (what we below call) finking situations not arise.

ditionality rather than the particular conditionals selected—is a claim of much greater philosophical moment.

After an elaborate but somewhat obscure generic formulation of the conditional fallacy by Robert Shope,[4] the practice in the literature on the conditional fallacy has been to proceed primarily through consideration of specific examples. Thus C. B. Martin gives us the case of the electro-fink. Suppose it is suggested that the dispositional property of a wire's being *live* is equivalent to the wire's being such that *if it were touched by a conductor, electrical current would flow from the wire to the conductor.*

> Consider now the following case. [A wire] is connected to a machine, an *electro-fink*, which can provide itself with reliable information as to exactly when a wire connected to it is touched by a conductor. When such contact occurs the electro-fink reacts (instantaneously, we are supposing) by making the wire live for the duration of the contact. In the absence of any contact the wire is dead.
>
> . . . Consider a time when the wire is untouched by a conductor, for example t_1. *Ex hypothesi* the wire is not live at t_1. But the conditional

4. "A mistake one makes in analyzing, defining, or paraphrasing a statement p or in giving necessary and sufficient conditions for the truth of that statement, by presenting its truth as dependent, in at least some specified situations, upon the truth (falsity) of a subjunctive conditional, ϕ, of the form: 'If state of affairs a were to occur then state of affairs b would occur', when (Version 1) one has failed to notice that the truth value of p sometimes depends on whether a actually occurs and does not merely depend upon the truth value of the analysans, definiens, paraphrase or list of necessary and sufficient conditions; moreover, one has failed to notice this because one has overlooked the fact that, in some of the specified situations, (i) conditional ϕ is true (false), (ii) the analysans, definiens, or paraphrase is true or the necessary and sufficient conditions are fulfilled, (iii) state of affairs a does not occur, and (iv) if a were to occur then the occurrence of a or the occurrence of b or their combination (the occurrence of a or the absence of b or their combination) would help make p true, although it is actually false [or help make p false, although it is actually true and although the analysans, definiens, or paraphrase would remain true or the necessary and sufficient conditions would remain fulfilled] or the states of affairs in question would occur at least partly because p would be true [or would be false] or because of what would make p true [or would make p false] or would together with the truth [or falsity] of p form at least part of a reason for some other occurrence; or (Version 2) one has overlooked the fact that, in some of the specified situations, statement p is actually true but if a were to occur then the occurrence of a would be at least part of what would make b absent (make b occur) or a would occur at least partly because of the absence of b (occurrence of b) or because of what makes b absent (makes b occur), or would together with the absence of b (occurrence of b) form at least part of a reason for some other occurrence." (Shope 1978, 402–3)

> ["If the wire were touched by a conductor, electrical current would flow from the wire to the conductor"] is true of the wire at t_1. (Martin 1994, 2–3)

Again, our point is not to assess the status of one or another specific, alleged example of conditional fallacy. We are interested rather in the status of arguments against the very possibility of conditional analyses. Our method, therefore, will not be primarily to respond directly to such examples. Instead we proceed more generally, by developing a general characterization of the conditional fallacy and by making clear in exactly what sense it is a *fallacy*. From that starting point, we proceed to examine the prospects for using the conditional fallacy as a generic tool for defeating conditional analyses. Our goal is not to give instruction on the proper production of conditional analyses of particular dispositional properties, but rather to examine whether the very idea of conditionality makes such analyses subject to a characteristic style of objection and hence unsuitable for use in analysis. The result of this examination will be a picture of what conditionals must be like for such an analysis to be free of fallacy. This result can be taken negatively via the additional assumption that no plausible conditional could be like *that*. Or it can be taken positively as a guide toward a philosophically fruitful conditional.

1. What Is the Conditional Fallacy?

The general form of a conditional analysis of a disposition is:

- **COND** $\Box(d \leftrightarrow (c \Rightarrow m))$

for some disposition d, some conditions of manifestation c, some manifestation m,[5] and some conditional $\Rightarrow$. Much of the literature on the conditional fallacy has taken for granted that $\Rightarrow$ is the Lewis (1973) counterfactual conditional. However, since our discussion of the conditional fallacy will focus on its implications for the logic of the analyzing conditional, we will adopt a more abstract perspective from which we make no such specific assumptions about the conditional employed.

We follow the lead of F. P. Ramsey (1931) and Gerhard Gentzen (1969) in taking as essential to conditionality the thought that conditionals serve to introduce antecedent-oriented circumstances and then per-

5. Strictly speaking, some claim d that a certain object possesses a certain dispositional property, some claim c that certain conditions of manifestation obtain, and some claim m that certain manifestations occur.

form on those circumstances consequent-determined tests.[6] This conception of conditionality takes as fundamental two inference patterns concerning the introduction and exploitation of conditionals. We will assume that $\Rightarrow$ supports a form of conditional proof that proceeds by introducing some auxiliary assumption ϕ, deriving ψ under this auxiliary assumption, and then inferring, without the assumption, $\phi \Rightarrow \psi$. If $\Rightarrow$ is a modal conditional of some sort (as we will throughout assume it is), then conditional proof proceeds on the semantic assumption that $\phi \Rightarrow \psi$ holds in virtue of the behavior of ψ on some distinguished class of ϕ-oriented worlds—in the case of the Lewis counterfactual, on the closest ϕ worlds according to the system of spheres.

We will thus appeal to rules of $\Rightarrow$I and $\Rightarrow$E. $\Rightarrow$I serves to complete conditional proofs, and hence implements the "test" conception of conditionals. $\Rightarrow$E allows, given $\phi \Rightarrow \psi$ *outside* a conditional proof with conditional assumption ϕ, introduction of ψ into the conditional proof; it thus marks the role of conditionals as "inference tickets." $\Rightarrow$I and $\Rightarrow$E together mark out the basic *conditional* nature of $\Rightarrow$. Any logical connec-

6. See Ramsey 1931, 247:

> If two people are arguing 'If p, will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q,

although we abstract from the epistemic aspects of Ramsey's characterization. See also Gentzen 1969, 80:

> To every logical symbol . . . belongs precisely one inference figure which "introduces" the symbol—as the terminal symbol of a formula—and one which "eliminates" it. . . . The introductions represent, as it were, the "definitions" of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. . . . We were able to introduce the formula $\mathfrak{A} \supset \mathfrak{B}$ when there existed a derivation of $\mathfrak{B}$ from the assumption formula $\mathfrak{A}$.

Thus:

- A Lewis counterfactual introduces a collection of nearest antecedent-supporting worlds, and then tests each of those worlds for support of the consequent.
- An epistemic indicative conditional introduces a hypothetical belief state, created by updating an evaluating agent's belief state by the content of the conditional's antecedent, and then tests that belief state for support of the content of the consequent.
- The material conditional, as a limiting case, introduces either the actual world (if the antecedent in fact holds) or a null situation (if the antecedent fails) and then tests that circumstance for the truth of the consequent (where null circumstances trivially pass all tests).

tive allowing these two rules we will take to be a conditional. By themselves, $\Rightarrow$I and $\Rightarrow$E impose only a very weak logic on conditionals. (They guarantee the validity of the inference from $\phi \Rightarrow \psi$ to $\phi \Rightarrow \theta$ when ϕ, $\psi \vdash \theta$.) Conditionals thus form a diverse family in which sundry additional inferential features can be possessed by varying members of the family. Although the other inference rules we introduce for $\Rightarrow$ will come under critical examination as we proceed, $\Rightarrow$I and $\Rightarrow$E will be presupposed.

Martin's electro-fink example, as he summarizes it in the *second* paragraph quoted above, is indeed one in which the proposed conditional analysis would be false since it purports to describe a situation in which "*ex hypothesi*" (as he puts it) both (a) the wire is not live ($\neg d$) and (b) if the wire were touched by a conductor, then electrical current would flow ($c \Rightarrow m$). But can a general recipe for defeating conditional analyses be extracted from Martin's example? What Martin offers is the possibility of a divergence in truth value of the two sides of the biconditional in **COND**. Of course the falsity of **COND** would follow immediately. However, one does not expose a *generic fallacy* in the project of conditional analysis merely by bruiting the possibility of a thing being such that $\neg d$, but also such that $c \Rightarrow m$. Unless that project has already been defeated, the conditional analyst will have no reason to accept universally the bruited possibility. The possibility may be non-question-beggingly acceptable in the particular electro-fink example (then again, it may not—one may reasonably hold that if the wire is indeed such that if it were touched by a conductor then electrical current would flow, then it is a live wire, and thus that Martin mischaracterizes [in the second quoted paragraph] what electro-finking *is* possible), but this simple approach to the location of a fallacy cannot produce the desired recipe.[7]

7. Nevertheless, Fara (2005, 45) does take this reading of the electro-fink example "to show that the simple conditional account is just wrong":

> Suppose that in its natural state the wire is not disposed to conduct electricity (it's made of rubber, say). An electro-fink is attached, making true the conditional "if the wire were touched by a conductor, it would conduct electricity." In order to save the simple conditional account one would have to defend the claim that, just by being attached to the electro-fink, the wire gained the disposition to conduct electricity. It is hard to imagine how this claim could be defended.

But it may in fact be plausible to suppose that "just by being attached to the electro-fink, the wire gained the disposition to conduct electricity." It's worth keeping in mind here one of the reasons for which we care about the category of dispositional properties in the first place. We might, for example, be engaged in practical reasoning regarding the wire, trying to determine whether certain electrical configurations would pose a safety

But another, more persuasive, recipe can be extracted from Martin's electro-fink example, especially from his description of that example in the *first* paragraph quoted above. We can drop the insistence, question-begging when generalized, that the wire be both dead and such that if it were touched by a conductor, electricity would flow, and instead say only that the wire is both dead and such that if it were touched by a conductor, it would be *live*. This description of finking matches Lewis 1997's account of how dispositions and lacks of dispositions can be finkish and requires no direct rejection of the conditional analysis, but instead only the relatively modest claim that dispositional properties can "come and go" under various circumstances—including, crucially, the circumstance of the presence of the conditions of manifestation of the disposition. So far, there seems to be nothing the conditional analyst ought to, or plausibly could, object to. But if it were then to *follow*, in conjunction with **COND**, that the wire was in fact live, a problem for the conditional analysis would have emerged.

We thus offer the following general characterization of the conditional fallacy. Suppose there is some set Σ of claims such that the following two conditions are met:

1. Σ in conjunction with **COND** entails some conclusion ϕ that makes some assertion about the disposition-possessing object (so, for us, paradigmatically ϕ is either d or $\neg d$). This is the *inferential condition.*
2. Σ together with $\neg\phi$ form an independently plausible scenario. This is the *possibility condition.*

If both the inferential and the possibility conditions are met, we have reason to reject **COND**—the compossibility of Σ and $\neg\phi$ is incompatible with the argument Σ, **COND** $\vdash \phi$. The *fallacy* committed by the condi-

---

risk. Live wires might require insulation, whereas dead wires would not. But in such a context, surely it would be right to categorize the wire, electro-finked in this manner, as live. It ought to be insulated since it would otherwise pose a safety risk.

Fara (2005, 47) says that in order to resist the electro-fink example, as he characterizes it, "one would have to insist that the presence or absence of an electro-fink makes an actual difference to the dispositions of a wire, not just a counterfactual one." Implicit in this remark is the thought that the counterfactually deferred claim *if the wire were touched by a conductor, it would be live* is more plausible than first-order claims about the dispositions of the wire. Below we use a version of this thought to present a stronger version of the electro-fink example—one that shows that even the more plausible deferred claim threatens to lead to trouble for the conditional analyst.

tional analyst is that of neglecting the (perhaps obscure, until brought to explicit attention) compossibility of Σ and $\neg\phi$—thus the "overlooking" and "failing to notice" terminology used by Shope.

On one version of the electro-fink example, Σ is simply $c \Rightarrow m$, and ϕ is d. The inference Σ, **COND** $\vdash \phi$ is, of course, trivial, so the inferential condition is met. However, Σ and $\neg\phi$ fail to form,[8] in the eyes of the conditional analyst, an independently plausible scenario, since they are together the straightforward rejection of the conditional analysis. The more persuasive electro-fink argument takes Σ to be $c \Rightarrow d$. The possibility condition is now less controversially satisfiable. The inferential condition, on the other hand, becomes more demanding, requiring $c \Rightarrow d$, COND $\vdash d$ in order to create problems for the conditional analyst. The justifiability of the inferential condition will accordingly be a focus as we proceed.

The distinction between these two styles of conditional-fallacy argument can be brought out by considering some further prominent examples from the literature:

- Johnston's case of a fragile cup with "a support which when placed inside the glass cup prevents deformation so that the glass would not break when struck." (Johnston 1993, 120)
- Lewis's briefly mentioned case (1997), further pursued by Bird (1998), of a poison taken simultaneously with its antidote, such that the poison would not kill if it were ingested.
- Fara's case of a barrel, parked by a brick, which is "disposed to roll when pushed" (Fara 2005, 54) but such that if it were pushed, it would not roll.

Each of these examples purports to give a case in which the presence of a dispositional property is combined with the absence of a proposed analyzing conditional feature. As taken, each thus follows the first interpretation of Martin's electro-fink (that following the wording of the *second* paragraph quoted from Martin above)—each takes Σ to be $c \Rightarrow \neg m$ and ϕ to be $\neg d$. These examples thus attack conditional analyses by proposing *explicit counterexamples* to them. When these counterexamples are plausible—and we have suggested already that they will not always be, especially from the point of view of the conditional analyst—they are

8. *Likely* fail to form an independently plausible scenario, in the case of a particular proposed conditional analysis of a particular disposition (else the conditional analyst would not have offered that particular analysis). *Certainly* fail to form an independently plausible scenario, in the case of the generic project of offering conditional analyses.

of course decisive in refuting a proposed conditional analysis. Were the philosophical project specifically to show that "if it were struck, it would break" was inadequate as an analysis of fragility, Johnston's internally supported cup might well do the job.

But the conditional fallacy strategy is out for bigger game—it seeks, again, a *generic fallacy* in the very idea of conditional analysis of dispositional properties. No particular counterexample can bring down this game. In response to the proffering of particular (convincing) counterexamples, the conditional analyst can respond by:

1. Conceding that the particular conditional analysis proposed for the dispositional property was the *wrong analysis.* Perhaps fragility is associated with the conditional "if struck *while unsupported,* it would break," rather than the first draft proposal. The Johnston counterexample may successfully bring out the need for the revised conditional, but it does not show that *no* conditional can succeed. Thus Lewis (1997, 153) says that his case of a poison taken together with its antidote shows:

   > The specifications both of the response and of the stimulus stand in need of various corrections. To take just one of the latter corrections: we should really say if ingested without its antidote. Yet the need for this correction to the analysis of poison *teaches no lesson about the analysis of dispositionality in general.* (emphasis added)

2. Conceding that the particular dispositional property at issue does not have a conditional analysis. Perhaps *fragility* lacks a conditional analysis—it does not yet follow that, for example, *solubility* does not have a conditional analysis.

What is lacking here, of course, is a reason to think that the strategy pursued in the particular cases generalizes and provides a recipe for defeating conditional analyses. But a conditional fallacy argument that takes Σ to be $c \Rightarrow \neg m$ and ϕ to be $\neg d$, as in these cases, cannot provide such a recipe. That argument says in effect that conditional analyses always fail because they are always subject to counterexamples; it so far gives no reason to think that the counterexamples are always forthcoming.

We will thus consider a variety of more indirect conditional-fallacy arguments that we think provide better prospects for finding a persuasive general recipe. We distinguish two strategies for the construction of such arguments:

- Following the **Johnston Strategy**, conditional-fallacy arguments focus on situations in which, in certain anomalous situations, placing an object in a disposition-manifesting situation gives rise to some disruption in the manifestation-related features of the object (so, in one simple case, removing from the object the conditional connection between *c* [being in the conditions of manifestation] and *m* [exhibiting the manifestation]).
- Following the **Lewis Strategy**, conditional-fallacy arguments focus on situations in which, in certain anomalous situations, placing an object in a disposition-manifesting situation gives rise to some disruption in the dispositional features of the object (so, in one simple case, removing from the object the dispositional property *d*).[9]

A conditional-fallacy argument that holds out hope of full generalization must place some distance between conditions of manifestation and the manifestation. In the simplest cases, the Lewis Strategy is thus more promising since it uses change to the dispositional property *d* to provide that distance. Thus in section 2 we construct our arguments accordingly, adapting some of the examples and classifications of Johnston and others pursuing the Johnston Strategy to the structure required by the Lewis Strategy. Then in section 4 we return to arguments using the Johnston Strategy, in which the requisite distance is achieved by appeal to a suitable intermediary.[10]

9. Thus the wording of the second paragraph quoted above from Martin suggests a pursuit of the Johnston Strategy, whereas the wording of the first paragraph suggests a pursuit of the Lewis Strategy. Our own reading is that, despite some imprecision in the language, the dominant approach in Martin is the Lewis Strategy (thus his introduction of the paper with the assertion that "the dispositions of a thing can change"). Note also that Lewis's case of the poison taken together with its antidote represents his consideration of the Johnston Strategy, not an application of the Lewis Strategy.

10. That said, we will in this note set out the analysis of some simple conditional-fallacy arguments following the Johnston Strategy. These remarks will draw heavily on the discursive framework to be established in sections 2 through 4 of the essay and are not intended to be self-standing at this point. The analyses are provided for those who feel the draw of some simple instance of the Johnston Strategy and will wish, in light of the subsequent discussion, to see the formal details worked out.

First, a simple version of the argument, following the structure of the examples above, proceeds as follows:

2. Masking and Mimicking

The potential varieties of the conditional fallacy are thus as varied as the potential premises Σ entailing, together with **COND**, facts about the distribution of dispositional properties. We begin by focusing on two varieties that have received considerable attention in the literature.[11]

---

1. **COND**	A
2. $c \Rightarrow \neg m$	A
3. $\neg(c \Rightarrow m)$	**Exclusion**, 2
4. $\neg d$	**Taut**, 3

This argument, requiring only **Exclusion**, places fewer demands on the logic of $\Rightarrow$ than any we consider in the main text but, for the reasons discussed above, is unpromising as a *general* recipe for defeating conditional analyses.

Second, another argument following the Johnston strategy appeals to the possibility of items having a dispositional property *d*, being exposed to the conditions of manifestation *c*, but failing to produce the manifestation *m*. The argument then proceeds as follows:

1. **COND**	A
2. $c \wedge \neg m$	A
3. d	A (for reduction)
4. $c \Rightarrow m$	**Taut**, 1,3
5. c	**Taut**, 2
6. m	**MP**, 4,5
7. $\neg m$	**Taut**, 2
8. d	Reductio, 5,6

The central inferential assumption of this argument is that the conditional $\Rightarrow$ supports modus ponens; it thus occupies the same location in logical space as the mimicking argument to be given in section 2.2.

11. The categories of masking and mimicking derive from Johnston (1993), though we have converted them into instances of the Lewis Strategy. Johnston also identifies a category of *altering* cases. Both masking and altering cases (suitably adjusted) involve the loss of the dispositional property in the counterfactual circumstances of manifestation; the difference between the two is that in altering cases, the loss of the dispositional property is effected by a change in the intrinsic structure of the object, whereas in masking cases, the loss of the dispositional property is effected by an extrinsic change in the circumstances of the object. (A complete characterization of logical space would require a similar distinction within mimicking cases between those that require alteration of the intrinsic structure of the object and those that do not, a distinction that Johnston

In a *masking* situation, an object possesses a dispositional property, but is such that if the conditions of manifestation of that disposition were realized, the object would not have the disposition. The relevant Σ is thus the sentence $c \Rightarrow \neg d$. Martin's "reverse-cycle" electro-fink is a masking example: in it, the wire is live, but if it were to touch a conductor, it would become dead, and its disposition is thus masked. A fragile glass watched over by a divine agent that renders it no longer fragile whenever it is struck has its fragility masked. Masking arguments will then require the inference $c \Rightarrow \neg d$, **COND** $\vdash \neg d$.[12]

In a *mimicking* situation, an object lacks a dispositional property, but is such that if the conditions of manifestation of the lacked dispositional property were realized, the object would have the disposition. The relevant Σ is thus the sentence $c \Rightarrow d$. Martin's original electro-fink is a mimicking example: in it, the wire is dead, but if it were to touch a conductor, it would become live, and it thus mimics liveness. A nonfragile glass watched over by a divine agent that renders it fragile whenever it is struck mimics fragility. Mimicking arguments will then require the inference $c \Rightarrow d$, COND $\vdash d$.

2.1 The Masking Argument

Can a masking argument meet the required inferential condition? The answer depends, of course, on the inferential details of the conditional used in the conditional analysis. We will thus give a schematic proof and then consider what inference rules the conditional would have to support to make sense of certain steps in the proof. Consider the proof presented in figure 1. (**Taut** is used as a general rule of tautological consequence, including simple consequences of modal logic.)[13]

---

does not make.) Since the distinction between intrinsic and extrinsic alteration will play little role in our examination of the logic of the conditional fallacy (although see section 4.3 for some discussion), we consistently group altering cases together with mimicking cases. We thus need not attend to Johnston's distinction between "bare" and "constituted" dispositions. We should note, however, that Johnston indicates in a note that the relevant passages "no longer represent" his thinking.

12. In the presence of **COND**, either of these occurrences of d could be replaced with $c \Rightarrow m$.

13. The double indentation lines mark conditional subproofs. We assume throughout that necessary truths (such as **COND**) can be imported into conditional subproofs.

1. **COND**	A
2. $c \Rightarrow \neg d$	A
3. d	A (for reductio)
4. $c \Rightarrow m$	**Taut**, 1,3
5. c	A (for $\Rightarrow$I)
6. m	$\Rightarrow$E, 4,5
7. $c \Rightarrow m$	$?_1$, 5,6
8. d	**Taut** 1,7
9. $c \Rightarrow d$	$\Rightarrow$I, 5,8
10. $\neg(c \Rightarrow d)$	$?_2$, 2
11. $\neg d$	Reductio, 3,9,10

Figure 1

Given rules to plug the gaps marked by the two question marks, this derivation yields the inferential condition needed by the masking argument.

- The gap marked by '$?_1$' requires a rule allowing the inference from c and m to $c \Rightarrow m$. For the Lewis counterfactual, this inference corresponds to the semantic constraint of strong centering—the requirement that each world be the unique member of the minimal sphere in the system of spheres about that world. Henceforth we refer to this rule as **C**.
- The gap marked by '$?_2$' requires a rule allowing the inference from $\phi \Rightarrow \neg\psi$ to $\neg(\phi \Rightarrow \psi)$. Henceforth we refer to this rule as **Exclusion**. Of course, not every conditional will treat $\phi \Rightarrow \psi$ and $\phi \Rightarrow \neg\psi$ as contraries. If, for example, $\Rightarrow$ is the Lewis counterfactual, they are contraries only on the condition that $\Diamond\phi$.

The use of **Exclusion** thus typically amounts to the imposition of some further constraint on the possibility condition. Although this additional constraint—the requirement, for example, that the glass can be struck—is surely wholly uncontroversial, it can also be rendered unnecessary. Masking arguments can begin with $\neg(c \Rightarrow d)$ rather than with $c \Rightarrow \neg d$. On any conditional obeying **Exclusion**, the former will be a weaker assumption than the latter. Call the argument from **COND** and $c \Rightarrow \neg d$ to $\neg d$ *inefficient masking* and the argument from **COND** and $\neg(c \Rightarrow d)$ to $\neg d$ *efficient masking*. Efficient masking shows that trouble looms for the conditional analyst on the mere assumption that the dispositional property *need not* be preserved under counterfactual occurrence of the conditions of manifestation, without going as far as the assumption that the dispositional property *would not* be preserved under those conditions.

Alternatively, an efficient masking argument can proceed as presented in figure 2.

1. **COND**	A
2. $\neg(c \Rightarrow d)$	A
3. d	A (for reductio)
4. $c \Rightarrow m$	**Taut**, 1,3
5. $c \Rightarrow (c \Rightarrow m)$	?, 4
6. $\neg(c \Rightarrow (c \Rightarrow m))$	**Taut**, 1,2
7. $\neg d$	Reductio, 3,5,6

Figure 2

This argument requires the following inference rule for $\Rightarrow$:

- The gap marked by '?' requires a rule allowing a conditional to be expanded through a repetition of its antecedent. The requisite rule thus allows the inference from $\phi \Rightarrow \psi$ to $\phi \Rightarrow (\phi \Rightarrow \psi)$. Henceforth we refer to this rule as **Expansion**. In the context of a Lewis-style counterfactual, **Expansion**, like **C**, is a manifestation of the semantic constraint of strong centering.[14] Without strong centering, there can be a world w_1 whose unique closest q world is w_2. If w_2 is an r world, but shares its own minimal sphere with a $q \wedge \neg r$ world w_3, then $q \Box\!\!\rightarrow r$ will hold at w_1, but fail at w_2, leading to the failure of $q \Box\!\!\rightarrow (q \Box\!\!\rightarrow r)$ at w_1—a counterexample to **Expansion**.

This efficient masking argument requires of $\Rightarrow$ only that it support the rule **Expansion**. An efficient version of the first masking argument

14. Given $\Rightarrow$I, $\Rightarrow$E, and **C**, **Expansion** is straightforwardly derivable.

requires only the rule **C**. Since, in a broadly Lewisian framework for the conditional, both **C** and **Expansion** are manifestations of strong centering, it is tempting to think that the success of masking arguments is intimately connected to the imposition of strong centering on the conditional.[15]

2.2 The Mimicking Argument

In the mimicking argument, we derive the actual possession of a dispositional property from the combination of **COND** and the mimicking claim $c \Rightarrow d$ (the claim that if the conditions of manifestation were to obtain, the object would have the relevant disposition). Thus consider the argument presented in figure 3.

1. **COND**	A
2. $c \Rightarrow d$	A
3. c	A (for $\Rightarrow$I)
4. d	$\Rightarrow$E, 1,3
5. $c \Rightarrow m$	**Taut**, 1,4
6. m	?, 3,5
7. $c \Rightarrow m$	$\Rightarrow$I, 3,6
8. d	**Taut**, 1,7

Figure 3

This derivation requires the following inference rule governing the $\Rightarrow$ conditional:

- The rule filling the gap marked by '?' allows the application of modus ponens to $\Rightarrow$; henceforth we refer to this rule as **MP**. **MP** should not be conflated with the earlier $\Rightarrow$E; it makes a more controversial assumption about $\Rightarrow$. **MP** will, for example, fail on a Lewis counterfactual that lacks the semantic constraint of weak centering. If the world of evaluation is not a member of its own minimal sphere, then the fact that it is a ϕ world, together

15. Given his own ambivalence about strong centering, it is surprising that Lewis does not mark this.

with the truth of $\phi \mathbin{\Box\!\!\rightarrow} \psi$, will not entail ψ. $\Rightarrow$E, however, remains valid on all Lewis counterfactuals, whether weakly centered or not.

Alternatively, we can reason as in figure 4.

1. **COND**	A
2. $c \Rightarrow d$	A
3. $c \Rightarrow (c \Rightarrow m)$	**Taut**, 1,2
4. $c \Rightarrow m$	?, 3
5. d	**Taut**, 4

Figure 4

This argument requires the following inference rule for $\Rightarrow$:

- To fill the gap marked by '?', an inference rule is needed allowing the inference from $\phi \Rightarrow (\phi \Rightarrow \psi)$ to $\phi \Rightarrow \psi$, contracting a repeated antecedent. Henceforth we refer to this rule as **Contraction**. For a Lewis counterfactual, **Contraction** requires weak centering. Given weak centering, **Contraction** follows from the validity of modus ponens across $\Box\!\!\rightarrow$. Without weak centering, let w_1 and w_2 be two worlds such that w_2 is the unique world closest to w_1, and w_1 is the unique world closest to w_2. If w_1 is a $q \wedge r$ world, then $q \mathbin{\Box\!\!\rightarrow} r$ holds at w_2, so if w_2 is a $q \wedge \neg r$ world, then $q \mathbin{\Box\!\!\rightarrow} (q \mathbin{\Box\!\!\rightarrow} r)$ holds at w_1 but $q \mathbin{\Box\!\!\rightarrow} r$ does not, and **Contraction** fails.

Both versions of the mimicking argument thus rely on an inference rule that, in a broadly Lewisian framework for the conditional, is intimately connected to the imposition of weak centering.

3. The Heart of the Conditional Fallacy

A successful conditional-fallacy argument has two prongs. To satisfy the inferential condition, it must adduce some Σ that in conjunction with **COND** yields substantive conclusions about the distribution of dispositional properties. To satisfy the possibility condition, it must make independently plausible the possibility of that same Σ holding without that particular distribution of dispositional properties. The deductions set out above trace the impact of the first prong by showing that the simplest popular styles of conditional-fallacy argument—those that employ either masking or mimicking considerations—must assume inference rules flowing from strong or weak centering (respectively) of the conditional in order to meet the inferential condition. Since the Lewis counterfactual does have both strong and weak centering, it becomes a natural target for conditional-fallacy arguments. We take up in section 5 below the question of what a conditional lacking strong and weak centering would look like.

What of the possibility condition? Why should masking and mimicking situations be plausible? The thought here is that possession of dispositional properties should be a contingent matter—the sort of thing that can come and go as various counterfactual situations obtain. Thus Martin (1994)'s opening remark is "the dispositions of a thing can change," and Lewis (1997, 143) begins his discussion of finkishness with the claim that "dispositions come and go, and we can cause them to come and go." In particular, masking and mimicking scenarios require for their plausibility that there be a difference between:

(A) An object's possessing a dispositional property, and
(B) An object's being such that, if it were placed in a disposition-manifesting situation, it would have that dispositional property.

Masking scenarios rely for their intuitive plausibility on the thought that there is no inference from (A) to (B)—if there were such an inference, then the disposition could not be masked.[16] Mimicking dispositions rely for their intuitive plausibility on the thought that there is no inference from (B) to (A)—if there were such an inference, then the mimicking of the disposition would entail the actual presence of the disposition.

But notice the close similarity between:

- The inference from (A) to (B) and the inferential pattern of expansion. (Both rely on a move from what is true in the actual world to what is true in worlds relevant to *c*-antecedent counterfactuals.)

16. Thus Wright (1992, 137) says:

> While lacking the appropriate disposition, it might be that the object, under certain conditions, would be so altered by the implementation of *C*-circumstances that it acquired it; and that would not vindicate ascribing the disposition to it in the actual, non-*C*-circumstances.

Wright's comment is directed toward *altering* cases; see note 11 for the relation between such cases and masking cases.

- The inference from (B) to (A) and the inferential pattern of contraction. (Both rely on a move from what is true in worlds relevant to *c*-antecedent counterfactuals to what is true in the actual world.)

Given **COND** the close similarity becomes outright identity. This similarity suggests a *diagnosis* of the conditional fallacy. When a conditional-fallacy argument attempts to meet the possibility condition, it does so by exploiting the failure of certain inferential features of dispositional properties to craft scenarios that strike us as plausible. (Thus masking arguments exploit the failure of [an analogue of] expansion; mimicking arguments exploit the failure of [an analogue of] contraction.) When that argument attempts to meet the inferential condition, it does so by presupposing that the conditional exhibits those same inferential features whose failure in the case of dispositions was just exploited to meet the possibility condition, and then observing the discrepancy between the two.

Something indeed goes wrong here. But the proper lesson is not yet that the conditional analysis should be rejected. Rather, we extract the following general moral:

> **(General Moral)** The conditional used in a conditional analysis should not be loaded down with inferential features that we reject for dispositions.

If the inferential behavior of the conditional is properly adjusted to the inferential behavior of the dispositions, then the attempted derivation of ϕ from Σ and **COND** will fail due to an unavailable inference rule in *exactly those cases* in which the unavailability of that inference rule allows a plausible Σ and $\neg\phi$ scenario.

The General Moral suggests that conditional-fallacy arguments trade on a perhaps artificially imposed difference between the inferential capacities of dispositions and the inferential capacities of appropriate conditionals. Opponents of conditional analyses will, of course, deny that the difference is artificial and claim instead that the conditional-fallacy arguments merely elicit real inferential differences that show the difference between dispositional and conditional properties. Some purchase on the resulting disagreement can be achieved by applying the following test:

> **(The Canonical Test)** Suppose a conditional-fallacy argument is advanced, with a certain justification Ξ being given, by way of meeting the possibility condition, for the compossibility of Σ and ϕ. Let Σ', ϕ', and Ξ' be the result of replacing all mention of the dispositional property with mention of the purported analyzing conditional. If the resulting justification Ξ' suffices to make plausible the compossibility of Σ' and ϕ', then the conditional-fallacy argument fails.

Suppose a certain masking scenario is justified by observing *that the glass is fragile* does not entail *that were the glass struck, it would be fragile.* We should then ask ourselves if we also agree that *that were an object struck, it would break* does not entail *that were an object struck, it would be such that if it were struck, it would break.*

Dispositions can come and go, we are told. *But counterfactual features too can come and go.* Lewis (1997, 143) gives us examples of the ephemerality of dispositions:

> Glassblowers learn to anneal a newly made joint so as to make it less fragile. Annoyances can make a man irascible; peace and quiet can soothe him again.

But consider the result of employing the Canonical Test and making the relevant replacement in Lewis's hortatory language:

> Glassblowers learn to anneal a newly made joint so as to make it less [prone to break if struck]. Annoyances can make a man [such that if he were bothered, he would become angry]; peace and quiet can [make him] again [such that if he were bothered, he would remain calm].

The plausibility of the remarks is not obviously diminished by these substitutions. Accordingly, the proffered conditional-fallacy argument could be recast to show the impossibility of analyzing "if the glass were struck, it would break" via the conditional "if the glass were struck, it would break," as in the following similarly altered passage from Lewis (1989, 117):

> Imagine that a surface now has just the molecular structure that [makes things such that they would reflect light if light were to fall on them]; but that exposing it to light would catalyze a swift chemical reaction and turn it into something [not such that it would reflect light if light were to fall on it]. So long as it's kept in the dark, is it [such that it would reflect light if light were to fall on it]?—I think so; but its [being such that it would reflect light if light were to fall on it] is what Ian Hunt once called

> a 'finkish' [counterfactual feature], one that would vanish if put to the test. (So a simple counterfactual analysis of [counterfactuals] fails.)

If a conditional-fallacy argument fails the Canonical Test, then the mismatch between dispositional and conditional inferential potency to which the General Moral draws attention should be rejected. This rejection can proceed in two ways:

1. The inferential capacities of the disposition can be expanded to match those posited of the conditional. If this route is taken, the previous justification of the possibility condition will be rejected, and what we will broadly call the *finking situation* will now be rejected as impossible. In a mimicking argument, for example, this route will amount to holding that a glass that is such that if it were struck, it would be fragile is *already* a fragile glass. One might call this a bullet-biting response, but some bullets should be bitten.[17] (To one reading of the electro-fink example, the conditional analyst may respond by biting the bullet and holding that a wire that is such that if it were touched by a conductor, electrical current would flow through it is indeed a live wire.)
2. The inferential capacities of the conditional can be limited so as to match those posited of the disposition.

Different responses will be appropriate for different instances of failure of the Canonical Test, of course. Our primary interest will be in cases in which the second response is appropriate. This response has been largely neglected in the literature on the conditional fallacy; we return to it in section 5 below.

4. Weaker Inference Rules: Intermediaries

Masking and mimicking arguments are only two instances of a general form of argument. Conditional arguments generically can be organized along a rough spectrum corresponding to the strength of the inference rules required of the conditional. At one end of the spectrum lie conditionals with extremely strong inferential structures—for such conditionals, the possibility condition can be met very easily. At the other end of

17. See again note 2—one appropriate response to a *particular* mimicking argument is always to reject the *particular* conditional analysis given. The mere existence of inadequate conditional analyses should not be conflated with the generic failure of the project of conditional analysis.

the spectrum lie conditionals with extraordinarily weak inferential structures—for such conditionals, the required possibility condition will be either apparently impossible or clearly question-begging.[18]

We will focus below on a collection of arguments that appeal to inference rules weaker than, and hence to finking scenarios somewhat more difficult to motivate than, those used in masking and mimicking arguments. These arguments will revolve around the idea of *mediated masking*. Masking cases discussed in the literature often explicitly involve an *intermediary*, such as a divine agent or an electro-fink, that disrupts the usual connection between conditions of manifestation and manifestation. In the masking and mimicking arguments given above, no special inferential role was given to that intermediary; the arguments ran directly on conditional connections between conditions of manifestation and the presence or absence of dispositions. Mediated cases represent a sophisticated application of what we earlier called the Johnston Strategy.

18. The above discussion of masking and mimicking arguments suggested that masking arguments turn on a conditional's support of strong-centering-based inference principles, whereas mimicking arguments turn on a conditional's support of weak-centering-based principles. Since strong centering is a stronger semantic constraint than weak centering and gives rise to stronger inference principles, we should expect masking scenarios to be more plausible than mimicking scenarios. The general preference in the literature for masking over mimicking arguments suggests that this expectation may be correct. Thus Wright (1992, 137) notes that "it is actually rather difficult to think of plausible (non-supernatural) examples of mimicking." Mimicking arguments may be somewhat too prone to the thought that a glass that is such that if it were struck, it would be fragile would thus in fact *break* if struck (being then fragile), and thus be *actually fragile.* (Of course, this line of reasoning can be resisted if weak-centering principles are rejected.)

The issue is clouded by the fact that the strong-centering principle of **Expansion** is equivalent to the permissibility of the weak-centering principle of **Contraction** in *negated* contexts. Any tendency to overlook the distinction between negated and unnegated contexts will thus lead to a conflation of the masking and mimicking arguments. Use of dispositional properties whose negations are also dispositional properties may encourage such a tendency. In Martin's electro-fink example, both *live* and *dead* can be read as dispositional properties. Given that the mimicking inference and the *inefficient* masking inference are isomorphic under replacement of d with $\neg d$, if the negation of every dispositional property is itself a dispositional property (a claim that admittedly seems unlikely), there will be no difference between mimicking and inefficient masking. The efficient masking argument is thus useful for drawing more sharply the difference between masking and mimicking.

4.1 Mediated Masking

Suppose we make explicit the role of the intermediary in the reasoning. Let a *mediated masking* argument be one that takes as the relevant Σ the pair of conditionals $c \Rightarrow i$ and $i \Rightarrow \neg m$, where i is the claim that some appropriate intermediary intervenes.[19] In a fragility case, for example, the two conditionals are *were the glass struck, the divine agent would intervene* and *were the divine agent to intervene, the glass would not break.* We can then reason as shown in figure 5.

1. **COND**	A
2. $c \Rightarrow i$	A
3. $i \Rightarrow \neg m$	A
4. $c \Rightarrow \neg m$	?, 2,3
5. $\neg(c \Rightarrow m)$	**Exclusion**, 4
6. $\neg d$	**Taut**, 1,5

Figure 5

The gap marked by '?' can be filled as follows:

- With a transitivity principle, allowing the inference from $\phi \Rightarrow \psi$ and $\psi \Rightarrow \theta$ to $\phi \Rightarrow \theta$. Henceforth we refer to this rule as **Trans**.

This mediated masking argument is unlikely to worry any proponent of conditional analyses, because the rule **Trans** fails on the relevant sorts of counterfactual conditionals (paradigmatically, it fails on every version of the Lewis counterfactual).[20]

However, there are close variants of the mediated masking argument that will work against many conditionals, including Lewis-style counterfactuals. These variations replace **Trans** with various weakened transitivity principles that are valid for Lewis counterfactuals:

19. Similar mediated mimicking arguments can be constructed from the combination $c \Rightarrow i$ and $i \Rightarrow m$. Some examples involving an intermediary are cases in which Σ contains $c \Rightarrow i$ and $i \Rightarrow \neg d$ (for masking; or $c \Rightarrow i$ and $i \Rightarrow d$, for mimicking). These are not "mediated" cases as we define them here; they do not pursue the Johnston Strategy. They are better seen as possible instances (*pace* an apparent reliance on the **Trans** rule defined below) of the basic cases of masking and mimicking discussed above, which we interpreted as versions of the Lewis Strategy.

20. Another mediated masking argument can be constructed as follows:

A. We might add the assumption $i \Rightarrow c$. (This is the assumption that if the intermediary were to occur, the conditions of manifestation would occur.) The inference rule $\phi \Rightarrow \psi, \psi \Rightarrow \theta, \psi \Rightarrow \phi \vdash \phi \Rightarrow \theta$—that we henceforth refer to as **Weak Trans A**—is valid on all versions of the Lewis counterfactual.

B. We might add the assumption that $i \Diamond\!\!\rightarrow c$. (If the intermediary were to occur, the conditions of manifestation might occur.) The inference rule $\phi \Rightarrow \psi, \psi \Rightarrow \theta, \psi \Diamond\!\!\rightarrow \phi \vdash \phi \Rightarrow \theta$—we henceforth refer to as **Weak Trans B**—is valid on all versions of the Lewis counterfactual.

C. We might replace the assumption $i \Rightarrow \neg m$ with $(i \wedge c) \Rightarrow \neg m$. (If the intermediary and the conditions of manifestation were to occur, the manifestation would not occur.) The inference rule $\phi \Rightarrow \psi, (\phi \wedge \psi) \Rightarrow \theta \vdash \phi \Rightarrow \theta$—which we henceforth refer to as **Weak Trans C**—is valid on all versions of the Lewis counterfactual.

D. We might replace the assumption $i \Rightarrow \neg m$ with the strict conditional $i \prec \neg m$. (In every possible world, if the intermediary occurs, the manifestation does not occur.) Call an intermediary supporting such a strict conditional *fail-safe.*

---

1. **COND**	A
2. $c \Rightarrow i$	A
3. $i \Rightarrow \neg m$	A
4. d	A (for reductio)
5. $c \Rightarrow m$	**Taut**, 1,4
6. c	A (for $\Rightarrow$I)
7. i	$\Rightarrow$E, 2,6
8. m	$\Rightarrow$E, 5,6
9. $\neg m$	?, 3,7

Here we suppose that the contradiction inside the modal subproof suffices to complete the reductio (thereby implicitly presupposing $\Diamond c$). The required inference rule to replace '?' is one that allows use of a $\Rightarrow$ conditional $\phi \Rightarrow \psi$ inside a modal subproof whose assumption is not ϕ, but within which ϕ has been derived. This inference rule is equivalent to the rule **Trans**. Transitivity principles can, in general, be recast as principles controlling the exploitation of external conditionals within $\Rightarrow$I subproofs. Given **MP**, the recasting can be made somewhat more perspicuously in terms of principles controlling the importation of external conditionals into $\Rightarrow$I subproofs.

E. We might replace $i \Rightarrow \neg m$ with the strict conditional $i \prec \neg m$, and also replace the assumption $c \Rightarrow i$ with the strict conditional $c \prec i$. (In every possible world, if the conditions of manifestation occur, the intermediary occurs, and if the intermediary occurs, the manifestation does not occur.) Call an intermediary supporting both of these strict conditionals *attentive* and *fail-safe.*

Each of these variants relies on the fact that although the Lewis counterfactual is not transitive, it supports inference principles that are closely related to transitivity. Given premises tailored to suit any of these inference principles, we can construct mediated masking arguments that do not depend on centering.

Consider the argument presented in figure 6.

1.	**COND**	A
2.	$c \Rightarrow i$	A
3.	$i \Rightarrow \neg m$	A
4.	$i \Rightarrow c$	A
5.	$c \Rightarrow \neg m$	**Weak Trans A**, 2,3,4
6.	$\neg(c \Rightarrow m)$	**Exclusion**, 5
7.	$\neg d$	**Taut**, 1,6

Figure 6

This argument meets the inferential condition for a conditional-fallacy argument against any conditional supporting **Weak Trans A.**[21] To complete the conditional-fallacy argument, the possibility condition must also be met. Thus we need some reason to think that it is jointly possible that:

21. Following the pattern of note 20, the use of **Weak Trans A** here could be replaced with a principle allowing the use of external conditionals within $\Rightarrow$I subproofs. Here the required condition is that the external conditional have an antecedent that is tantamount, in the sense defined in Bonevac 2003, 412, to the assumption of the $\Rightarrow$I subproof. Recasting the argument in this way, we can then see that the justification offered for the possibility condition will concern the failure of some combination of dispositions and conditionals regarding dispositions, all to be exploitable/importable to a suitable finking world, and application of the Canonical Test to this justification will provide clear motivation for abandoning **Weak Trans A**. Similar considerations apply to **Weak Trans B** and **Weak Trans C**.

1. An object is (for example) fragile.
2. A divine agent would act if that object were struck.
3. The object would not break if the divine agent were to act.
4. The object would be struck were the divine agent to act.[22]

What could this reason be?

To understand the dialectic here, consider first how the possibility condition for the earlier simple mediated masking argument (in section 4.1, proceeding via **Trans**) could be met. We can then proceed to determine what would need to be added to justify the possibility condition for mediated masking arguments using **Weak Trans A**. Why should we think there could be an object that was fragile, but that would be protected by a divine agent if struck, where the divine agent is such that the object would not break were the agent to act?

Consider first an illegitimate answer. Suppose the conditional analyst is presented with the following line of reasoning:

> Dispositions come and go. Perhaps the object is fragile *right now*, but it, or its situation, will be crucially altered by the divine agent in striking cases. After that alteration, it loses the dispositional property of fragility, so of course it does not break. Its counterfactual failure to break is thus irrelevant to its current fragility.

The scenario sketched here presents a real challenge to the conditional analyst, but not one that turns on issues of mediation. On one reading, the finking scenario is, in this justification, simply being strengthened to include $c \Rightarrow (i \wedge \neg d \wedge \neg m)$—the success of the intermediary in preventing the manifestation is built into the scenario. Of course, the conditional analyst cannot accept the possibility of *this* situation because $c \Rightarrow \neg m$, and the falsity of **COND**, follow immediately from it. But the conditional analyst *shouldn't* accept this possibility any more than the earlier question-begging form of the electro-fink example.[23] If medi-

22. It may be tempting to rephrase this last as "The divine agent would act only if the object were struck." But it is not clear whether "ϕ only if ψ" should be understood as "$\phi \Rightarrow \psi$" or as "$\neg\psi \Rightarrow \neg\phi$." If the latter, then it would be an illicit replacement—the inference $\phi \Rightarrow \psi \vdash \neg\psi \Rightarrow \neg\phi$ need not hold. (It fails for the Lewis counterfactual, for example.)

23. Perhaps our views on dispositions are supposed to incline us to accept that, so long as the object has certain categorical features (a certain microstructure, perhaps) in the actual world, we are to count it as fragile, *regardless of whether it breaks under any circumstances whatsoever.* Such an inclination would certainly allow the above justification to go through (it would, in fact, allow us to endorse $c \Rightarrow (i \wedge \neg d \wedge \neg m)$), but it is hard

ated masking cases are to be persuasive, the intermediary must carry an inferential load and not just serve as a third wheel in a simple assertion of failure to manifest.

Perhaps, then, we are meant to reason from the failure of disposition in the counterfactual conditions of manifestation to the failure of manifestation. This would be to replace the question-begging $c \Rightarrow (i \wedge \neg d \wedge \neg m)$ with the weaker $c \Rightarrow (i \wedge \neg d)$. But i is again superfluous here, and the necessary inference follows the pattern of simple (unmediated) masking cases above (see section 2.1). Conditional analysts whose accounts have survived simple masking cases have found some way to deal with the thought that fragility may be lost whenever the object is struck. Perhaps they have given up **Expansion** for their conditional; perhaps they have bitten the bullet. If they have bitten the bullet, they will presumably be unmoved by the current justification. If they have abandoned **Expansion**, then they accommodate the possibility of the loss of fragility by observing that the counterfactual loss of fragility does not entail the counterfactual failure to break (this entailment—with adjustment for the distinction between efficient and inefficient masking—is precisely what **Expansion** would have ensured).

The conditional analyst who rejects **Expansion** rejects also the "of course it does not break" of the above justification. It is the burden of mediated masking arguments to find a way to foist that conclusion onto the conditional analyst through independently acceptable means. The right way for the proponent of the mediated masking argument to reason is thus the following:

> Dispositional properties don't always manifest successfully; sometimes interfering circumstances arise. So there could be an object that was fragile, but whose fragility would be interfered with should conditions of manifestation occur. Moreover, it could be that that interference would succeed.

But a conditional analyst who rejects **Trans** can simply accept all of this and observe that it might be that the appearance of conditions of manifestation are much less plausible than the appearance of the interfering

to see why it is a reasonable inclination. What motivates treating certain categorical features as sufficient for fragility if there is no link between the possession of those features and possible breakage? Appeal to categorically similar items would seem insufficient since such an appeal would seem to leave us with no tools for determining what categorical differences (in intrinsic or relational features) mattered to dispositional properties. See section 4.4 for further discussion.

agent, and thus that the truth of $i \Rightarrow \neg m$ might be confined to worlds much closer to the actual world than the nearest c worlds, allowing for $i \wedge m$ worlds in the more distant c regions. (This, of course, is the standard Lewis story for the failure of transitivity.) Application of the Canonical Test to the above justification will show that the resulting justification brings out the unacceptability of **Trans**.

Now consider how this justification could be extended to the argument employing **Weak Trans A**. The required justification could proceed:

> Dispositional properties don't always manifest successfully; sometimes interfering circumstances arise. So there could be an object that was fragile, but whose fragility would regularly be interfered with should conditions of manifestation occur. It could also be that were interference to occur, it would be in a case in which the conditions of manifestation occurred—*the nearest conditions-of-manifestation worlds and the nearest interference-worlds could be the very same.* Moreover, it could be that that interference would regularly succeed.

Why might one be inclined to accept such a justification? Why, in particular, might one think that an object that failed to break with such reliability when struck was nevertheless fragile? Here are two possible reasons:

1. One might take the *regularity* appealed to in the justification to involve less than universality over the relevant space of worlds. If fragility's being regularly interfered with means just that some threshold percentage or measure of the nearest c worlds are i worlds, then the chaining of regularities becomes relevant. Some percentage (for example) of the nearest c worlds is made up of i worlds; some percentage of the nearest i worlds, $\neg m$ worlds—the resulting connection between c and $\neg m$ can be quite weak and certainly may fail to support $c \Rightarrow \neg m$. Thus if $\Rightarrow$ is a *threshold conditional*, rather than a *universalizing conditional*, the justification can be acceptable.

   But a threshold conditional will not support the inference rule **Weak Trans A**, and hence will not allow the inferential condition of this particular conditional-fallacy argument to be met.

Thus, if one accepts the General Moral, a careful examination of the justification for the possibility condition permits the conclusion that if that justification is to be endorsed, one's view of the conditional ought to be adjusted to match. But one thereby blocks the proof needed for the inferential condition. Again, application of the Canonical Test helps to reveal this—recasting the above argument entirely in terms of $\Rightarrow$ transfers whatever persuasive force the justification had to an argument against **Weak Trans A**.

2. One might drop the italicized phrase *the nearest conditions-of-manifestation worlds and the nearest interference-worlds could be the very same* from the justification. Doing so would require a conception of the conditional that allowed $c \Rightarrow i$ and $i \Rightarrow c$ without mandating the rejected coincidence of worlds. Such a conception requires, if a universalizing rather than threshold conditional is to be employed, abandoning deep features of the Lewis counterfactual. Suppose the conditional employs some method of picking out the relevant antecedent worlds—this can be thought of as a selection function f mapping a pair of a world and an antecedent proposition (thought of as a set of worlds) to a set of relevant worlds. Given a space of worlds W and a selection function $f: W \times \wp(W) \mapsto \wp(W)$, we then have (in the style of, for example, Chellas 1980), truth conditions of the form:

   - $w \vdash \phi \Rightarrow \psi$ if and only if $f(w, [[\phi]]) \subseteq [[\psi]]$

   One particularly simple method of providing a selection function—the one used by the Lewis counterfactual—is built off of a global structuring of the worlds (a system of spheres) and the truth or falsity of the antecedent at given worlds. Two natural formal consequences of the simple method are:

   a) $f(w, [[\phi]]) \subseteq [[\phi]]$
   b) Given $\phi \Rightarrow \psi$ and $\psi \Rightarrow \phi$ at a world w, it follows that $f(w, [[\phi]]) = f(w, [[\psi]])$.

   Call a conditional satisfying condition (a) a *bounded* conditional and a conditional satisfying condition (b) a *proximity* conditional. Given an understanding of the conditional according to which it is not a proximity conditional, we can thus drop the italicized phrase from the justification and understand the situation in the same way that the earlier justification for the simple transitivity argument was understood.[24]

Again, the result of going in this direction will be the rejection of Weak Trans A. The possibility condition is thereby secured, but only (in keeping with the General Moral) at the cost of giving up the tools needed to meet the inferential condition.

The situation with **Weak Trans B** is similar. We can give almost exactly the same argument, as shown in figure 7.

1. **COND**	A
2. $c \Rightarrow i$	A
3. $i \Rightarrow \neg m$	A
4. $i \mathrel{\Diamond\!\!\rightarrow} c$	A
5. $c \Rightarrow \neg m$	**Weak Trans B**, 2,3,4
6. $\neg(c \Rightarrow m)$	**Exclusion**, 5
7. $\neg d$	**Taut**, 1,6

Figure 7

The justification of the possibility condition will have to proceed along the same lines as above, and the same responses are available. Moving to a threshold conditional makes the justification acceptable even more easily than before since the occasional exceptions to $c \Rightarrow i$ may happen to coincide with the occasional instances of $i \mathrel{\Diamond\!\!\rightarrow} c$, leaving no particular connection between the relevant c and i worlds.[25] And if $\Rightarrow$ is not a proximity conditional, the fact (given by $i \mathrel{\Diamond\!\!\rightarrow} c$) that the c worlds come as "close" to the actual world as the i worlds will simply fail to engage with the question of whether the relevant c worlds and relevant i worlds are the same.

In sum, attempts to support the claim that the relevant mediated masking examples are *possible* can be seen, in accord with the General Moral, as undermining the weakened transitivity rules required for satisfying the inferential condition. No general argument against conditional analyses appears forthcoming.

24. We return in section 5 below to the question of what a conditional rejecting proximity would look like.

25. That people find **Weak Trans B** intuitively less plausible than **Weak Trans A** (its validity for the Lewis counterfactual can come as a bit of a surprise) is perhaps a small piece of evidence favoring threshold conditionals over universalizing conditionals.

4.3 Causal Bases

Earlier (section 4.1) we listed five variations on the mediated masking argument. The third of these replaced the premise $i \Rightarrow \neg m$ with $(c \wedge i) \Rightarrow \neg m$. Lewis presents a form of the conditional fallacy that follows this third variation. His argument makes crucial use of the idea of *causal bases*:

> A finkish disposition is a disposition with a finkish base. The finkishly fragile glass has a property B that would join with striking to cause breaking; and yet the glass would not break if struck. Because if the glass were struck, straight away it would lose the property B. And it would lose B soon enough to abort the process of breaking. (Lewis 1997, 149)

Suppose, then, that b expresses the possession by the glass of some causal basis on which fragility depends, so that we accept $(c \wedge b) \Rightarrow m$ and $(c \wedge \neg b) \Rightarrow \neg m$. Following Lewis, suppose that the glass loses its causal basis when it is struck. Then we have a mediated masking argument, where Σ is the set of $(c \wedge \neg b) \Rightarrow \neg m$ and $c \Rightarrow \neg b$.[26] We can reason as shown in figure 8.

1. **COND**	A
2. $(c \wedge \neg b) \Rightarrow \neg m$	A
3. $c \Rightarrow \neg b$	A
4. $c \Rightarrow \neg m$	**Weak Trans C**, 2,3
5. $\neg(c \Rightarrow m)$	**Exclusion**, 4
6. $\neg d$	**Taut**, 1,5

Figure 8

What of the possibility condition? Why are we supposed to believe that a glass that loses the causal basis of fragility whenever a fragility-manifesting condition obtains still counts as fragile? Is it really so obvious that the World's Funniest Joke is indeed funny if the hearing of it kills the hearer, thereby removing the causal basis for manifestations of amusement?[27]

26. $(c \wedge b) \Rightarrow m$ is not necessary for the argument, although it will play a role in analogous mediated mimicking arguments.

27. One must prescind from the details of the joke killing via uncontrollable laughter, of course, as well as from an inclination to entertain the hypothesis of the joke heard by the unkillable auditor. Perhaps more to the point is the joke the hearing of which brings about a shift in personality that removes the basis for reactions of amusement (the racially insensitive joke the hearing of which leads one not to be amused by racially stereotyped anecdotes, for example). The case is analogous to

The claim requires some justification. Suppose, then, that a certain glass is fragile, by way of possession of a certain causal basis. Suppose the glass would lose the causal basis were it struck, and were it struck without the causal basis, it would not break. But suppose that the glass is one that could be struck only in highly anomalous conditions. If the conditional of analysis is a bounded conditional, this supposition will be of no help—the relevant worlds, no matter how anomalous, will still be $\neg b$ ones, and hence $\neg m$ ones, defeating the conditional analysis.

But if $\Rightarrow$ is not a bounded conditional in the sense defined in section 4.2 above, then it may be (since we thus drop the requirement that $f(w, [[\phi]]) \subseteq [[\phi]]$) that the worlds relevant to c-antecedent counterfactuals include some $\neg c$ worlds. If the mark of an anomalous ϕ situation is that it leads to *many* $\neg\phi$ worlds being relevant to ϕ-antecedent conditionals, then the anomalousness of striking may allow us to make sense of the fragility of the glass in terms of breakage at relevant $\neg c$ worlds. Such a situation would, of course, be very odd. The conditional analysis of the envisioned sort thus entails that a glass in a mediated masking case of this sort would not be fragile unless it were fragile in a very odd way—but this does indeed seem an apt summary of the situation.[28]

4.4 Competent and Perfect Finkishness

Suppose we replace the assumption $i \Rightarrow \neg m$ in the original mediated masking argument with the strict conditional $i \prec \neg m$. Then we can reason as presented in figure 9.

1. **COND**	A
2. $c \Rightarrow i$	A
3. $i \prec \neg m$	A
4. $c \Rightarrow \neg m$	?, 2,3
5. $\neg(c \Rightarrow m)$	**Exclusion**, 4
6. $\neg d$	**Taut**, 1,5

Figure 9

Kripke's Killer Yellow—interestingly, the inclination to say that Killer Yellow is indeed yellow seems much stronger than the inclination to say that this World's Funniest Joke is indeed funny.

28. Alexander Bird's argument from antidotes (Bird 1998, 228) presents another mediated masking case of the same logical sort. Assume **COND** and suppose that the glass is fragile. Nevertheless, if the glass were to be struck in the presence of an anti-

The rule required to replace '?' we henceforth call **Weakening**; it allows the replacement of the consequent of a conditional with any necessary consequence of that consequent. **Weakening** thus represents a deep feature of the conditional—any view of the conditional that treats the consequent as a test on modal space will support this rule since the strict conditional ensures that the passing of the one test entails the passing of the other test.

Call an argument of the above sort a *competent finking* argument. The interfering condition is fail-safe, but not necessarily attentive—there can be worlds (albeit not relevant to $c \Rightarrow i$[29]) in which the conditions of manifestation are met but the interfering conditions fail to manifest. We can also consider cases of *perfect finkishness* in which the interfering condition is both attentive and fail-safe. The following argument then results, as shown in figure 10.

1. **COND**	A
2. $c \prec i$	A
3. $i \prec \neg m$	A
4. $c \prec \neg m$	**Taut**, 2,3
5. $c \Rightarrow \neg m$	?, 4
6. $\neg(c \Rightarrow m)$	**Exclusion**, 5
7. $\neg d$	**Taut**, 1,6

Figure 10

The rule required to replace '?' we henceforth call **Strictness**; like **Weakening** it represents a deep feature of the conditional. Since the strict conditional $\phi \prec \psi$ guarantees that every ϕ world is a ψ world, any bounded conditional will support **Strictness**.

dote—canceling sound waves, for example—it would not break. Say that it is struck and the antidote is applied. Since the glass is struck, it breaks (because it is fragile), but it also does not break (since it is struck and the antidote is applied)—a reductio of **COND** results. Bird intends this argument against Lewis's complex analysis—showing that an appeal to a causal basis fails to prevent finking—but from our more abstract point of view, it provides another instance of a conditional-fallacy argument via **Weak Trans C** and shows that it is not essential that the mediator b be related to intrinsic microstructure. The relevant distinction is akin to that between altering and masking. A structurally isomorphic diagnosis will then be available to the conditional analyst able to give up **Weak Trans C**: antidotes are possible but, in conjunction with strikings, can be applied only anomalously.

29. Unless $\Rightarrow$ is a threshold rather than a universalizing conditional.

Given **Weakening** and **Strictness**, the competent and perfect finkishness arguments can meet the inferential condition. But what of the possibility condition? It is hardly implausible to think that the conditional-fallacy proponent has here reached inference rules so weak that the requisite finking scenarios are, in the end, simply implausible. When striking and failing to break (for example) are *so* tightly correlated, the object may just not be fragile. In a competent finking case, if the glass were struck, a divine agent would act and would *invariably* succeed in preventing the glass from breaking. All the worlds relevant to striking conditionals are interfering worlds and hence—via the strict conditional of fail-safety—nonbreaking worlds. Thus it is not the case that if the glass were struck, it would break. So, the glass is not fragile. The divine agent is overprotective; his infallible protection eliminates the very fragility it means to be protecting against.[30]

Nevertheless, if someone is indeed attracted to the plausibility of these finking scenarios, the conditional analyst *can* make room even for them. Surprisingly, the accommodation turns out to be easier in the case of perfect finkishness than in that of competent finkishness despite the apparently stronger conditions laid down in the perfect case. If, as in the case of the causal basis arguments above, some $\neg c$ worlds can be relevant to the evaluation of c-antecedent conditionals, then the conditional analyst can endorse the fragility of the perfectly finked object, and hence endorse the conditional $c \Rightarrow m$, despite the universal world-by-world correlation between c and $\neg m$, by resting on the relevant $\neg c$ worlds. Again, the background thought will be that the striking worlds are anomalous, and hence such that nonstriking worlds become relevant to questions of what *would* happen in striking cases. Accommodating competent finking cases, on the other hand, requires giving up the "test" conception of the conditional. Although non-consequent-test conditionals are available,[31] our own (considerable) generosity regarding the distribution of dispositional properties gives out at this point, and we are prepared to join the bullet-biting camp.[32]

30. If you doubt this, imagine trying to convince someone otherwise. If you were to hit the glass with a hammer, it wouldn't break. If you were to throw it onto the pavement, it wouldn't break. If you were to back your car over it, it wouldn't break. And you insist that it's fragile? It is intrinsically just like fragile glasses, perhaps, but thanks to the divine agent, it is not fragile.

31. Such as dynamic conditionals in which the consequent can reshape rather than merely test the contextually relevant worlds.

32. "Joining the bullet-biting camp" here amounts to endorsing the necessary inference principle for the analyzing conditional (**Weakening**), thereby accepting the

The major reason not to bullet-bite at this point seems to be a supervenience principle: two microphysically identical objects must have the same dispositional properties, and hence the finked and unfinked glass must be alike in fragility. Thus Lewis says that to share the microstructure of a fragile thing is to be fragile. This idea goes back at least as far as Aristotle, who observes that the very term 'disposition' indicates something dependent on "the arrangement of that which has parts" (*Metaphysics* 1022b1).

Competent and perfect finking cases show that one cannot maintain that fragility is both microstructural and conditional. Admittedly, normally things are fragile if and only if they have certain microstructural properties. But normally things are fragile if and only if they would break if struck. Under ordinary circumstances, microstructure and conditional behavior go together. A divine agent (or other interferences), however, can anomalously cause them to come apart.

We see at least three general stances one might adopt in response:

A. Fragility is really a microstructural property.[33] The glass protected by the divine agent is fragile, even if it could never break. The simple conditional analysis of *fragility* then fails. The failure of the conditional analysis of this particular property, however, does not entail that the conditional analysis of a different property must also fail.[34]

B. Fragility is really a simple conditional property. The supervenience of fragility on microstructure fails. The glass is not fragile despite its similarity to other, fragile glass. Ordinarily, fragility depends on microstructure. In this case, however,

logic of the competent finking argument and rejecting the Possibility Condition when the argument is advanced. But it should be noted that there is a bullet to be bitten only when there is, in fact, a suitable intermediary. Perhaps intermediaries are always available, but an argument for such a conclusion is wanting. (Could there be a competent-finking "evil demon"? Absent such a demon, competent finking arguments are at best once again a successful means of refuting particular conditional analyses, not the source of a general strategy for refutation.)

33. See Quine 1960, 223:

> What we have seen dissolve in water had, according to [a theory of subvisible structure], a structure suited to dissolving; and when now we speak of some new dry sugar lump as soluble, we may be considered merely to be saying that it, whether destined for water or not, is similarly structured. Fragility is parallel.

34. See again note 2.

the glass's (lack of) fragility depends on the divine agent's anomalous protection.

C. Fragility has no strict definition. Ordinarily, it satisfies both microstructural and conditional characterizations. In extraordinary cases, when these come apart, we are not sure what to say. Fragility is ambiguous or vague. Generally, we might count instances satisfying one characterization but not the other as borderline cases of fragility. If we are called upon to be more precise, we might count them as fragile in one sense but not in another. So, is the glass fragile? Sort of (the general answer); in one sense, but not in another (the more precise answer). On this view, the simple conditional analysis succeeds in delineating one sense of fragility, or, perhaps, one component of a multifaceted concept of fragility.

Opponents of conditional analyses must argue that only the first of these three responses is *ever* acceptable. The Canonical Test is again useful here. Is the thought that two objects identical in microstructure cannot differ in dispositional properties more compelling than the thought that two objects identical in microstructure cannot differ in the counterfactuals true of them? If the answer is "no," then a general conditional-fallacy argument cannot succeed here.

5. Conditional Archaeology

When presented with a conditional-fallacy argument, the conditional analyst who does not wish to bite the bullet on offer must see to it that the analysis employs a conditional that rejects at least one of the inferential principles used in satisfying the inferential condition of the fallacy argument. Seen in this light, conditional-fallacy arguments serve as challenges to various inferential features of conditionals. Although some attention has been given to the possibility of responding to the conditional fallacy by weakening conditional logic, the scope of the issue has not been fully appreciated.[35] Some masking arguments can be defeated by giving up **C**, and some mimicking arguments can be defeated by giv-

35. Recently Gundersen 2004 advocates a conditional logic that gives up weak centering, and hence **MP** and **C**, but that retains the various weak transitivity principles, as well as **Strictness** and **Weakening**. Morreau 1997 advocates a normality conditional with a quite weak logic; see section 5.1 below for further discussion. Fara 2005 argues against the use of **MP**-rejecting conditionals for conditional analysis; for his argument against normality conditionals see also section 5.1 below.

ing up **MP**. But consideration of the full range of possible conditional-fallacy arguments shows that each of **C**, **Expansion**, **Contraction**, **MP**, **Exclusion**, **Weak Trans A**, **B**, and **C**, **Strictness**, and **Weakening** is called into question. Although some of those calls may be rejected as the conditional analyst determines that the possibility condition has not been adequately met (competent and perfect finkishness scenarios may fail to persuade, for example, leaving **Strictness** and **Weakening** untouched), careful attention to the logic of the conditional will be needed to preserve **COND**.

Any need for revisions to standard accounts of the counterfactual conditional could reasonably be taken as an argument against conditional analyses of dispositions. Indeed, we view much of our discussion as serving to clarify the force of conditional-fallacy arguments when suitably generalized. That force is considerable—commitment to any of a number of prima facie plausible inferential features will allow one or another of a family of conditional-fallacy arguments to go through and block the use of conditionals with such a logic in the analysis of dispositions. If one takes the disjunction of these inferential features to be *essential* to conditionality (or implicit in natural language conditionals), then—if one finds the necessary finking situations generically available—one should conclude that dispositions cannot be given conditional analyses (or cannot be given such analyses with the resources of natural language).

Another response is also possible. We have adopted an abstract perspective on conditionality, guided by the Ramsey-Gentzen conception, with conditionals forming a broad class delimited by the satisfaction of $\Rightarrow$I and $\Rightarrow$E. Accordingly, the General Moral invites the following line of thought. We do indeed find many finking scenarios plausible, and the Canonical Test shows that we continue to find those scenarios plausible when they are recast in terms of the possession of explicitly conditionally defined properties rather than dispositional properties. We thus have reason to endorse a logic of conditionals that rejects sufficiently many of the inference rules used in the derivations satisfying the inferential condition. The conditional fallacy becomes a tool of conditional archaeology, allowing us to unearth unexpected logical features of conditionals.

In this last section, we tentatively examine some of the archaeological findings. We begin by discussing *normality conditionals*, as developed by Nicholas Asher and Michael Morreau (1995) and Morreau (1997), showing that they meet (most of) the inferential burdens conditional-fallacy arguments propose to reveal but then suggesting that they create their own difficulties for understanding dispositional properties. We conclude by pointing toward a region of logical space in which resides an alternative approach to the semantics of conditionals, an approach that matches the inferential successes of the normality conditional but avoids the accompanying difficulties.

5.1 Normality Conditionals

Normality conditionals express generic connections between states of affairs. Thus the normality conditional:

- If that were a bird, it would (normally) fly,

expresses a generic connection between *being a bird* and *flying*. The notion of a generic connection is intended to tolerate a range of exceptions. If that were a penguin, it would be a bird but would not fly. This exception fails to falsify the normality conditional because being a penguin is an "abnormal" way of being a bird. If we think of (for example) a finked fragile glass as one that has been placed in an abnormal situation, a conditional designed to tolerate abnormal exceptions may seem exactly what is needed to salvage the conditional analysis.

The normality conditional $\phi > \psi$ can be given a semantic interpretation that follows the guiding idea of the Lewis counterfactual, but with the understanding that the conditional checks for ψ at the *normal* ϕ worlds rather than at the *nearest* ϕ worlds. The shift from proximity to normality leads to two significant alterations in the Lewis semantics:

1. Since there is no assumption that the actual world is normal in any respect, weak centering is abandoned. Just because the actual world is a ϕ world, it does not follow that it is relevant to the evaluation of $\phi > \psi$.
2. The idea of a *perfectly normal* world is of dubious coherence. Normal birds fly, so a perfectly normal world ought to contain no penguins among its birds. On the other hand, normal penguins swim rather than fly, so a perfectly normal world must contain swimming, rather than flying, penguins, and surely those penguins are birds. Normality conditionals thus appeal to the notion of *normality for a condition*. $\phi > \psi$ is evaluated at ϕ-normal worlds.

Normality conditionals employ a selection function that can be thought of as imposing a vast number of (very coarse) sphere systems about each world, as opposed to the single Lewis sphere system, or as imposing a vast number of accessibility relations.

Because the normality conditional does not impose weak centering (which would require the condition $w \in f(U, w)$ for all $U \subseteq W$), the rules **C**, **Expansion**, **MP**, and **Contraction** all fail for >. For example, if Tweety is a bird, then normally Tweety flies. But it does not follow that Tweety actually does fly because Tweety may be injured or a penguin. Normality conditionals thus have the resources necessary to resist masking and mimicking arguments.

Because normality is antecedent-relative, **Weak Trans A**, **B**, and **C** all fail as well. Antecedent-relative normality amounts to an abandonment of proximity for the normality conditional.[36] The method of selecting the worlds relevant to the evaluation of $\phi > \psi$ does not depend simply on the global structuring of the worlds together with the truth value of ϕ at worlds, but instead depends on a structuring of modal space specific to the antecedent ϕ. Thus (for example) imposing both $\phi > \psi$ and $\psi > \phi$ (as in **Weak Trans A**) does not ensure that the conditional-relevant ϕ worlds be the same as the conditional-relevant ψ worlds since $\phi > \psi$ may, in looking at ϕ-normal worlds, run on an entirely different region of modal space than $\psi > \phi$, which examines ψ-normal worlds.

The failure of the weak transitivity principles means that the normality conditional is also immune to the various mediated masking arguments. However, **Weakening** and **Strictness** both hold for normality conditionals, so anyone moved by the plausibility of competent and perfect finkishness scenarios should be unhappy with the normality conditional as a tool for analyzing dispositions.

Morreau 1997's term "fainthearted conditionals" picks up on a use of that term by Nelson Goodman (1954) in his discussion of conditionals and dispositions. Goodman's criticism of such fainthearted conditionals and their need for an account of "propitious conditions" then finds an echo both in Martin 1994 and in Fara 2005's objection to Morreau.[37] Michael Fara (2005, 58), noting that normality conditionals appeal to the notion of antecedent-normal worlds, argues that

> a circumstance is normal with regard to some disposition only if it is not a circumstance in which that disposition is being masked—that much is trivial. But then, without any more being said about what it is for a cir-

36. The normality conditional is, however, bounded—normal-ϕ worlds must always be ϕ worlds.

37. For responses to the criticisms of Goodman and Martin, see Mumford 1998.

> cumstance to be normal, any view according to which disposition ascriptions entail their corresponding conditionals restricted to normal circumstances is likewise trivial, and will be of no help when it comes to giving an informative statement of the truth conditions of disposition ascriptions.

The objection is well understood as going beyond the paradox of analysis. The question is whether a defender of normality conditionals has a grasp of normality that is independent of, or prior to, his or her grasp of dispositions. Otherwise, Fara's triviality complaint will stick. If "normal" simply *means* "such that the disposition is not being masked (or, more generally, finked)," the (normality-)conditional analyst understands "X has disposition D" to mean "If X were placed in such-and-such conditions of manifestation in which its disposition D were not being masked/finked, then it would display such-and-such a manifestation." The now trivial proposed analysis of the disposition would itself contain an ascription of the very same disposition.[38]

Thus far an explanatory burden has been located; it has not been shown that the normality conditionalist *cannot* or *does not* have a suitable conception of normality. However, the very semantic moves that have allowed the normality conditional to shed the troublesome inference rules also make more difficult the task of providing the requisite notion of normality. This is so for two reasons:

38. Fara's own account of dispositional ascriptions is in terms of habitual sentences, his semantics for which appeals to the notion of *permissible exceptions*. (In Fara 2001, 76, the notion of a permissible exception is in turn defined in terms of the notion of a normal condition; in Fara 2005, the notion of a permissible exception is left undefined.) Fara does thus incur the same sort of explanatory burden as the normality conditionalist. He (Fara 2005, 30) says:

> I am offering an account of disposition ascriptions stated in terms of ordinary, non-dispositional sentences of English: habitual sentences. I am not offering an account stated in terms of expressions that may appear in a semantics for these ordinary sentences.

The normality conditionalist, too, will want to claim simply to be offering an account of dispositional ascriptions stated in terms of ordinary, nondispositional sentences of English: conditionals. That others, unused to normality conditionals and overly influenced by the legacy of the Lewis semantics, insist on inserting explicit mention of normality in these conditionals should not (they might add) hide the fact that normality is just a notion used in the elucidation of the semantics for normality conditionals (in the same way that similarity is used in the elucidation of the semantics for Lewis counterfactuals).

1. In abandoning weak centering, the normality conditional abandons the use of the categorical features of the actual world as a guide to normality. It is consistent with the semantics of the normality conditionals that the ϕ-normal worlds, for any (or all) ϕ be *maximally dissimilar* to the actual world—things as they actually are can be unlimitedly abnormal. This possibility significantly weakens our intuitive grasp on the notion of normality. A Lewis conditional with weak centering allows us to use some notion of similarity to the actual world as a guide to the proximity structuring of modal space; the uncentered normality conditional permits no such guidance.
2. In defeating the weak transitivity rules by the move to antecedent-relativized normality, the normality conditional greatly increases the explanatory burden. Without an overarching global measure function on modal space (as the Lewisian system of spheres provides), we require multiple independent notions of normality-for-various-ϕ's.[39]

5.2 Neighborhood Conditionals

The two difficulties just cited are perhaps not insuperable. But if the conditional used in analyses of dispositions could preserve weak centering, we could appeal to the relatively well-understood notion of world-to-world similarity for its structuring of modal space. If, however, weak centering is preserved, how can simple mimicking arguments (with their appeal to weak-centering-driven **MP** and **Contraction**) and mediated masking arguments be answered?

Suppose $\Rightarrow$ is a conditional that is not bounded. How could this be? How could non-ϕ worlds be relevant to the evaluation of ϕ-antecedent conditionals? Lack of boundedness can be made intelligible by abandoning an even more fundamental aspect of the Lewis counterfactual—the assumption that it is evaluated *pointwise.* Instead we could suppose that $\Rightarrow$ is a *neighborhood conditional,* whose purpose is to test the coordination of features across *regions* of modal space. Thus $\phi \Rightarrow \psi$ would mean, roughly, that some relevant ϕ-region of modal space was in a ψ region.

39. This may understate the problem; a solution to the "drowning problem" may require not only an independent notion of normality for each $[[\phi]]$ but also an independence measure for each $[[\phi]]$—$[[\psi]]$ pair. See Asher 1995.

Instead of understanding "If that were a bird, it would fly" as asserting that all normal *that's a bird* worlds are *that flies* worlds, we might think of it as asserting that *that flies* is sufficiently thick about the neighborhood of *that's a bird* worlds. A conditional with such truth conditions would allow for the possibility of exceptions in a way that differs significantly from that of existing normality conditionals.

To be a neighborhood conditional along these lines would be to have the following consequences:

1. For a region of modal space to be a ϕ region needn't be a matter of every world in that region being ϕ; it can instead depend on the general character of the region being ϕ-like. It could thus be particularly natural for $\Rightarrow$ to be a threshold conditional. This allows **MP**, **Contraction**, and the weak transitivity principles to fail.
2. That the relevant ϕ region is imbedded in a ψ region and the relevant ψ region is imbedded in a ϕ region needn't imply that the relevant ϕ and ψ regions are identical. Thus $\Rightarrow$ need not be a proximity conditional, again allowing the weak transitivity principles to fail.
3. The relevant ϕ-region's being in a ψ region may depend on a suitable ψ presence in the region *about the ϕ region.* ϕ itself may not hold in that vicinity about the ϕ region, so the imposition of the ψ test need not be restricted to ϕ worlds. Thus $\Rightarrow$ is not a bounded conditional, and **Strictness** fails.

The earlier consideration of nonproximity and unboundedness contemplated abandoning popular *constraints* on the selection function f; the current possibility of nonpointwise conditionals contemplates incorporating the selection function into a more general apparatus that undermines the prior motivation for those constraints. The truth condition for a neighborhood conditional would thus generalize the (already quite abstract) semantics for conditionals in terms of selection functions. Rewriting

- $w \vdash \phi \Rightarrow \psi$ if and only if $f(w, [[\phi]]) \subseteq [[\psi]]$

as

- $w \vdash \phi \Rightarrow \psi$ if and only if $f(w, [[\phi]]) \in \wp([[\psi]])$,

we then generalize, thinking of a set of worlds as a region, and let g be any function from regions to sets of regions:

- $w \vdash \phi \Rightarrow \psi$ if and only if $f(w, [[\phi]]) \in g([[\psi]])$.

Intuitively, $g([[\psi]])$ is the set of regions about which ψ is thick. The resulting neighborhood conditional would then be able to reject the inferential conditions required by a vast array of conditional-fallacy arguments including competent finking arguments: if g is not monotonic, the rule **Weakening**, which we have largely implicitly accepted throughout, fails.[40]

Obviously a great deal of work must be done before these schematic remarks can be converted into anything like a plausible account of a conditional. Perhaps, in the end, this schematic neighborhood conception fails to lead to a plausible conditional. But if we view the General Moral as introducing a project of conditional archaeology, systematic application of the Canonical Test leads us, through consideration of various conditional-fallacy arguments, to a collection of (perhaps surprising) constraints on the logic of this conditional and bequeaths to us a philosophical project of making sense of it. Normality and neighborhood conditionals would be two attempts to make good on the project, but the project itself does not stand or fall with these two attempts.

We think reports of the death of conditional analysis are greatly exaggerated. As emphasized throughout, we do not defend any specific conditional analysis. But the thought that such analyses, in general, are bound to commit the conditional fallacy is, we think, premature. To reveal this, we first characterize the conditional fallacy generally, showing how it might be thought a fallacy and how that thought might be differentiated from a general, possibly question-begging, claim that for any conditional analysis offered there will be a counterexample. In particular, our account characterizes the fallacy as involving a set of claims satisfying an inferential condition and a possibility condition. Our discussion then focuses on the inferential condition. Conditional-fallacy arguments rely on an optional asymmetry between the inferential behavior of the analysans and the analysandum. In particular, conditional-fallacy arguments load down the conditional used in the analysis with inference rules of controversial generality. The role of centering, for example, in masking and mimicking arguments has not been adequately appreciated. And even in subtler, mediated cases, in which the inference rules are less

40. Perhaps, though, some of the lost inferential features can be regained via a nonmonotonic consequence relation that serves in effect to "filter out" anomalous counterexamples such as, perhaps, finking scenarios. Note that normality conditionals, which do not allow **MP** monotonically, do allow defeasible application of that rule.

bold, use of our Canonical Test can be seen as transforming the dialectic from rejection to development: far from serving as a general recipe for the refutation of conditional analyses, "finking" cases can serve to expose presuppositions of some of our conditional modal thought. The relevant examples can be used as a tool for conditional archaeology: intuitive responses to the finking examples support investigation of the possibility of a new sort of conditional, a "neighborhood" conditional that does not satisfy some of the traditional presuppositions about conditionals. Unmasked, finking scenarios can open rather than close philosophical doors. The conditional fallacy is only conditionally fallacious.

References

Aristotle. 1936. *The Metaphysics*, Books 1–9, trans. Hugh Tredennick. Loeb Classical Library. Cambridge, MA: Harvard University Press.

Asher, Nicholas. 1995. "Commonsense Entailment: A Conditional Logic for Some Generics." In *Conditionals: From Philosophy to Computer Science*, ed. G. Crocco, Luis Fariñas del Cerro, and Andreas Herzog, 103–46. Oxford: Clarendon.

Asher, Nicholas, and Michael Morreau. 1995. "What Some Generic Sentences Mean." In *The Generic Book*, ed. Greg N. Carlson and Francis Jeffry Pelletier, 300–338. Chicago: University of Chicago Press.

Bird, Alexander. 1998. "Dispositions and Antidotes." *Philosophical Quarterly* 48: 227–34.

Bonevac, Daniel. 2003. *Deduction*. Malden, MA: Blackwell.

Carnap, Rudolf. 1936. "Testability and Meaning I." *Philosophy of Science* 3: 419–71.

Chellas, Brian. 1980. *Modal Logic: An Introduction*. New York: Cambridge University Press.

Fara, Michael. 2001. "Dispositions and Their Ascriptions." PhD diss., Princeton University.

———. 2005. "Dispositions and Habituals." *Noûs* 39: 43–82.

Gentzen, Gerhard. 1969. "Investigations into Logical Deduction." In *Collected Papers of Gerhard Gentzen*, ed. M. E. Szabo, 68–131. Amsterdam: North-Holland.

Goodman, Nelson. 1954. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Gundersen, Lars B. 2004. "Outline of a New Semantics for Counterfactuals." *Pacific Philosophical Quarterly* 85: 1–20.

Johnston, Mark. 1992. "How to Speak of the Colors." *Philosophical Studies* 68: 221–63.

———. 1993. "Objectivity Refigured: Pragmatism and Modern Idealism." In *Reality, Representation, and Projection*, ed. Crispin Wright and John Haldane, 85–130. New York: Oxford University Press.

Lewis, David. 1973. *Counterfactuals*. Cambridge, MA: Harvard University Press.

———. 1989. "Dispositional Theories of Value." *Proceedings of the Aristotelian Society* 63: 113–37.

———. 1997. "Finkish Dispositions." *Philosophical Quarterly* 47: 143–58.

Martin, C. B. 1994. "Dispositions and Conditionals." *Philosophical Quarterly* 44: 1–8.

Morreau, Michael. 1997. "Fainthearted Conditionals." *Journal of Philosophy* 94: 187–211.

Mumford, Stephen. 1998. *Dispositions*. Oxford: Clarendon.

Quine, W. V. O. 1960. *Word and Object*. Cambridge, MA: MIT Press.

Ramsey, F. P. 1931. "General Propositions and Causality." In *Foundations of Mathematics and Other Logical Essays*, ed. R. B. Braithwaite, 237–57. London: Routledge and Kegan Paul.

Ryle, Gilbert. 1949. *The Concept of Mind*. New York: Barnes and Noble.

Shope, Robert. 1978. "The Conditional Fallacy in Contemporary Philosophy." *Journal of Philosophy* 75: 397–413.

Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.