In What Sense (If Any) Is Logic Normative for Thought?

John MacFarlane

Draft of April 21, 2004
For presentation at the Central Division APA 2004

Abstract

Logic is often said to provide norms for thought or reasoning. Indeed, this idea is central to the way in which logic has traditionally been defined as a discipline, and without it, it is not clear how we would distinguish logic from the disciplines that crowd it on all sides: psychology, metaphysics, mathematics, and semantics. But it turns out to be surprisingly hard to say how facts about the validity of inferences relate to norms for reasoning, and some philosophers have concluded that the whole idea is confused. In this talk I will survey a space of possible "bridge principles" connecting logical facts with norms for reasoning. After discussing some considerations relevant to choosing between these bridge principles, I will defend two of them. I will then consider the implications of various choices of bridge principle for the long-standing debates about the roles of relevance, necessity, and formality in our notion of logical consequence. The methodological aim of the talk is to provide an alternative to the usual brute appeals to our "intuitions" about logical consequence in these fundamental debates.

1 The issue and its importance

Logic is often said to provide norms for thought or reasoning. Indeed, this idea is central to the way in which logic has been demarcated as a discipline, and without it, it is hard to see how we would distinguish logic from the disciplines that crowd it on all sides: psychology, metaphysics, mathematics, and semantics. (Try saying how logic differs from geometry without mentioning thought or reasoning, and try saying how logic differs from psychology without mentioning norms.) But in what sense, exactly, is logic normative for thought?

Before tackling this question, I want to say a bit about why I think it is important. One of the most central topics in philosophy of logic—and one of the most confusing—is the concept of validity. A logic is supposed to tell us which inferences are valid. But what exactly does it *mean* to say that the inference from A to B is "valid," or that B "follows logically" from A? There is an intuitive modal characterization of validity: necessarily,

if the premises are true, the conclusion is true as well. But it is not universally accepted. Some think an additional relevance condition is needed. Others think that only those inferences that are necessarily truth-preserving "in virtue of their forms" are valid—and there is further debate about what "in virtue of their forms" means. Still others reject the modal characterization entirely. Nor is the modal characterization particularly clear. When we say that the inference from A to B is valid, what are A and B: propositions, utterances, statements, or sentences? Is the modality in the modal characterization alethic or epistemic? How strong is it? Can it be explained without using the notion of logical validity (or a close relative)? How should the conditional in the characterization be understood?

There are, of course, many clear *technical* explications of validity, both proof-theoretic and semantic, but it is not at all clear how these relate to the intuitive notion. For example, what is the point of quantifying over models in the standard model-theoretic definition of validity? Are we trying to capture the modal element in the intuitive modal characterization, so that the models represent possible worlds or situations? Or are we trying to capture the idea that logical validity depends only on form, and not on the meanings of non-logical terms? (As Etchemendy 1990 has shown, there seem to be problems with both answers.) And why does the model-theoretic definition treat some terms (the "logical constants") differently from others? What is the principled basis for this distinction?

These are hard questions. The dominant methodology for addressing them involves frequent appeals to our "intuitions" about logical validity. I do not think it should surprise us that this methodology leads different investigators in different directions. For our intuitions about logical validity, such as they are, are largely the products of our logical *educations*. Anyone who has taught elementary logic will know that it can take a week or more to get students to distinguish questions of validity from questions of soundness. Even after they've caught on, their intuitions about validity are likely to depend heavily on the glosses they have been offered. Consider these two dialogues, both eminently realistic.

> *Teacher:* So does "$2 + 2 = 4$" follow logically from "Today is Tuesday"?
> *Student:* Huh?
> (a)
> *Teacher:* Well, could the conclusion be false if the premise was true?
> *Student:* I guess not, because the conclusion couldn't be false period. I guess it does follow.
> (b)
> *Teacher:* Well, could the conclusion be validly inferred from the premise?
> *Student:* No, that would be crazy. I guess it doesn't follow.

Professional philosophers and logicians tend to have more settled intuitions. But that is because they have had more time to become indoctrinated. If they were all indoctrinated in the same way, there might be some point to the intuition-mongering methodology. But as things stand, there is very little convergence, and lots of deadlock.

I doubt that we will make much progress on any of these questions until we get beyond intuitions and get a clearer understanding of what the concept of logical validity is *for*. Why

do we bother studying this notion at all? Surely it is because we think there is some connection between logical validity and the evaluation and criticism of reasoning. If we could get clearer about this connection, we could transpose questions about logical validity into questions about how we ought to think. And that is something about which our views are considerably less dependent on education and indoctrination. It's not that we're completely immune to false subtleties in this area. But thinking is something we do all the time, and something we have a great interest in doing properly. Our views about correct thinking are the product of our own experience and reflection, not just education. And I would wager that there is considerably more convergence in these views than there is in our intuitions about logical validity. The validity of *ex falso quodlibet* ($P, \neg P / R$) is hotly debated, but *no one* thinks that when one finds oneself with contradictory beliefs one ought to conclude that one is a pumpkin.

For the sake of concreteness, let's look at the issue of relevance in a bit more detail. Relevantists hold that *ex falso quodlibet* and disjunctive syllogism ($P \vee Q, \neg P / Q$) are invalid, because their premises are not properly relevant to their conclusions.[1] Classicists hold that these inference forms are valid, on the grounds that they cannot take one from true premises to a false conclusion. Clearly the dispute is, at root, a dispute about what logical validity *is*. For relevantists will concede that *ex falso quodlibet* is necessarily truth-preserving, so that it *would* be valid if the classicists were right that logical conesquence is necessary truth-preservation. It's just that that's not what validity is.

How do relevantists support their view about what logical validity is (or is not)? In the main, they appeal to our putatively shared intuitions about "what follows from what."

> Surely it does not *follow* from your having eaten eggs this morning that all logicians are logicians! That conclusion may be true, and necessarily so, but there's no connection! Anyway, why would we call such cases "paradoxes" of strict implication if there weren't something funny about them?

I have already explained why this kind of appeal to intuitions cannot carry much argumentative weight. Indeed, relevantists have plenty of reason to be wary of it. Their own claim that disjunctive syllogism is invalid—a claim they must make in order to coherently reject *ex falso quodlibet*—strikes most people as extremely unintuitive. And they themselves charge that the classicists' intuitions have been corrupted by the indoctrinating effects of education.

I suspect that relevantism would not have the following it does if its motivation rested solely on intuitions about validity. There is a crypto-motivation for relevantism that appeals instead to our intuitions about good reasoning. We can see it, for instance, in this quote from Graham Priest, criticizing the classical account of validity:

> For the notion of validity that comes out of the orthodox account is a strangely perverse one according to which any rule whose conclusion is a logical truth is

[1] I follow Anderson and Belnap in using the term "relevantist" for someone who takes this controversial position. A "relevance logician" studies systems that make this possible, but may or may not be a relevantist.

> valid and, conversely, any rule whose premises contain a contradiction is valid. By a process that does not fall far short of indoctrination most logicians have now had their sensibilities dulled to these glaring anomalies. However, this is possible only because logicians have also forgotten that logic is a normative subject: it is supposed to provide an account of correct reasoning. When seen in this light the full force of these absurdities can be appreciated. Anyone who actually reasoned from an arbitrary premise to, e.g., the infinity of prime numbers, would not last long in an undergraduate mathematics course. (Priest 1979: 297)

Priest assumes here that if an argument is valid, it is always correct to reason from its premises to its conclusion. If this assumption is sound, it puts the case for relevantism on a much firmer footing, appealing to robust intuitions about correct reasoning rather than weak (and less widely shared) intuitions about validity.[2] Some may think that Priest's assumption is innocuous, but others have rejected it, and even suggested that without it, the whole case for relevantism dries up. I think the truth is somewhere in between. One thing should be clear, however: in order to make further progress in this debate, we need to get much clearer about the relation between validity and norms for thought and reasoning.

2 Normative for what?

Let's start with the slogan "logic is normative for reasoning." "Reasoning" can be used in either a formal or an informal sense. Informally, reasoning is just "reasoned change in view" (as Harman 1986 puts it).[3] To reason is to figure out what to believe. One reasons well if one revises one's beliefs as one ought to (where this revision can involve both additions to and subtractions from one's set of beliefs). In a more formal sense, reasoning is a process of drawing out the consequences of a given set of premises. One need not believe the premises: one might just be investigating them, or using them in a conditional proof or *reductio ad absurdum*. To distinguish this process from reasoning in the sense of "reasoned change in view," we might call it "inferring" (though "inferring" may be subject to the same kind of ambiguity as "reasoning"). In which sense of "reasoning" does logic provide norms for reasoning? Is logic normative for belief revision, or for inferring?

I think it is relatively uncontroversial that logic provides norms for inferring (in the narrow sense of drawing out consequences). For the proof rules of a logic are *explicitly*

[2] Priest now repudiates this argument, but not because he rejects the way it links validity to norms for reasoning. Rather, he rejects its assumption that a model-theoretic account of validity must inevitably validate *ex falso quodlibet*. As he notes in Priest 1999, this depends on whether there are models representing nontrivial but inconsistent situations (202).

[3] Cf. Harman 1984: "There is a tendency to identify reasoning with proof or argument in accordance with rules of logic. Given that identification, logic obviously has a special role to play in reasoning. But the identification is mistaken. Reasoning is not argument or proof. It is a procedure for revising one's beliefs, for changing one's view. ... The question, then, is whether logic has a special role to play in this procedure of belief revision." (107)

normative: for example, the $\supset$-elimination rule says that if you have already written down A and $A \supset B$, you may write down B. These proof rules *license* or *permit* certain inferences.

So here is a clear sense in which logic is normative for reasoning. But this sense isn't going to help us much with the problems we looked at in the last section. Our intuitions about when it is permissible to infer a conclusion from some premises (in the narrow sense) have the same sources as our intuitions about logical validity: primarily, our logical training. (Indeed, it takes some logical training in order to engage in the practice of "inferring" at all: one must be trained not to use information not contained in the premises, for instance, and not to worry about whether the premises are true.) Thus these intuitions are likely to be subject to just the same "indoctrination biases" as our intuitions about validity. A classicist will take it to be correct to infer anything from a contradiction in formal argumentation, while a relevantist will not. If we are to get beyond this kind of conflict of intuitions, we need to talk about norms for reasoning in the broader sense: norms for belief and belief change.

Here is another way to see the same point. Formal argumentation—the controlled drawing of consequences from a set of premises—is a *tool*. We engage in it (and train our students to engage in it) not for its own sake, but because we think it is useful for telling us what we ought to believe. We infer *correctly* when we infer in a way that is conducive to this goal. In what way, then, does formal argumentation help us in revising our beliefs? Presumably it does so be telling us what follows logically from what. But how does knowing *this* help us in revising our beliefs? Now we are back to the question of how logical validity relates to norms for belief (or reasoning in the broad sense). That is what we must understand.[4]

But this is much harder to understand. Unlike proof rules, claims about logical validity are not explicitly normative in their content. If they give rise to norms, it is by way of some bridge principle linking claims about logical validity with norms for belief. But as Gilbert Harman has pointed out, it is not at all obvious what this bridge principle should be. If we are unclear about the distinction between reasoning in the narrow sense (inferring) and reasoning in the broad sense, we might suppose that the fact that $A, A \supset B \models B$ *licenses* anyone who believes A and $A \supset B$ to believe B. But it does not. If B is absurd or clearly false (in light of one's other beliefs), one should instead abandon one's belief in A or $A \supset B$. It is even less plausible that one is *obligated* to believe all of the logical consequences of one's beliefs. In addition to the consideration above, there is a worry about "clutter avoidance:" am I really obligated to believe all of the infinitely many trivial logical consequences of my beliefs?[5] There is also a worry about "excessive demandingness." Are the norms of logic so demanding that no human being could possibly satisfy them? Yet no human being could believe *all* of the logical consequences of the Peano axioms for arithmetic. One might retreat to the position that logic only makes *negative* demands on

[4]Cf. Harman 1984 "So I thought that I might be able to understand how argument and calculation facilitate reasoning if I could understand how the appreciation of implications can facilitate reasoning." (112–13).

[5]For both points, see Harman 1984:113, Harman 1986.

belief: one ought not believe things that are inconsistent, for example. But even this can be questioned:

> . . . even the rule 'Avoid inconsistency!' has exceptions, if it requires one not to believe things one knows to be jointly inconsistent. On discovering one has inconsistent beliefs, one might not see any easy way to modify one's beliefs so as to avoid the inconsistency, and one may not have the time or ability to figure out the best response. In that case, one should (at least sometimes) simply acquiesce in the contradiction while trying to keep it fairly isolated. I would think this is the proper attitude for most ordinary people to take toward many paradoxical arguments.
>
> Furthermore, a rational fallible person ought to believe that at least one of his or her beliefs is false. But then not all of his or her beliefs can be true, since, if all of the other beliefs are true, this last one will be false. So in this sense a rational person's beliefs are inconsistent. (Harman 1984: 108–9)

In this way, even the contraction of logical norms to an injunction against inconsistent beliefs seems too strong. In another way, though, it seems too weak. It seems false that logic makes only negative demands on belief. We criticize people not merely for having inconsistent beliefs, but for failing to accept logical consequences of their beliefs.

The issues are tangled here. Let's try to sort them out more systematically.

3 The bridge principle

We need a bridge principle of the following form:

> BRIDGE PRINCIPLE: If $A, B \models C$, then (normative claim about believing A, B, and C).

The question is what the consequent should look like. We can generate a nice set of options by varying three parameters:

1. *Type of deontic operator.* Do facts about logical validity give rise to strict *obligations*, *permissions*, or (defeasible) *reasons* for belief?

2. *Polarity.* Are these obligations/permissions/reasons *to believe*, or merely *not to disbelieve*?

3. *Scope of deontic operator.* These norms are in some sense conditional: what one ought/may/has reason to believe with respect to C depends somehow on what one believes, or ought/may/has reason to believe, with respect to A and B. Does the deontic operator govern the *consequent* of the conditional ($P \supset O : Q$), or both the antecedent and the consequent ($O : P \supset O : Q$), or the whole conditional ($O : (P \supset Q)$)?

Table 1: If $A, B \models C$, then ...

`C` Deontic operator embedded in consequent.

- `o` Deontic operator is strict obligation (ought).
  - `Co+` if you believe A and you believe B, you ought to believe C.
  - `Co-` if you believe A and you believe B, you ought not disbelieve C.
- `p` Deontic operator is permission (may).
  - `Cp+` if you believe A and you believe B, you may believe C.
  - `Cp-` if you believe A and you believe B, you are permitted not to disbelieve C.
- `r` Deontic operator is "has (defeasible) reason for."
  - `Cr+` if you believe A and you believe B, you have reason to believe C.
  - `Cr-` if you believe A and you believe B, you have reason not to disbelieve C.

`B` Deontic operator embedded in both antecedent and consequent.

- `o` Deontic operator is strict obligation (ought).
  - `Bo+` if you ought to believe A and believe B, you ought to believe C.
  - `Bo-` if you ought to believe A and believe B, you ought not disbelieve C.
- `p` Deontic operator is permission (may).
  - `Bp+` if you may believe A and believe B, you may believe C.
  - `Bp-` if you may believe A and believe B, you are permitted not to disbelieve C.
- `r` Deontic operator is "has (defeasible) reason for."
  - `Br+` if you have reason to believe A and believe B, you have reason to believe C.
  - `Br-` if you have reason to believe A and believe B, you have reason not to disbelieve C.

`W` Deontic operator scopes over whole whole conditional.

- `o` Deontic operator is strict obligation (ought).
  - `Wo+` you ought to see to it that if you believe A and you believe B, you believe C.
  - `Wo-` you ought to see to it that if you believe A and you believe B, you do not disbelieve C.
- `p` Deontic operator is permission (may).
  - `Wp+` you may see to it that if you believe A and you believe B, you believe C.
  - `Wp-` you may see to it that if you believe A and you believe B, you do not disbelieve C.
- `r` Deontic operator is "has (defeasible) reason for."
  - `Wr+` you have reason to see to it that if you believe A and you believe B, you believe C.
  - `Wr-` you have reason to see to it that if you believe A and you believe B, you do not disbelieve C.

`-k` (As suffix to one of the above:) antecedent of bridge principle is "If *you know that* $A, B \models C$ ...."

There are eighteen variations in all, represented and named in Table 1.

A few comments on the table:

1. In the naming scheme, the first letter indicates the scope of the deontic operator, the second letter indicates the type of the deontic operator, and the third letter indicates the polarity. So, for example, `Wo-` is

   > If $A, B \models C$, then you ought to see to it that if you believe A and you believe B, you do not disbelieve C,

   where the deontic operator is obligation (`o`), taking wide scope over the whole conditional (`W`), and the polarity is negative (`-`). Some have thought that only *known* facts about logical validity give rise to norms for thought.[6] We can add these positions by introducing the suffix `k`, which changes the antecedent of the bridge principle to "If *you know that* $A, B \models C$." Thus, `Br+k` is

   > If you know that $A, B \models C$, then if you have reason to believe A and have reason to believe B, you have reason to believe C.

   That gives us 36 versions of the bridge principle in all. (Note: the conditionals in all these versions should be understood, for now at least, as material conditionals.)

2. The force of "you ought not Φ" is not "it is not the case that you ought to Φ," but rather "you are forbidden to Φ," which following Belnap, Perloff, and Xu 2001 I understand as "it is obligatory that you refrain from Φing," that is, "it is obligatory that you see to it that you do not Φ."

3. "Disbelieving" is a mental state that stands in the same relation to believing as denying does to asserting. It is not the same as "not believing": there are many things one neither believes nor disbelieves. The negative polarity versions of the bridge principle say "not disbelieve C" instead of "not believe $\neg C$" for two reasons. First, it seems undesirable to bring particular logical constants into a general principle about how logics relate to thought. It makes perfect sense to talk of the logic of a negation-free fragment of a language, and there ought to be a way of connecting such a logic to norms for thought without introducing new vocabulary. Second, and more important, the Fregean identification of "disbelieving p" with "believing not-p" is controversial. Dialetheists reject it, because they think one should sometimes believe both p *and* not-p (when p is both true and false), though one should never both believe and disbelieve p.[7] A dialetheist could reasonably propound

   `Wo-` If $A, B \models C$, then you ought to see to it that if you believe A and believe B, you do not disbelieve C,

[6] For example, Sainsbury 2002: 3.

[7] See Priest 1998: 425.

but not

`Wo`$\neg$ If $A, B \models C$, then you ought to see to it that if you believe A and believe B, you do not believe $\neg C$.

Since the paraconsistent logician accepts $A \models A$, `Wo`$\neg$ would yield a global prohibition against believing anything and its negation.[8]

4. The difference between obligations and reasons is that obligations are strict, whereas reasons are defeasible and *pro tanto*.[9] If one ought to do A and does not do A, one is thereby subject to criticism; not so if one has reason to do A and does not do A. If one has reason to do A and no reason not to do A, then presumably one ought to do A. But often one has reason both to do A and not to do A. In that case, the stronger reason determines what one ought to do. One ought to go to class even though one has reason to take a walk instead, for one has an even more compelling reason to go to class. Conflicting reasons of this kind are the norm in practical reasoning; conflicting obligations, if they are possible at all, are much rarer. In a genuine case of conflicting obligations, one cannot emerge blameless. No matter what one does, one is remiss for failing to do what one ought to. Let us leave it open whether it is possible for obligations to conflict in this way.

What considerations tell for and against the various possible ways of filling in the bridge principle? We can eliminate the `C`'s straightaway. We have already noted Harman's point that if one finds that one's beliefs logically imply an absurdity, one should should generally revise one's premises rather than accepting the absurd conclusion. Broome 2000 gives an elegant argument against `Co+`. Any reasonable logic will contain $A \models A$ as a theorem. Thus `Co+` implies that for every A, if you believe A, you ought to believe A. But as Broome points out, belief is not "self-justifying" in this way (85). Sometimes we believe things that we ought *not* believe. Similar considerations tell against `Co-`, `Cp+`, and `Cp-` (and against their `-k` variants).

`Cr+` and `Cr-` are not as obviously wrongheaded. `Cr+` says that we have a defeasible reason to believe each of the logical consequences of our standing beliefs. This is not at all absurd; indeed, some philosophers hold just this view. The trouble is, they accept it because it is a consequence of two other things they believe:

(1) If you believe A, you have (defeasible) reason, or "default entitlement," to believe A. (The fact that you have a belief is a *pro tanto* reason to retain it.[10])

[8] Hartry Field also rejects the identification of disbelieving with believing the negation, but on somewhat different grounds: he thinks one should sometimes disbelieve (or, in his terminology, "reject") both p and not-p, though one should never believe both not-p and not-not-p. For Priest, disbelief is stronger than believing the negation; for Field, it is weaker.

[9] See Broome 2000:79–81.

[10] For this view, see, for example, Harman 1986.

(2) If $A, B \models C$, then if you have reason to believe A and believe B, you have reason to believe C. ("Reason to believe" is closed under logical consequence.)

But (2) is just Br+. Could anyone rationally accept Cr+ but *not* Br+? This would mean denying that "has a reason" is transmitted to logical consequences in general, while asserting that it is transmitted to the logical consequences of the beliefs one happens to have. It is difficult to see what could be said in favor of such a view, so I will assume that anyone who accepts Cr+ will accept Br+ as well. Similar considerations allow us to dismiss the other Cr variants.

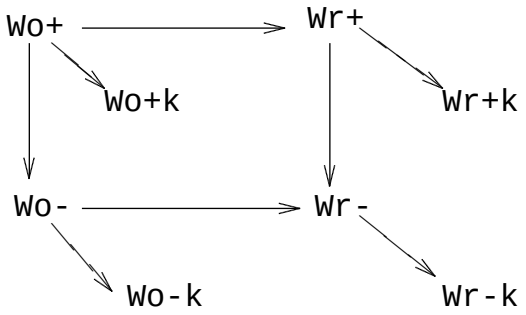
Let us now turn to the B's and W's. I think there is a clear sense in which the B's are too weak. According to the B's, logical consequence is a channel through which existing norms for belief (obligation, permission, reason) can be *extended*. For example, according to Bo+, if you ought to believe the premises of a valid argument, you ought to believe its conclusion. But what if you believe the premises, though you shouldn't believe them? Then Bo+ says absolutely nothing about what you should or may believe. Similarly, if you do not have reason to believe A and believe B, but you believe them anyway, Br+ says nothing at all about whether you have reason to believe their conjunction, $A \wedge B$. According to the B's, then, logic is only normative for those whose beliefs are already in order—that is, for those who believe what they ought to believe (or may believe, or have reason to believe). To the unfortunate others, logical norms simply do not apply. Maybe that is a consequence we can swallow, but it seems to me a strong reason to prefer the W's to the B's. If the W's turn out to be untenable, though, we have the B's to retreat to.[11] The B's are at least free of the objectionable features of the C's.

The W's also avoid the problems that face the C's. Suppose you believe A and believe B, and suppose you come to see that A and B logically imply C. According to Wo+, you must see to it that *if* you believe A and B, you also believe C. Clearly, you can do this either by coming to believe C or by abandoning your belief in A or B. So Wo+ accounts for the fact that the demands of logic can be satisfied by subtractions or changes in belief as well as additions. And, unlike the B's, the W's state norms that apply to all believers, not just to those who already believe what they ought to (or may, or have reason to).

Before we get down to the really hard choices, however, we can do some more pruning. The Wp's can be rejected right away. As we have seen, rules of proof—that is, rules of reasoning in the narrower, formal sense—are permissions. But logical norms for belief—for reasoning in the broad sense—are surely *constraints* of some kind. According to Wp+, you *may* see to it that if you believe A and $A \supset B$, you believe B. But you aren't required to do so. For all Wp+ says, you don't even have a *reason* to do so! The difference between the Wp's and the position that there are no logical norms for belief seems slim indeed.

[11] Note that while the Wo's imply the Bo's, under reasonable assumptions, the Wp's do not imply the Bp's or the Wr's the Br's. It could be that one has a reason to believe A and believe B and an incompatible reason to see to it that if one believes A and believes B, one believes C; in that case it would not follow that one has a reason to believe C. Similarly, it could be that one is permitted to see to it that one believes A and believes B and is permitted to see to it that if one believes A and believes B, one believes C, but not to do both at the same time.

We are left with eight contenders: the Wo's and the Wr's. Given the reasonable assumption that you have a reason to do what you ought to do, the Wo's imply the corresponding Wr's. Given the reasonable assumption that you ought to see to it that if you believe C, you do not disbelieve C, the positive polarity variants imply their negative polarity partners. And given that 'know' is factive, the k variants are implied by the corresponding non-k variants. So they can be arrayed by strength as follows:



What can be said for and against each of these variants? Here are some considerations for and against.

1. EXCESSIVE DEMANDS. Wo+ implies that you ought either to cease believing the axioms of Peano Arithmetic or come to believe all the theorems as well. We humans are presumably incapable of the latter. Even if a genie could grant us the capacity for arithmetical omniscience, it's not clear we'd have reason to accept it. Only a small number of the theorems are likely to be of any practical or theoretical use to us; why must we clutter up our minds with all the rest?[12] So it seems that the only feasible way for a human to comply with the norm would be by ceasing to believe the Peano axioms. The same argument could be made with regard to any inferentially fertile set of axioms. But it's crazy to think that logic (together with these premises about human capacities) forbids us from believing such axioms, even if they are consistent and true!

Wo+k avoids the problem, because we are not omniscient about what follows logically from the axioms. It does not seem unreasonable to say that if you *know* that something is a theorem of PA, you ought to believe it (or give up your believe in the axioms). Wo- avoids the problem in a different way. It forbids us to disbelieve the theorems while believing the axioms, but it allows us to take a neutral stance, neither believing nor disbelieving, to most of them. Wr+ implies only that if you believe the axioms, you have a *pro tanto* reason to believe each of the theorems. Since this reason can be overridden by other, more compelling reasons, Wr+ does not have the same objectionable upshot as Wo+.

2. THE PARADOX OF THE PREFACE. You have written an authoritative book about sea turtles. You believe each claim you make in the book. Yet you also believe, on general inductive grounds, that at least one of these claims is false. So you don't believe the conjunction of the claims in the book. Indeed, you disbelieve it, despite the fact that it is a (known) logical consequence of other things you believe. And your position seems quite rational. It would be crazy to give up either your belief in all of the claims you make

[12] For these "clutter" considerations, see Harman 1986, Sainsbury 2002: 1.

in the book, *or* your well-founded belief that at least one of them is false, just to avoid inconsistency. Yet it looks as if this is just what the `Wo`'s say you *must* do.

The `Wr`'s look better here. Perhaps you do have a *reason* to revise your beliefs. But because you have stronger, overriding reasons to keep them as they are, despite the inconsistency, the `Wr`'s do not imply that you ought to give them up.[13]

3. THE STRICTNESS TEST. Broome 2000 argues that "The relation between believing p and believing q [a logical consequence of p] is strict. If you believe p but not q, you are definitely not entirely as you ought to be" (85). The `Wr`'s do not capture this strictness. They allow that one might believe p but not its logical consequence q and still be just as one ought to be. To use Broome's terminology, believing p "recommends" believing q but does not "normatively require" it. Perhaps, in light of the preface paradox, this is the right result. If you are "entirely as you ought to be" in believing each of your claims while disbelieving their conjunction, then the preface case is a counterexample to Broome's strictness test. Still, there seems to be something to the strictness test, at least in ordinary cases. And perhaps hard cases make bad law.

4. THE PRIORITY QUESTION. According to the `-k` variants, we are subject to logical norms only in so far as we have logical knowledge. The more ignorant we are of what follows logically from what, the freer we are to believe whatever we please—however logically incoherent it is. But this looks backwards. We seek logical knowledge so that we will know how we ought to revise our beliefs: not just how we *will* be obligated to revise them when we acquire this logical knowledge, but how we are obligated to revise them even now, in our state of ignorance.

5. LOGICAL OBTUSENESS. Suppose someone believes A and believes B but just refuses to take a stand on their conjunction, $A \wedge B$. Intuitively, there is something wrong with her: she is being illogical. Yet according to `Wo-`, she is not violating any logical obligation in failing to revise her beliefs, and according to `Wr-`, she does not even have a reason to revise them—at least, not a reason that has anything to do with logic. That seems wrong, and it suggests that the negative-polarity variants are too weak.

We can sum up the rather confusing picture in tabular form:

	`Wo+`	`Wo+k`	`Wo-`	`Wo-k`	`Wr+`	`Wr+k`	`Wr-`	`Wr-k`
Demands	x							
Preface	x	x	x	x				
Strictness					x	x	x	x
Priority		x		x		x		x
Obtuseness			x	x			x	x

(where the x's after each consideration marks the variants it seems to rule out as too strong or too weak). It is not clear how one should weigh the various considerations against each

[13] Sainsbury 2002 uses the preface paradox to motivate a move from something like `Wo-k` to `Wr-k`.

other when they pull in different directions, as they frequently do. (It is notable that no column is free of a black mark.) I am curious to see what others think about this.[14]

My own temptation is to go for a combination of `Wo-` and `Wr+`. (Thus, I accept `Wo-`, `Wo-k` (which it implies), and all the `Wr`'s. I reject `Wo+` and `Wo+k`.) Let me explain why. I think that the excessive demand consideration shows that `Wo+` is too strong, and the priority consideration suggests that the `-k` variants are too weak. That leaves `Wo-` and the `Wr`'s. `Wo-` has two black marks on our chart, from the preface paradox and from logical obtuseness. The obtuseness worry is, I think, ameliorated if we *also* have `Wr+`. This allows us to say that the logically obtuse person at least has a *reason* to revise her beliefs. So the real problem is the preface paradox. If the preface paradox really shows that `Wo-` is too strong, then we have to rest content with the `Wr`'s. And that means learning to live with the idea that logical norms are not, as Broome puts it, "strict." But does the preface paradox really show what we took it to show?[15]

If we take strictness seriously, we must say that someone who believes A_1, A_2, ..., and A_n but disbelieves their conjunction is not "entirely as she ought to be." There is something *wrong* with her beliefs, and she would be remiss if she did not take steps to fix it. The preface paradox puts pressure on this idea by showing that it might be perfectly *rational*, all things considered, to retain this pattern of beliefs and disbeliefs. We would not criticize the preface believer for failing to revise her beliefs in this situation. Indeed, we *would* criticize her if she responded by giving up her belief in all of her claims about sea turtles *or* by giving up her well-founded fallibilism. Given that she is taking the best course of action available to her, she is surely "entirely as she ought to be."

But this reasoning is not sound. From the fact that it is more rational for our writer to keep her logically incoherent pattern of belief and disbelief than to revise it *in the ways now open to her*, it does not follow that she is not obligated to revise her beliefs and disbeliefs in a way that restores logical coherence. For it may be that making her beliefs coherent in any of the ways now open to her is *also* forbidden. She cannot give up *all* of her individual beliefs about sea turtles, because in each case she is under an obligation to believe that for which she has compelling evidence. But she cannot give up her general inductive disbelief in her own infallibility, either, because it too is well grounded in evidence. Thus she is under conflicting obligations. Whatever she does, she will not be "entirely as she ought to be."

[14] There is quite a bit of difference of opinion in the literature. Harman 1984 and 1986 seems to think none of them are any good. (For a remark on the `Wr`'s, see 1984: 109.) Broome 2000 goes for `Wo+`. Sainsbury 2002 goes for what is essentially `Wr-k`. Many writers talk about logical norms for reasoning in such a vague way that I am not sure what their positions are.

[15] Perhaps what the tension between the strictness test and the preface paradox shows is that we need a bridge principle of an entirely different form from any of those we have considered, one that is sensitive to the kinds of differences that distinguish the preface case, in which it is permissible to hold inconsistent beliefs, from ordinary cases, in which it is not. But it seems to me that formulating such a principle would require us to abandon talk of "belief" and "disbelief" for talk of "degrees of belief," and perhaps to generalize our talk of "logical consequence" to talk of "degrees of evidential support." This is a worthy project, but I want to confine myself here to the non-quantitative framework with which we started.

Thus it is only if one denies (or ignores) the possibility of conflicting obligations that the preface paradox appears to count against `Wo-`. In some contexts it is reasonable to deny that there can be conflicting obligations. Many philosophers have thought, for example, that genuine moral obligations (unlike moral reasons) cannot conflict with each other: there is always a morally permissible action available to the agent, no matter what her situation. I am not sure why we should believe this, but even if it is right, the point cannot be generalized to all kinds of obligation. Clearly there can be conflicting *legal* obligations, if the law is badly designed. Even more clearly, obligations of one kind can conflict with obligations of another kind: for example, moral obligations can conflict with legal obligations, so that one must choose between doing what is forbidden by the law or doing what is morally forbidden. So it would seem rash to assume that logical obligations cannot conflict with other kinds of broadly epistemic obligations. But it is only if we assume this that the preface paradox tells against the `Wo`'s.

Indeed, the preface puzzle is just one of a large number of structurally identical puzzles about conflicting obligations in non-ideal circumstances. Let's look at another one, which has nothing to do with norms for belief. You are the manager of a widget factory. According to the laws of your country, you are legally obligated not to employ any illegal aliens. However, the law (unlike current U.S. law) does not set up any formal document-checking procedures. You check each applicant's credentials carefully, but you know that this process is not infallible, and based on your experience with other widget factories, you believe with certainty that at least one of your employees is an illegal alien. What should you do? At present, you have no way to determine which employees are illegal, beyond the screening you have already done. Firing them all would violate your obligation to the company's board and shareholders. Firing some of them arbitrarily would not guarantee compliance with the law, and would violate moral and perhaps legal obligations to the fired workers. It seems best, all things considered, not to fire any of them. Yet that means violating a legal obligation. Thus you are not "entirely as you ought to be." What we have here, as in the preface paradox, is a situation in which

1. You ought to see to it that P.
2. You ought to see to it that Q.
3. None of the actions currently available to you will bring about both P and Q.

No matter what one does, in such a case, one fails to do what one ought to do.

What about the fact that we don't think a person in such a situation merits *criticism* for her failure to satisfy both of the putative obligations? Doesn't that show that one of them was not really an obligation at all? No, it does not. The argument assumes too direct a relation between obligations and criticism. One can avoid criticism for a failure to live up to one's obligations if one has the right kind of *excuse*. And conflicting obligations are often exculpatory in this way. (Not always, though: it may be the agent's own fault that she was forced into the choice.) Imagine yourself again as the manager of the widget factory. Would anyone seriously maintain that because you do not merit criticism for your

actions—you're doing the best thing you can do in the circumstances—you do not in fact have a legal obligation not to employ illegal aliens?

So far I have employed the widget factory analogy to show how little force the preface paradox carries against Wo-. But there is an important point of disanalogy that offers an even more direct response to the preface paradox. In the widget factory case, you really could satisfy the obligation by firing everyone. Perhaps that is what you *would* do if you weren't under any other obligations in addition. The analogous course of action in the preface case would be to abandon all of your beliefs about sea turtles. But unlike firing all your employees, this is not really an open course of action at all. If you really believe all these things about sea turtles, you can't just *decide* to stop believing them. Unless, perhaps, you're insane, you can't just say, "Well, of course there are compelling reasons for thinking these claims are true, but I'm not going to believe them." Belief is not voluntary in that way.[16] Nor can you just decide to stop disbelieving their conjunction, about which you have plenty of reason for scepticism. So the problem isn't just that although both of these courses of action are available to you, both violate (conflicting) obligations. The problem is that neither of them is really available to you at all.

Suppose you wanted to obey the norm implied by Wo-:

> You must see to it that either you do not disbelieve the conjunction of your claims about sea turtles, or you do not believe any (or most) of these claims.

How would you go about it? You would step up critical examination of these claims. You would do more studies, try harder to embed them in established theory, publish them so that others can scrutinize them, and do all of the things a good scientist does. That is, you would change your beliefs indirectly, by seeking new evidence, not by force of will. This is the only course of action open to you that could conceivably count as seeing to it that your beliefs and disbeliefs are revised for coherence. But doing this is not foolish at all. This is precisely what you *ought* to do. What would be foolish is to sit back and say, "Well, there's an inconsistency in my beliefs, but so what? The preface paradox shows that that's okay."

To sum up, then, the preface paradox does not show that the Wo's are too strong. When one finds oneself with preface-like beliefs and disbeliefs, one ought to see to it that they change. What this involves is not *deciding* to change the beliefs, however, but seeking further evidence that might change one's mind. What is more, the argument from the preface paradox against the Wo's seems to assume without justification that there are no genuinely conflicting epistemic norms. That is not to say that Wo+ is tenable: it is ruled out by its excessive demandingness. Since the priority considerations rule out Wo+k and Wo-k, we are left with Wo-. The only remaining problem with Wo- is that it does not censure the "logically obtuse" person who refuses to believe the conjunction of two things she believes, even when there is no reason not to. We solve that problem by accepting Wr+ as well.

[16]The locus classicus for this point is Williams 1973.

4 Three applications

I now want to consider (very briefly) the implications of various versions of the bridge principle for some central questions about the notion of logical validity. Once we have settled on a view about the relation between validity and norms for belief, we can appeal to our independent grip on the latter to help resolve questions about the former. There won't be time here for an adequate resolution of any of these questions; my comments are intended only to illustrate a methodology that I find more satisfactory than the usual intuition-mongering.

4.1 Relevance

Let us start with the debate between relevantists and classicists. Classicists hold that for any A and B, no matter how tenuously related in subject matter, the inference from A and $\neg A$ to B is valid. Relevantists deny this. Can we help resolve this debate by considering the norms to which the debated consequence would give rise?

It would be easy to argue for relevantism using `Cp+` as a bridge principle. One need only note that people who find themselves with contradictory beliefs (perhaps for preface-like reasons) are not thereby entitled to believe *everything*. But as we have seen, there are plenty of reasons to reject `Cp+`. Can a "normative" case for relevantism also be made using a more defensible version of the bridge principle, like `Wo-`, `Wo+k`, or `Wr+`? Take a typical "irrelevant" instance of *ex falso*:

(I) Frank is six feet tall, Frank is not six feet tall / Lisa has a pet iguana.

If (I) is valid and, let us suppose, known to be valid, then our three candidate bridge principles generate the following norms:

($N_{\texttt{Wo}-}$) You ought to see to it that: either you do not believe both that Frank is six feet tall and that Frank is not six feet tall, or you do not disbelieve that Lisa has a pet iguana.[17]

($N_{\texttt{Wo}+\texttt{k}}$) You ought to see to it that: either you do not believe both that Frank is six feet tall and that Frank is not six feet tall, or you do believe that Lisa has a pet iguana.

($N_{\texttt{Wr}+}$) You have reason to see to it that: either you do not believe both that Frank is six feet tall and that Frank is not six feet tall, or you do believe that Lisa has a pet iguana.

Is it plausible to think that there are such norms? Do I really have an obligation, or even a reason, to adjust my beliefs so that either I do not have contradictory beliefs about Frank or I believe that Lisa has a pet iguana? If any of our bridge principles is correct, that is what (I) implies. So we can argue against (I) by arguing against one of these putative norms for belief and employing the appropriate bridge principle.

The classicist might defend these norms as follows:

[17] I've converted the material conditional into a disjunction, for clarity.

(1) You ought to see to it that your beliefs are consistent (because inconsistent beliefs cannot all be true).

(2) If you believe both that Frank is six feet tall and that he is not six feet tall, then your beliefs are not consistent.

(3) So you ought to see to it that you do not believe both that Frank is six feet tall and that he is not six feet tall. [from (1) and (2)]

(4) But if you do not believe both that Frank is six feet tall and that he is not six feet tall, then it follows logically that either you do not believe both that Frank is six feet tall and that he is not six feet tall or you believe that Lisa has a pet iguana. [logic[18]]

(5) Thus if you see to it that you do not believe both that Frank is six feet tall and that he is not six feet tall, then you see to it that either you do not believe both that Frank is six feet tall and that he is not six tall or you believe that Lisa has a pet iguana. [from (4)]

(6) Thus if you ought to see to it that you do not believe both that Frank is six feet tall and that he is not six feet tall, then you ought to see to it that either you do not believe both that Frank is six feet tall and that he is not six tall or you believe that Lisa has a pet iguana. [from (5)]

(7) Thus you ought to see to it that either you do not believe both that Frank is six feet tall and that he is not six tall or you believe that Lisa has a pet iguana. [from (3) and (6)]

In this way, the classicist might seek to show that we are forced to accept ($\mathrm{N}_{\mathtt{Wo+k}}$) and the other odd norms that follow from *ex falso quodlibet* and our bridge principles. (The argument can be modified in obvious ways to apply to ($\mathrm{N}_{\mathtt{Wo-}}$) and ($\mathrm{N}_{\mathtt{Wr+}}$).)

But the reasoning in this little argument is fallacious. The step from (4) to (5) makes tacit use of the principle:

(P) If you see to it that A, and B follows logically from A, then you also see to it that B.

This is a principle that ought to be rejected (and *is* rejected in the most fully developed logic of agency I know of, that of Belnap, Perloff, and Xu 2001). I saw to it that I ate cereal for breakfast this morning: that was something that I did. But I did not see to it that I either ate cereal for breakfast or did not eat cereal for breakfast. I had no role in making-true this tautology. Thus the classicist's argument fails, and the classicist is left with no explanation of why ($\mathrm{N}_{\mathtt{Wo-}}$), ($\mathrm{N}_{\mathtt{Wo+k}}$), or ($\mathrm{N}_{\mathtt{Wr+}}$) should be true.[19]

[18]Note that the inference from A to 'A or B' is valid in the relevance logic of first degree entailment, so the classicist isn't begging any questions against the relevantist here.

[19]Note that because the dubious step does not involve the deontic modality at all, the same fallacy infects the versions of the argument directed at ($\mathrm{N}_{\mathtt{Wo-}}$) and ($\mathrm{N}_{\mathtt{Wr+}}$) as well.

4.2 Necessity

It is generally assumed that a valid inference must be materially truth preserving:

(MTP) For all A, B, C, if $A, B \models C$, then it is not the case that A and B are true and C false.

A minimal notion of validity would take this necessary condition to be sufficient as well:

(Min) $A, B \models C$ iff it is not the case that A and B are true and C false.

Most people think that (Min) is utterly unacceptable as an account of validity. To say that C follows from A and B, they say, is to say at least that the transition from A and B to C is *necessarily* truth-preserving. My question is this: how might we justify this claim, other than by appealing to our "intuitions" about validity or consequence?

We might naturally appeal to the relation between validity and norms for reasoning. According to (Min),

$2 + 2 = 4 \models$ New Mexico is larger than Arizona.

just because the latter happens to be true. But surely belief in the former does not compel, permit, or even give reason for belief in the latter. Therefore, (Min) is inadequate as an account of validity.

It should be clear by now that this reasoning is too fast. It assumes an account of the normativity of logic—the `C`'s—that we have seen reason to reject. Can we reconstruct it using a more plausible version of the bridge principle: `Wo-`, `Wo+k`, or `Wr+`? It turns out that the soundness of the argument depends crucially on the bridge principle we use. `Wo+k`, together with (Min), gives us

> If you know that it is either false that $2 + 2 = 4$ or true that New Mexico is larger than Arizona, then you ought to see to it that you do not both believe that $2 + 2 = 4$ and not believe that New Mexico is larger than Arizona.

There is nothing wrong with this: indeed, it looks straightforwardly true. If you know that it is not the case that the premise is true and the conclusion false, then you should see to it that you do not believe the former without believing the latter. So if we adopt `Wo+k` as our bridge principle, we will be hard pressed to explain why (Min) is not an adequate account of validity. Someone who is antecedently prejudiced against (Min) might take this as another reason to reject `Wo+k`.

If we adopt `Wo-` or `Wr+`, the argument against (Min) looks much more plausible. `Wo-`, together with (Min), gives us

> If it is either false that $2 + 2 = 4$ or true that New Mexico is larger than Arizona, then you ought to see to it that you do not both believe that $2 + 2 = 4$ and disbelieve that New Mexico is larger than Arizona.

Since the antecedent is true, we can detach and get

> You ought to see to it that you do not both believe that $2+2=4$ and disbelieve that New Mexico is larger than Arizona.

This norm is supposed to apply even to those who don't know *anything* about the truth values of the premise and the conclusion or how they are related. Surely that's implausible.

Worse yet, we can substitute *any* truth for "New Mexico is larger than Arizona" and get another norm. Thus the combination of (Min) and `Wo-` requires not just logical coherence in our beliefs (like `Wo-` and regular logical validity), nor even just logical closure (like `Wo+` and regular validity), but a kind of *factual* perfectionism! Anyone who believes $2+2=4$ and disbelieves any true proposition at all is in violation of the norm. That's about as plausible as a general epistemic norm forbidding false belief. To avoid commitment to such norms, we must reject either (Min) or `Wo-`. Since there are good reasons for accepting `Wo-`, we should reject (Min).

A similar argument can be run with `Wr+` in place of `Wo-`. It is implausible to think that someone with no information about the relative sizes of New Mexico and Arizona has any reason to see to it that she she does not both believe $2+2=4$ and not believe that New Mexico is larger than Arizona. I suppose someone might try to defend this consequence of `Wr+` and (Min) as follows:

(1) Our agent has a reason to see to it that she believes what is true.

(2) It is true that New Mexico is larger than Arizona.

(3) Thus, our agent has a reason to see to it that she believes that New Mexico is larger than Arizona. [from (1) and (2)]

(4) But if one has a reason to see to it that A, one has a reason to see to it that $\neg(B \wedge \neg A)$.

(5) Thus, our agent has a reason to see to it that she does not both believe that $2+2=4$ and not believe that New Mexico is larger than Arizona. [from (3) and (4)]

But there are two flaws in this line of thought. First, the move from (1) and (2) to (3) is invalid. (1) is ambiguous between

(1a) For all A, if A is true, then our agent has reason to see to it that she believes A, and

(1b) For all A, our agent has reason to see to it that if A is true, she believes A.

(1a) is implausible; it is the availability of (1b) that makes (1) sound like a plausible premise. But (1b) and (2) do not imply (3).[20]

Second, (4) is false. Consider the following counterexample. Joe has reason to work overtime every week. (He has an apartment that is a bit beyond what his regular salary can

[20] If this is not obvious, consider the following counterexample. I have reason to see to it that if the window is open, the heater is turned off. The window is open, and I'm freezing cold. Do I have reason to turn off the heater?

pay for.) It doesn't follow that he has reason to see to it that he does not both get a higher paying job and not work overtime every week.

Thus the argument has only surface plausibility. We are left without any reason to suppose that we are bound by the norms that (Min) and `Wr+` imply. Since `Wr+` is plausible—even more plausible than `Wo-`—we have good reason to reject (Min).

4.3 Formality

Consider the following inference:

(1) Cicero talked the talk.

(2) Tully walked the walk.

(3) Someone walked the walk and talked the talk.

Is it logically valid? The premises are certainly relevant to the conclusion, and the inference is necessarily truth-preserving.[21] Yet many philosophers would deny that the conclusion is a logical consequence of the premises, on the grounds that even though the inference is necessarily truth-preserving, it does not owe this status to its *form*. To get a genuinely valid inference, we would need the proper names in (1) and (2) to be identical:

(1) Cicero talked the talk.

(2*) Cicero walked the walk.

(3) Someone walked the walk and talked the talk.

What is the motivation for this additional condition on logical validity? It seems to me that it must be something like this: the connection between (1), (2*) and (3) is *transparent* in a way that the connection between (1), (2) and (3) is not. Someone who did not know that 'Cicero' and 'Tully' were names for the same person would not be able to discern the necessary connection between (1), (2) and (3). By contrast, the connection between (1), (2*) and (3) is evident even to someone who has no idea who 'Cicero' refers to.

But why is it important that logical validity be transparent in this way? I would like to suggest that it is important because of the *normative* implications of logical validity. Someone who accepted (1) and (2) but did not know that Cicero was Tully would not have a reason to accept (3) or revise her belief in (1) and (2). Nor would she have an obligation not to reject (3). Plausibly, then, we require logical validity to be formal because

[21] Indeed, it is necessarily truth-preserving in at least two reasonable senses. (a) In any possible context of utterance at which the premises are true, the conclusion is true (because 'Cicero' and 'Tully' are not context-sensitive terms). (b) At any possible circumstance of evaluation at which the premises are true, the conclusion is true (because 'Cicero' and 'Tully' are rigid designators).

we require it to be transparent, and we require it to be transparent because of the reasons and responsibilities to which it gives rise.[22]

If this line of thought is correct, then the connection between formality and transparency is a very tight one. If the whole *raison d'être* of the formality requirement is to ensure transparency, then the formal structure of an inference had better always be transparently discernible. The problem is that most accounts of formal structure make it difficult to see how this can be the case.

Let us focus, for concreteness, on the formal difference between the two inferences above: (1-2-3) and (1-2*-3). Intuitively, it is this: in (1-2*-3), the same name, 'Cicero', occurs twice whereas in (1-2-3), two different names occur. But what does it mean to say that the "same name" occurs? Surely it is not enough that the letters C-I-C-E-R-O appear twice: different people can have the same name in that "orthographic" sense. Nor is it enough that it is the same orthographic name with the same reference: we can concoct cases where that happens by coincidence, and in any case coreference is not, in general, transparent.[23] We might seek to individuate names more finely than orthography and reference, say, by grouping uses of them into chains linked together by speakers' intentions to corefer. But because these individuating factors are largely sociological, it will not in general be transparent whether two name tokens belong to the same type in this sense. Clearly, if we are to have transparency, we're going to have to look inside the head, at the way the thinker *thinks* the names. Accordingly, some have suggested that the relevant sameness relation is sameness of *mentalese* name: syntactic sameness in the language of thought.[24] But it is not at all clear that *this* is transparent, either. How am I supposed to know whether the subpersonal representational vehicles of two thoughts have a component in common? Isn't that for the experimental psychologist to determine?[25] (Of course, someone whose belief revision did not manifest sensitivity to sameness of representational vehicle would be cognitively defective, but not in a way that would warrant epistemic blame. She could more appropriately be said to be *malfunctioning*; she would need a brain doctor, not criticism.)

Where does this leave us? We could settle for a way of understanding formal structure on which it is not fully transparent. But if we go that way, we lose the coherent rationale for the formality requirement on logical validity that I sketched above. Why should we insist that logical validity be formal if formal structure is not fully transparent? Why should we *care* about formality? I have never heard a satisfactory answer to this question.[26]

---

[22] I am not sure I want to *endorse* this line of thought. Perhaps an alternative would be to say that the obligations are the same in both cases, but in the (1-2-3) case the agent has an *excuse* for failing to meet them.

[23] One can know to whom A refers, and to whom B refers—in a perfectly good ordinary sense of "knowing who"—and still not know that they refer to the same person. If we insisted that in order to understand a proper name, one had to know to whom it referred in a way that precluded this kind of ignorance, no one would ever count as understanding a proper name.

[24] For this approach, see Fodor 1990: 176 n. 10: "…according to the present view, questions of rationality are assessed with respect to the vehicle of a belief as well as its content; whereas questions of truth are assessed with respect to content alone …."

[25] For a similar point, see Millikan 1994: 92.

[26] One sometimes hears the following *un*satisfactory answer: "Logic is an *a priori* discipline and cannot possibly be expected to tell us about the validity of inferences like 'Cicero was Roman, therefore, Tully was

I want to suggest that we made a wrong turn at the very beginning of this line of thought, in thinking of formal validity as a *species* of validity. That's what led to the quixotic quest for epistemically transparent formal structure in concrete inferences. So let's not think of formal validity as a kind of validity, or as property of inferences at all. Instead let's think of it as a property of inference *schemata*.[27] There's no interesting question about whether the formal structure of an inference *schema* is transparent, because grasping an inference schema just *is* grasping a formal structure.[28] If the normativity of logic has its source not in the formal validity of inferences, but in the formal validity of inference schemata, we can avoid the transparency problem completely. Our bridge principles will need to have a slightly more complex form:

> BRIDGE PRINCIPLE: If [you know that[29]] the schema S is formally valid and you apprehend the inference A, B / C as an instance of S, then (normative claim about believing A, B, and C).[30]

On this view, all logical norms have their source in the thinker's "apprehension" of inferences as having a certain formal structure.[31] Nothing turns on whether the inferences *really* have this formal structure, or even on whether there is a fact of the matter about this. So the view is compatible with a wide range of views about the formal structure of *inferences*, including the view that inferences don't have formal structures, the view that they have structures but these are not epistemically transparent to us, and the view that it can be indeterminate what the formal structure of an inference is. Even if the classical view that each inference has a unique and epistemically transparent formal structure is correct, the fact that an inference *has* a particular formal structure does not give rise to any norms; what matters is what formal structure the inference is apprehended as having.

My talk of "apprehension" is meant to be a placeholder for a fuller account. My own view is that apprehension should not be intellectualized to the extent that it requires a completely explicit understanding of what an inference schema is, the kind one would get

---

Roman.' Thus logic must concern itself only with *formal* validity." But this this is a bit like saying that theoretical physics studies only *ideal* matter, on the grounds that a physicist won't be able to tell you exactly how a particular portion of non-ideal gas will behave. In general, a discipline that studies X will not concern itself with *every* X-related question. Chemists study acids and bases, but they can't tell you how acidic your soil should be if you want to grow lilacs. Cf. Coffa 1975, Read 1994.

[27]Note that it is inference schemata, not inferences, to which the model-theoretic definition of validity applies. The "nonlogical constants" are schematic, not interpreted terms. Otherwise there would be a real question why we should look at different "reinterpretations" of them in defininig validity (cf. Etchemendy 1990).

[28]In our usual logical systems, we adopt the following representational convention: if the same letter appears twice in an inference schema, it indicates a formal "link." Different conventions are possible—for example, we might use lines to indicate "links"—but this one is simple and convenient.

[29]Optional, for the `-k` variants.

[30]Or should the "apprehension" part be subsumed under the deontic operator? For now I'll stick with the simpler formulation.

[31]I do not use "apprehend" as a success word: from "X apprehends I as an instance of S," it does not follow that I *is* an instance of S. Perhaps this is a grammatical mistake. In that case, I need a new word.

from an encyclopedia article on the subject. It is something more basic than that. But it is important that apprehension be something for which one can take responsibility and give or receive criticism. So it presumably requires at least a rudimentary grasp of the system of concepts that would be spelled out more fully in the encyclopedia article. On this view, then, logical norms only apply to creatures who have acquired certain capacities for recognition, classification, and criticism: fairly sophisticated creatures. That seems right: it seems odd to criticize a dog for thinking *illogically*. For now, I want to leave open a wide range of possible views about apprehension—even the strong "Kantian" view that logical apprehension is necessary for any thought at all.

Let's look at how this view applies to a couple of cases. First, the case with which we started. Intuitively, the inference (1-2-3) gives you no reason to revise your beliefs, while (1-2*-3) does. On the present view, we can explain this as follows. The inference schema

(S1) $Fa, Ga/\exists x(Fx \wedge Gx)$

is formally valid, while the schema

(S2) $Fa, Gb/\exists x(Fx \wedge Gx)$

is not. So one who apprehends (1-2-3) as an instance of (S2) but not (S1) has no reason or obligation to revise her beliefs, while one who apprehends (1-2*-3) (or, for that matter, (1-2-3)) as an instance of (S1) does have such a reason. That seems just right.

Second, consider the case of Peter, who seems to believe both that Paderewski is a good musician and that Paderewski is not a good musician (Kripke 1979). It seems out of place to hold him responsible for having contradictory beliefs, when he cannot see that they are contradictory. The usual responses to the problem are (a) to concede that his beliefs are contradictory, but weaken the connection between contradictoriness and norms for belief, or (b) to keep this connection by arguing that his beliefs are not inconsistent, since Peter has two "internal names" for Paderewski. The present approach offers a third alternative. On this approach, Peter is not subject to the logical norm because he does not apprehend his beliefs as instances of the schema $P, \neg P$. The issue of whether he has satisfied logical norms thus comes apart from the issue of whether his beliefs really are contradictory, and we need not choose between the unattractive options (a) and (b).[32]

Third, consider cases of "ontological confusion" of the type discussed by Camp 2002.[33] You are watching your suitcase go down the conveyor belt at the airport. Right after putting it on the belt, you think to yourself "I did not zip that suitcase." You track it as it goes behind a small wall and reemerges. At this point you see that it is buckled. You conclude "That suitcase is buckled but not zipped." Unbeknownst to you, though, it is a different suitcase.

[32](a) is unattractive because it severs important links between logical relations (like contradictoriness) and norms for belief. (b) is unattractive because it ignores the semantic and syntactic importance of the fact that Peter intends his uses of "Paderewski" to be uses of the same name used by those from whom he picked it up.

[33]Cf. the "slow switching" thought experiments discussed in Boghossian 1994. But we need not go all the way to Twin Earth.

Yours fell off the belt behind the wall, and this is another, very similar one belonging to another passenger. Have you reasoned badly? On the one hand, you have done something very much like equivocating. At each moment you have intended to refer to the suitcase you are looking at, and this has not always been the same suitcase. On the other hand, you have intended to refer to the same suitcase in all of your "That suitcase" thoughts. From your point of view, at least, there is no equivocation. If we take the standard approach, on which norms for reasoning derive from facts (or perhaps known facts) about the validity or invalidity of inferences, then we cannot determine whether you have reasoned badly without determining whether your inference was *in fact* formally valid. And this requires determining whether the three instances of "That suitcase" are formally "linked." It's very difficult to give a definite answer to this question, since the "internal," anaphoric links pull in a different direction from the "external," informational links, and *both* seem relevant. If we take the approach I am recommending, however, the answer to our normative question depends only on whether you *apprehend* the inference as an instance of a valid form. Pretty clearly you do, so on this account you have not violated any logical obligations.

It might be objected that this account restricts the application of logical norms *too* far. Shouldn't we sometimes be held accountable for failing to apprehend logical structure that really is there, or for taking there to be logical structure that isn't there? Sure. But I am inclined to keep these norms *for* apprehension separate from the logical norms that arise *from* the apprehension of inferences as instances of formally valid schemata. The former seem to group together with general epistemic norms enjoining careful observation and thorough investigation, not with specifically logical norms.

Notice, finally, that if we modify the general form of our bridge principles in this way, the one compelling argument against `Wo+`—the argument from excessive demandingness—loses its force. Though it is possible to infer infinitely many consequences from the Peano axioms, we do not apprehend all of these potential inferences as instances of formally valid schemata. So `Wo+`—on the revised plan—does not obligate us to have deductively closed sets of belief. In the end, then, we can plump for the following bridge principle:

> BRIDGE PRINCIPLE: If schema S is formally valid and you apprehend the inference A, B / C as an instance of S, then you ought to see to it that if you believe A and you believe B, you believe C.

5 Conclusion

My main aim in this paper has been methodological. I want to illustrate a way in which fundamental issues in the philosophy of logic can be addressed systematically and fruitfully, without the constant recourse to "intuitions" about validity or consequence that inevitably leads to deadlock (your intuitions against mine, end of story). I've suggested that we can make progress on these issues by trying to get clearer about the normativity of logic: the relation between facts about validity and norms for reasoning (in the broad sense of "reasoned change in view"). To this end, I've explored various candidate "bridge principles"

and the considerations that favor some over others. Finally, I've shown how progress in this area can illuminate debates about the place of relevance, necessity, and formality in our concept of logical validity. Needless to say, all of this is just a beginning.

References

Belnap, N., M. Perloff, and M. Xu (2001). *Facing the Future: Agents and Choices in Our Indeterministic World*. Oxford: Oxford University Press.

Boghossian, P. A. (1994). The transparency of mental content. *Philosophical Perspectives 8*, 33–50.

Broome, J. (2000). Normative requirements. In J. Dancy (Ed.), *Normativity*. Oxford: Blackwell.

Camp, J. (2002). *Confusion: A Study in the Theory of Knowledge*. Cambridge: Harvard University Press.

Coffa, J. A. (1975). Machian logic. *Communication and Cognition 8*, 103–129.

Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (1990). *A Theory of Content and Other Essays*. Cambridge: MIT Press.

Harman, G. (1984). Logic and reasoning. *Synthese 60*, 107–127.

Harman, G. (1986). *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press.

Kripke, S. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and Use*, pp. 239–283. Dordrecht: Reidel.

Millikan, R. (1994). On unclear and indistinct ideas. *Philosophical Perspectives 8*, 75–100.

Priest, G. (1979). Two dogmas of quineanism. *Philosophical Quarterly 117*, 289–301.

Priest, G. (1998). What is so bad about contradictions? *Journal of Philosophy 98*, 410–426.

Priest, G. (1999). Validity. *European Review of Philosophy 4*, 183–206.

Read, S. (1994). Formal and material consequence. *Journal of Philosophical Logic 23*, 247–265.

Sainsbury, R. M. (2002). What logic should we think with? In A. O'Hear (Ed.), *Logic, Thought and Language*. Cambridge: Cambridge University Press.

Williams, B. (1973). Deciding to believe. In *Problems of the Self*, pp. 136–151. Cambridge: Cambridge University Press.