Forthcoming in *Truth 20/20*, edited by Drew Johnson and
Adam Podlaskowski, Synthese Library book series, Springer.

Validity as Truth-Conduciveness

Arvid Båve

Deductive logic begins with the observation that some deductive inferences are good and some bad, and goes on to systematize these first givens, partly by explicating the notion of goodness involved. At an early point in the history of deductive logic, it was decided that if there is a possible case in which the premises of a deductive inference are true but the conclusion false, then that general form of inference is among the *bad* inferences. Thus, a deductive inference rule was thought to be good only if it is *universally truth-preserving*. For present purposes, we can take this to mean that every instance of the rule (i.e., arguments whose premises and conclusions are real propositions or sentences, rather than schemata) is such that in every possible world in which its premises are true, its conclusion is true. (I thus commit, at least for present purposes, to a form of logical monism, on which possible worlds are the privileged kind of "case" in terms of which validity should be defined—see, e.g., Field (2009: 264ff.) for motivations.)

Following Thomas Hofweber (2007, 2009, 2021), let's call this property of an inference rule *strict validity*. This, or something very similar, has been assumed to be the central property common to the "good" deductive inferences, and the right explication of this pretheoretic notion. In this paper, I try to contribute to Hofweber's case against these traditional and still prevailing views.

He mainly bases his views on findings from the work on truth-paradoxes like the Liar and Curry's paradox. These paradoxes show that the classical inference rules and the unrestricted truth rules (i.e., "$p \Rightarrow \mathrm{T}(<p>)$" and "$\mathrm{T}(<p>) \Rightarrow p$") cannot all be strictly valid (on the plausible assumption that not every sentence is true). Hofweber, however, claims that they can still all be *generically* valid, and that this is the good-making property of inferences that is central to logic. (Proposals related to, but importantly different from Hofwever's, include Bruno Whittle (2021), who makes the stronger claim that *all* deductive rules have exceptions, as well as the "logical nihilist" views of Gillian Russell (2018) and Aaron Cotnoir (2018). Steven Daglish (2020) is another relevant contender.)

As indicated by this last paragraph, I will take "deductive inferences" to include not merely the usual simple inferences involving concepts like *and*, *or*, *some*, and *every*, but also inferences involving the concept *true*. This is only for convenience, however, and not intended to involve any contentious view about how "logicality" is to be demarcated. Thus, I am neutral on whether the two are different in kind, perhaps differing in that only the usual set of inferences are "formal" in some sense.

That a deductive inference rule R is generically valid means that the following generic is true:

Instances of R are truth-preserving,

where an instance of a rule is truth-preserving just in case in every world in which its premises are true, its conclusion is true. This, rather than strict validity, is what we should take as defining the notion of a good inference, according to Hofweber.

In parallel with this weakened notion of validity, I will be arguing that there is also a central *normative* characteristic of inference rules, namely, *defeasible rationality*. This property applies to a rule just in case one is always entitled to make an inference instantiating the rule, except where there is a defeater (for an overview of defeasibility in epistemology, see Grundmann (2011)). This notion will be considerably sharpened, with further preconditions added, in §8. I follow Dretske (2000) and others in taking entitlement as a fairly weak epistemic positive status, certainly weaker than that of *having a reason for* one's belief or inference. Since the notion of defeasible rationality and that of a defeater are central to epistemology, both are independently motivated.

Now, for an important class of generic statements of the form *Fs are G*, Hofweber says, if it is true then the inference *Fa* to *Ga* is defeasibly rational. I think we should assume that such an inference can be defeated even if the generic is still assertible. In any case, the generic figuring in the truth-preservation claim above, and thus in Hofweber's alternative definition of validity, should be taken to belong to this class. We will find reason to doubt this way of proceeding below.

Note that on the strict conception of logic, one can typically speak interchangeably of inference rules being good and of inferences being good. This is not so if one adopts the weaker notion of generic validity (it is weaker since an inference being strictly entails its being generically valid, but not *vice versa*). On the strict conception, a particular deductive inference is good just in case it instantiates a good rule. But if validity is rather something like generic validity, then an inference can be bad (both in the sense of not preserving truth and of being irrational) even if it instantiates a good rule. This is because a rule being good just means that it is valid in the weaker sense, which allows exceptions.

§1 lays out a fundamental assumption of the paper, that logic is concerned *inter alia* with norms of belief and inference and therefore intersects with epistemology. §2 criticizes Hofweber's notion of generic validity and recommends instead that logical validity be identified with truth-conduciveness. §3 discusses to what extent Hofweber's views are really substantive, and to what extent the matter is rather terminological. It concludes that there is a substantive position in the offing, but only if, in some sense, "there is no solution to the semantic paradoxes". In §4, I argue that there is an important sense in which there isn't, namely, in that no standard solution to the paradoxes, operating with strict validity, can be known. This is based on the conjecture that no such solution will score high enough relative to the desiderata on logic to merit justified belief. §5 makes some related points about some recent "substructural" solutions. §6 argues that the unknowability of which strict solution is true is best explained by its being *metaphysically indeterminate* which one is. This section also attempts to provide a broader explanation of how the unknowability of these matters of strict validity is unsurprising, and therefore a more palatable finding than it may first appear. §7 argues that a logic that takes validity to be truth-conduciveness performs markedly better with respect to the desiderata on logic. §8 discusses the notion of defeasible rationality, which is held as the central *normative* concept in logic, and argues that a logic operating with the twin concepts of truth-conduciveness and defeasible rationality has an attractive unity as well as a number of other important benefits. §9, finally, provides some additional arguments for thinking that some strictly invalid modes of inference can be rational.

1. Logic and epistemology

I will argue that validity should be identified with truth-conduciveness, rather than generic validity. Truth-conduciveness is central to reliabilist theories of epistemic justification, and is roughly the property of having a high ratio of true "output beliefs", given true premises (and "correct application", in the case of induction). Thus, perception is truth-conducive because, usually, when it looks to someone as if p, then p. Analogously for induction: when the relevant preconditions are met (the observations of particular cases being veridical, the sample being suitably unbiased, etc.), the generalization based on the inductive inference is usually true. In the case of deductive inference rules, I propose that the relevant property is that of having a (very) high ratio of necessarily truth-preserving instances. That is, an inference rule is truth-conducive just in case almost every instance of the rule is such that, for every possible world, if its premises are true in that world, so is its conclusion. Here, an instance of a rule is a collection of actual propositions or sentences that instantiate the relevant logical forms. Thus,

Snow is white and grass is green ⊢ Snow is white

is an instance of the rule A and $B \vdash A$, and so on.

This kind of "frequentist" understanding of reliability or truth-conduciveness raises difficult problems with respect to several types of belief-forming process, especially relating to choices of "totalities" of token processes relative to which we are to decide the ratios of true output beliefs. If, for instance, a person has been very unlucky with their inductive reasoning, should we conclude that induction is unreliable for them, or should we take the totality of inductive inferences to be those that have been undergone by humans in general? Or should we widen the totality further and take it to include every inductive inference in every possible world? But how then do we know whether the process really is reliable?

Many of these questions do not arise with respect to deduction, since we do not speak of inferences having been made at all, but, more abstractly, of the truth-values of premises and conclusions. A different problem does arise here, however, which is that there are uncountably many instances of any rule. I discuss this problem and attempt a solution in the next section. Apart from this, which concerns the present, rather idiosyncratic extension of truth-conduciveness to deductive inferences, I will not be discussing the familiar problems with defining truth-conduciveness or solving the generality problem. There is enough merit to the reliabilist view, especially if taken merely as a theory of entitlement, for warrant in using these notions even in the absence of a solution to all of its problems.

Defining validity as truth-conduciveness will strike many, especially logicians, as obviously wrongheaded, and is bound to provoke. It does indeed depart radically from standard definitions of validity found in formal logic. Those definitions are relatively “clean”, and use mainly logico-mathematical concepts. The present proposal also is similarly controversial in taking defeasible rationality to be a central notion to logic. This notion was briefly introduced in the Introduction and will be sharpened in §8. Note that this has nothing to do “defeasible logic”, which does not concern deduction, but rather the “logic” of uncontroversially defeasible or non-monotonic inferences, like induction or inferences based on generics. In any case, the present proposal thus introduces several concepts that may seem alien into its very heart.

Much of the controversy may depend on terminology, however. If “logic” is defined as involving normative questions about belief and inference, then it plausibly intersects with epistemology, where of course truth-conduciveness and defeasible rationality play central roles. This conception of logic was vehemently attacked by Gilbert Harman, who argued that there are no interesting connections between facts about validity and normative claims. We will return to this issue in §8, but, for now, let me simply declare my commitment to taking logic to be concerned with said normative issues. Given this commitment, it should no longer seem quite as radical, and provocative, to propose that it should operate with such empirical, normative, messy notions as truth-conduciveness and defeasible rationality, since those notions are central to epistemology. Or, perhaps we should say, the proposal will then only be as controversial as reliabilism about entitlement.

It is natural to take truth-conduciveness and defeasible rationality to be a kind of twin concepts. Are they even co-extensional? No, for a very complex “inference rule” could be truth-conducive—and even strictly valid—even though it would not necessarily be rational to infer in accordance with it. For, as argued by Ian Rumfitt (2000: II), it will in such cases be irrational to make the inference without first having gone through the necessary steps. Similarly, it is not rational to accept a very complex tautology “just like that” (although one could come to believe it rationally, by going through appropriate steps in the right order).

However, if we restrict ourselves to *basic* inference rules, like perhaps those of natural deduction, we can indeed propose that such a rule is truth-conducive just in case it is defeasibly rational. Here, we must of course specify what is meant by "basic". One idea is to take it to mean *meaning-constitutive* in the sense of Eklund (2007b) (similar to the notions defended in Båve (2012), (2018), and (2020)). But the idea that some inference rules are basic and others "derivative" in some sense, is rather common and we can stay neutral on how exactly to identify them. One might further propose other connections between the defeasible rationality and truth-conduciveness, e.g., that an inference rule being truth-conducive *grounds* its being defeasibly rational, but I stay neutral on this.

With strict validity, things are different, since the co-extensiveness claim fails even with the restriction to basic rules. This is for the simple reason that certain basic, *non*-demonstrative inferences, like induction or inference to the best explanation, can be rational although they are never strictly valid. If we restrict ourselves to deductive inferences, then it may seem more plausible to say that strict validity coincides with rationality. But, given the paradoxes, this entails that we take either some classical rule or the truth rules to be irrational. This is implausible, and I adduce some further arguments against it in §9.

A further thing to note concerning truth-conduciveness and strict validity is that a rule is strictly valid does not contradict the claim that it is generically valid, but actually entails it (on the plausible assumption that universal quantifications entail their corresponding generics). The contradiction in this discussion instead arises at a higher level: the claim that generic validity should be taken as the central notion in logic contradicts the claim that strict validity should.

Although it is clear that truth-conduciveness "plays a central role in epistemology" in the sense that it has been intensively debated there, it is still controversial that it *should* be seen as crucial to epistemology. That view presupposes a broadly externalist/reliabilist account of entitlement or rationality. But this is a relatively weak form of reliabilism, since entitlement, again, is a relatively weak notion, compared to that of having reasons.

The main rationale is that the alternative, internalist view of entitlement would face a dilemma. One horn of this dilemma is an infinite regress with ensuing unreasonable demands, and the other is dogmatism, according to which one can be entitled to reason a certain way simply because one is disposed to, even if that pattern of reasoning is not truth-conducive (see, e.g., Bergmann (2006)). On the kind of externalist/reliabilist view presupposed here, there are *normatively basic* modes of belief-formation, such that one can be entitled to form beliefs on their basis even in the absence of further justification for so doing. However, these modes of belief-formation must be truth-conducive in order to confer entitlement.

2. Truth-conduciveness rather than generic validity

Hofweber's idea of defining validity in terms of generics is clearly in the same spirit as the idea pursued here, but seems to me ill-advised, for several reasons. First, it seems that a generic, "*F*s are *G*", can be true even when only a small minority of the *F*s are *G* (cf. 'Mosquitoes carry malaria'). Hofweber is aware of this, but takes this to show merely that not all generics license default reasoning (2021, n. 17). But if we are to retain the idea that the central notion in logic should be understood in terms of generics, and also that it is essentially tied to default reasoning, we must somehow restrict the claim to cases (perhaps "readings" of generics) where default reasoning is indeed licensed. However, we do not currently understand why some generics appear true even in "minority cases", while others don't, e.g., through some classification of different readings of generics. Therefore, we cannot currently describe the needed restriction. One may think this is a merely temporary lacuna, but that response misses a more important point: there is no reason in the first place to think a clearer understanding of generics will lead to a clearer view of which notion to take to be central to logic. This is not to say that natural language constructions are in general illegitimate for defining validity. Rather, we should avoid constructions that are at once (a) poorly understood, (b) in need of disambiguation, and (c) lack an obvious connection to the target notion (validity).

It may be objected against this that generics, while badly understood in linguistics, are well-understood in ordinary use. But it is not clear that they are. Even if we do not worry about them in ordinary conversations, they may still be ambiguous in ways we do not recognize. If that is so, they are unsuitable for use in analyses of philosophical concepts, until disambiguated. Perhaps they are not ambiguous, but the mere possibility seems to be a reason against using them until we, as "linguists", understand them better.

Leaving aside the problematic "minority-case generics", it is still dubious that validity should be identified with generic validity. Even if all generics were true only in "majority cases", this still seems too weak to capture validity. The probability of a conclusion of a valid argument with certain premises should be more than just over .5. That is, if the probability of the conclusion of an argument, given the premises, is only, say, .7, then the argument should not be considered valid. As it turns out, that most Fs are G is neither necessary, nor sufficient for the truth of *Fs are G*. The mosquito case shows that it is not necessary, and 'Books are paperbacks' and "Canadians are right-handed' show it is not sufficient, since these sentences seem false although a clear majority of the Fs are G. This makes the identification of validity with generic validity even more problematic, since it means that an inference rule can be generically valid even if not even most instances thereof are truth-preserving.

Considering the candidate exceptions to inference rules that emerge from the Liar, i.e., the "pathological cases", they certainly seem to be very special and far between. This fact is plausibly relevant to our sense that inferences instantiating the rules to which they are exceptions, are typically rational. The obvious upshot, then, is to take validity to entail that, roughly speaking, *almost all* of the instances of valid rules preserve truth. It is for these reasons that I propose that we think of the relevant property of inference rules as truth-conduciveness, understood as something having to do with the proportions or frequencies of rule-instances that are truth-preserving.

Let us first discuss a minor worry about this idea. It may seem that any rule taking us from any premises to a given truth will count as truth-conducive, since any instance will just have that same truth as conclusion. But such an inference will of course not always be rational. The solution to this, I believe, is to simply stipulate 'truth-conducive' is to mean "non-vacuously truth-conducive", by which I mean that the inference rule has many instances, that it is schematic (i.e., neither premises nor conclusion may be a specific claim), etc. Or we may accept this "rule" as truth-conducive, and simply make the required qualification to the claim that all truth-conducive inference rules are (defeasibly) rational. In either case, the problem can and should be solved by ruling out these "rules".

The more serious problem with taking rational deductive inferences to be those that instantiate truth-conducive rules is that inference rules have uncountably many instances, and because it makes no sense to speak of proportions of an uncountable set (as opposed to a merely infinite set). But the uncountability of instances cannot reasonably be taken to show that our sense that the pathological cases are rare and special is simply mistaken. What it shows is rather that their rarity cannot be captured in any simple and clean way, e.g., by speaking of frequencies of instances, period.

A reasonable first step toward a definition of truth-conduciveness applicable to deductive rules is to relativize to *finite subsets of instances*. This can be achieved in many different ways. A quick and dirty way would be to take the relevant finite set to be that of propositions that have actually been expressed by living creatures (or sentences uttered). A better, more principled way would be to define the set recursively. Thus, we could

(1) opt for a certain degree of complexity or length as an arbitrary upper bound,

(2) specify some finite stock of primitive expressions, each coupled with a syntactic type (most naturally thought of as "the" expressions of a given natural language), and

(3) identify a set of formation rules.

The formation rules will now, in conjunction with the set of typified primitive expressions and the upper bound, determine a finite set of instances of a given rule.

These choices will of course be arbitrary to some extent. But that is not necessarily a serious cost. Firstly, such arbitrariness is a familiar feature of operationalizations. Further, we need not merely consider a single set of arbitrary parameters. We can consider whether a given rule comes out as having a high frequency of truth-preserving instances relative to a wide variety of sets of reasonably chosen arbitrary parameters. Here, "reasonably" chosen ones are those that determine a relatively large set of unbiased instances. Hence, the set can usefully be assessed the same way as induction bases are. In this context, "unbiased" means that its members have no specific relevance for the inference rules being considered. The two proposals above are surely unbiased in this sense: the condition of having been uttered has no relevant relationship with the truth rules, and nor does the condition of being formable with the rules of English and satisfying a given upper bound of length. If a given rule comes out as truth-conducive relative to several such unbiased finite sets, that is a significant empirical fact. Given a reliabilist outlook, this fact will also be epistemologically significant.

Although the task of characterizing truth-conduciveness, as applicable to deductive rules, is bound to be complex and substantial, it seems to me clearly required for the idea of taking the central notion in logic to be something weaker than strict validity, given the drawbacks of generic validity. We may also note that the difficulties with arbitrariness involved here arise also for reliabilist theories more generally, and go under the label of "the generality problem". While those difficulties are real threats to reliabilism, they do not immediately undermine it. This is because, firstly, they might have a solution, and, secondly, because there are independent reasons for thinking that some form of reliabilism must be correct. Thus, I will henceforth go on to speak of the truth-conduciveness of deductive rules, keeping in mind said difficulties, but also the possibility of overcoming them along the lines sketched above.

3. Is Hofweber's claim uncontroversial?

Although Hofweber's claims will appear radical at first, there is reason to suspect that, terminology and emphasis apart, it is actually something most logicians will accept, at least upon reflection. For instance, the claim that both classical and unrestricted truth rules are truth-conducive should not be controversial. Adherents of traditional solutions to the Liar—who try to solve it by restricting the "naïve" principles governing truth—all agree that *non*-pathological instances thereof should come out as true or truth-preserving as far as possible. Similarly, adherents of paracomplete or paraconsistent solutions take the exceptions to classical logic to be rare and special, and take pains to "recapture" classical logic for non-pathological discourses like mathematics. They should therefore also agree that they are truth-conducive in roughly the sense hinted at above.

By the same token, they should also agree that instances of these rules are rational barring defeaters. After all, this merely amounts to the plausible claim that non-pathological instances of the truth-rule, like ‘Snow is white; therefore, <Snow is white> is true’, are rational absent a special reason to doubt it, and similarly for clearly non-pathological instances of classical rules.

The claim that generic validity or truth-conduciveness should replace strict validity as the central notion of logic is surely more controversial, but, again, appearances may be deceptive. Hofweber is open to the possibility of a “straight” solution to the paradoxes, i.e., a well-motivated theory identifying which principles are strictly valid and which aren’t. If there is such a solution, strict validity would play an important role in logic alongside the weaker form of validity. None of the notions would be *the* central one, and which one is *more* “central” now seems purely interest-relative and insubstantial. Theorists focusing on rationality and other epistemological matters might take the weaker notion to be more central, but there will also be the abstract, “Platonic” facts about strict validity that are of intrinsic interest.

If there is a straight solution to the paradoxes, then, Hofweber’s claims lose much of their interest. But I will argue in the next section that, in an important sense, there isn’t. Even if that argument fails, however, there is still a more cautious claim to be made: *if* it should turn out that no “straight” solution is available (e.g., because no theory operating with strict validity can be virtuous enough to merit justification), then there is an alternative, “retreat position” that is interesting in its own right, and which would help us avoid any overly pessimistic conclusion about deductive reasoning. That view would provide an explanation of how it can sometimes be rational, even if it falls short of the traditional ideal of knowable strict validity.

4. The prospects of standard solutions to the truth-paradoxes

The most common kinds of solution to the semantic paradoxes are those that:

(1) deny that the truth rules are strictly valid (the "traditional approach"),

(2) deny that certain classical inferences are strictly valid, and

(3) deny these rules apply to the paradoxical sentences (perhaps because they are meaningless, ill-formed, etc.).

I will henceforth assume, with most logicians currently working on the paradoxes, that the third type of solution is not viable, and I will call a solution of either of the remaining types a *standard* solution/theory/etc. By a "standard approach", then, I shall mean a theory that decides, for each deductive inference rule relevant to the truth-paradoxes, whether it is strictly valid or not, and thereby avoids trivialization (that every sentences comes out as true). In the following section, I will discuss in a similar vein the more recent, "substructural" solutions to the paradoxes.

Now, the idea that "there is no" solution to the paradoxes is naturally interpreted as entailing that there is no standard theory is *true*. However, their failings do not quite establish such an ontic claim. Or, rather, they do so at most indirectly. What the failings indicate more directly, I will argue, is something epistemic, namely, that *there is no standard theory that meets the desiderata on logic to an extent sufficient for justified belief in the theory*. Given that knowledge entails justification, it is then *unknowable* which standard solution is the right one. As I will go on to argue later on, this unknowability is best explained by its being *metaphysically indeterminate* which standard solution is correct. This way, more ontic matters resurface, but in clearer guise.

Above, I said that the present proposal relies to some extent on conceiving of logic as intersecting with epistemology. However, the failures of standard solutions that is the topic of this section also provides some motivation for Hofweber's revamping of logic, and this motivation comes from *within* standard formal logic. It is thus independent of the epistemological conception of logic.

To see what is indicated by the failings of standard approaches, we must first give at least a minimal survey of the results of relevance for my purposes. This survey actually covers what is rather generally thought of as the most important results. The traditional approach to the paradoxes is to restrict the truth rules, or give a theory of truth given which only a limited, consistent set of T-sentences come out as theorems. However, it is a widespread view today that any such restriction will be to some extent *ad hoc*, i.e., that there is no principled restriction of the truth rules (or T-schema). This is reinforced by Vann McGee's proof (1992) that there are infinitely many consistent, maximal sets of T-schema instances, differing only arbitrarily.

Further, many proposed solutions turn out to face various *revenge paradoxes*. A revenge paradox for a theory T is a paradox that arise as we consider a sentence saying of itself that it is *F*, where being *F* is the property that T assigns to the original Liar sentence, and where T takes the truth rules or classical rules to fail for sentences that are *F*. The relevant predicate '*F*' might be, 'is not true', 'does not express a true proposition', etc. (see, e.g., Beall (2007) on revenge phenomena).

The other, now increasingly popular standard approach is to keep the truth rules unrestricted and reject classical logic in favour of a paraconsistent or paracomplete logic. However, ordinary versions of such logics are not weak enough to avoid Curry's paradox, which requires in addition that one reject the principle of *contraction*: "$(A\to(A\to B))\to(A\to B)$". But this feature, too, seems *ad hoc*, counter-intuitive, and makes the logic unattractively weak. It has also been argued that these non-traditional approaches, too, inevitably face revenge paradoxes (Murzi and Rossi (2020)).

These failings motivate the conjecture that *no* standard account will be good enough to merit justified belief. Here, "good enough" is a matter of satisfying to a sufficient extent such desiderata as,

(1) agreement with intuitive judgments of correctness of inference/assertion,

(2) simplicity/manageability/avoidance of *ad hoc* assumptions,

(3) strength,

(4) non-triviality (i.e., not every sentence is a theorem).

Some of the items in this list may well be redundant, i.e., they can be subsumed by others. Some perhaps should not be considered desiderata at all. One might, for instance, question desideratum (1), and argue that it is rather strength (short of triviality) that should be maximized. The list may also be incomplete. But, for present purposes, these examples will do.

The argument for the unknowability of any standard solution now runs:

(P1) Our justification for believing a standard theory is wholly determined by that theory's satisfying such desiderata as (1)–(4),

(P2) No standard theory satisfies such desiderata to an extent sufficient for justified belief in it.

(C1) No standard theory can be justified.

(C2) No standard theory can be known.

The idea is that (P1) and (P2) jointly entail (C1), which, via the assumption that knowledge requires justification, entails (C2). (P2) is intended in the strong, abstract sense, in which it is true just in case there is no set of propositions, whether actually proposed or entertained, meeting the condition. The reason (P2) is not itself enough to justify (C1) is that even if a theory performs poorly with respect to desiderata like (1)–(4), one could suppose that there are other (kinds of) desiderata that could tip the balance in favour of the theory, making us justified in believing it. (P1) says that this is not so.

Both of (P1) and (P2) are somewhat vague, in that we left open exactly which the relevant desiderata are. This is an imprecision that we cannot to resolve here. Fortunately, there is wide actual agreement on approximately which desiderata are relevant, and that current solutions to the paradoxes do not score well relative to them.

The idea, then, is that (P2) is supported by extrapolation from known failings of actual solutions to the paradoxes. To get a clearer view of this line of reasoning, we need to look closer at the specific desiderata listed, as well as of the failings. Simplicity is plausibly a non-redundant and rather crucial desideratum for the argument. For although we have little grounds for claiming that there is no strong, non-trivial, standard theory giving many desired specific results, e.g., accommodating specific intuitive judgments, it is unlikely that any such theory will also be simple. Traditional solutions tend to score especially low on this criterion, as witnessed, e.g., by the Revision Theory of Truth, which is arguably the state of the art of such solutions (see Gupta and Belnap (1993)). This theory is exceedingly complex, and nevertheless leaves many questions unanswered, e.g., with respect to revenge paradoxes. (On the revision theory, pathological sentences are held to be not "stably" true, and face a revenge paradox consisting of a sentence saying of itself that it is not stably true.) More generally, the idea that we are not likely to find a simple solution that also fares well with respect to other desiderata is mainly based on the sheer amount of work that has gone into this enterprise. If there were a simple solution, we would likely have found it already.

The desideratum of strength is also crucial. For it is of course easy to give a simple and perfectly intuitive logical theory, e.g., by having it state merely that ‘A ⊢ A’ is valid. The difficult task is that of finding a simple and intuitive theory that is also reasonably strong. Finally, the desideratum of non-triviality or non-explosion simply means that not every sentence should be a theorem, and requires no defence. Suffice it to say that it can be seen as an updated version of the traditional theoretical virtue of consistency—updated, that is, so as to remain neutral on paraconsistent theories. Even if this desideratum may seem obvious, it is not uncontroversial, as it is flouted on the “trivialist” view that *everything is true* (see Bueno (2007) and Kabay (2008)). David Liggins’ “radical restrictionist” view that *nothing is true* (2019) does not exactly conflict with non-explosion, but is equally radical. These views are worth mentioning in this context, for those who find the present proposal too radical. The fact that the Liar paradox has elicited such extreme responses is testament to those difficulties of solving it that motivates (P2), the claim that no standard theory can meet the desiderata enough to be justified.

Now, the above argument for (P2) of course relies heavily on extrapolation, and is best seen as a conjecture based on observed failings of extant proposals. It is hardly something we should expect to find a proof of. To me, the conjecture seems likely, given the failures of past attempts, but I realize this impression will not be universally shared. Perhaps the predicament is similar to the case of non-shared intuitions. Maybe such intuitions provide some warrant for those who have it, but not for others. Or perhaps the fact that the intuition isn’t universally (or widely) shared simply defeats any warrant it could have conferred.

Many will in any case find this conjecture premature. But even if it is, the mere *possibility* that it is true raises important questions of how to respond if it turns out to be. We can rephrase the main point as a conditional: *if* it should turn out that there is no sufficiently virtuous theory that solves the semantic paradoxes, then there is an alternative conception of logic that stands as a palatable rescue position. My own view, to be defended below, is the stronger one that this alternative conception is not merely palatable, but one that performs well with respect to desiderata (1)–(4), and should therefore be embraced. There is a second point emanating from this further line of thought, which is that we can say more definitely that it is not the case that there *must* be a straight solution to the paradoxes. There might be, but since there is a satisfactory theory consistent with its denial, it is not necessarily so. In other words, the line advanced here is not a defeatist one. This will likely be denied, too, by adherent of strict validity. Now, however, we are approaching more tractable issues concerning the relative merits of different kinds of theories of logical validity, and need no longer be consigned to blind guessing about whether there is a straight solution or not.

This last kind of consideration also answers a different objection (due to XXX??redacted), to the effect that the reasoning involves a kind of "pessimistic meta-induction" that risks over-generalizing. For instance, if the above reasoning is rational, couldn't we also rationally infer that there is no solution to radical scepticism, since all attempts to avoid it have failed? The reply is that *there is* a satisfying solution to the semantic paradoxes, just not a standard one. The "solution" is to simply note that sometimes, the rules we use have applications that face defeaters. They are still good rules, since they are truth-conducive. That there are defeaters to their application is no less tolerable than the fact correct applications of certain non-demonstrative rules, based on true assumptions, may deliver a false conclusion. More generally, many things we would say about induction and the like—for instance, in response to Humean scepticism—are things we can and should say about deduction. Indeed, a good way of summing up the perspective on logic advanced in this paper is that deduction is much more like induction that previously thought. (This is not to say, however, that they are not also different in kind in some respects.)

The objection from over-generalization is really due to the widespread assumption that a solution to the semantic paradoxes *just is* what I am here calling a "standard solution". But it is precisely this implicit identification that I am questioning. There is a solution of sorts, but it is not one that, like standard solutions, identifies a "culprit", i.e., a rule that fails to be strictly valid. It is rather one that takes the paradox to be an unsurprising fact about how rules introduced for certain practical purposes may turn out, when taken in conjunction, to have defeaters to their application. More about these "practical purposes" and how we shouldn't be surprised by the semantic paradoxes in §6.

Here is a more specific objection against the argument above, namely, that it presupposes that we are not justified in believing the best theory. Here, I mean *the* best theory, not *our* best theory, which makes the presupposition all the more questionable. But note that I am not quite saying that we are not justified in believing the best theory about logic, period. I am saying that we are not justified in believing the best *standard* theory, i.e., the best theory operating with strict validity. There is therefore no controversial presupposition about theory choice in general in play here.

Another, related question, has to do with the possibility of *ties* between standard theories. It seems reasonable that if the two best standard theories satisfy the desiderata (almost) equally well, then we are not justified in accepting either. This argument is rather different from the one developed above. That argument depended on taking each theory to fail to reach the required degree of desideratum-satisfaction, whereas the present consideration instead relies on a condition for a theory's being justified that it score considerably better than any rival. Some will find this condition to have intuitive appeal; others will take it to be reminiscent of Buridan's ass. The latter will take both theories to be justified, if they score well enough. But this seems to presuppose a rather pragmatic view of epistemic justification that may seem misplaced in this context. This is not the place to resolve this general dispute. Suffice it to note that the condition on justification of being considerably better than any rival could offer a further route to (C2), if rather different in spirit than the one pursued above.

5. Substructural solutions

A rather different kind of solution to the paradoxes, defended by David Ripley (2012, 2013) and Pablo Cobreros *et al.* (2013), is the substructural one. This kind of solution is similar to Hofweber's in that it endorses the validity of both classical rules and truth rules. But instead of redefining validity, they reject the standard set of structural rules governing the single turnstile. In particular, they reject the *Cut* rule:

$$(Cut) \quad \frac{\Delta \vdash A, \Gamma \qquad \Delta', A \vdash \Gamma'}{\Delta, \Delta' \vdash \Gamma, \Gamma'}$$

This rule can be seen as a generalization of *transitivity*, saying that if A ⊢ B and B ⊢ C, then A ⊢ C.

Substructural solutions stay silent on the (presumably vast majority of) instances of the *Cut* rule that preserve truth. That rule must therefore somehow be "recaptured" in order to legitimize ordinary mathematical (and other) practices, which rely heavily on *Cut*. Hofweber (2021: 269f.) argues that will be difficult, as it must rely on an assumption to the effect that one is working within a "consistent context". But it is hard to come by a *proof* that, e.g., mathematics is paradox-free. To this, I would add that it seems difficult to see what principled way there is for distinguishing different "contexts". Should mathematics as a whole count as a "context" or should, say, arithmetic be separated from different levels of higher mathematics?

Thirdly, Hofweber argues, even if these difficulties were to be solved, and if, in the context where we find ourselves, we have grounds for claiming it to be consistent, we wouldn't quite be reasoning with *Cut* anymore. Rather, we would be operating with an inference rule including, as one of its premises, something like the sentence, 'I am now reasoning within a consistent context' (Hofweber (2021: 269)). This is very different from *Cut*, and it is a rule very different in kind from any other rule in Logic, whether structural or ordinary. (The same kind of argument can of course be levelled at the more familiar non-classical solutions that aspire to "recapture" classical logic for certain discourses.)

I think an improved understanding of what the relevant rules should be like might alleviate these worries. For instance, it is reasonable that paraconsistent logicians do not literally mean to use classical rules when doing mathematics, but rather conditionalized versions of these rules, where the condition of their application is simply that the propositions involved are consistent (here meaning, "not both true and false"). The conditionalized version of *modus ponens* would then be,

$$\frac{<p> \text{ is consistent; } <q> \text{ is consistent; } p\text{; if } p \text{ then } q}{q}$$

Paraconsistencists could thus be content saying that the recapture in question is a matter of using these conditionalized rules. Further, it may be acceptable to use these rules in mathematics even in the absence of a *proof* of the consistency of mathematical propositions, as long as we are deafeasibly justified in taking them to be consistent. There is now also no special problem with distinguishing and individuating "contexts", since the rules only require that we distinguish propositions.

I therefore suspect that Hofweber's case against the substructuralists (and others aspiring to recapture classical logic for particular areas) can be neutralized. Yet there is still a worry about simplicity, since non-structural rules are assigned one type of validity, and certain structural rules, like *Cut*, are assigned another, weaker form of validity. The alternative proposed by Hofweber and myself is of course is to use only the weaker notion, and attribute it to all rules that are defeasibly rational.

The claim that an inference rule is truth-conducive does not entail that it is not strictly valid (just as with generic validity). Therefore, the claim we have been defending, i.e., that both the classical rules and the truth rules are truth-conducive, is consistent with the claim that the classical rules are strictly valid. Similarly, it is consistent with this claim to take *Cut* as strictly valid. The point is just that, given the difficulties of telling which rule(s) should be excluded from the privileged status of being strictly valid, and since we will likely obtain an overall better theory by simply taking the central property of inference rules to be truth-conduciveness, and accord this status to *all* the relevant rules, we do not need to try to establish the stronger claims, even if they are consistent with what we have been claiming.

6. The metaphysical indeterminacy of which standard solution is right

We have seen that Hofweber's view would be of greater interest if it turns out, in some sense, that there is no standard solution to the paradoxes, i.e., a logical theory on which validity is strict, which identifies either classical logic, the truth rules, or some structural rule, to be invalid. I have argued that no standard theory is knowable. Still, is any of them *true*? Firstly, this matter will of course be of less interest than ordinarily thought, given that it is unknowable. But I also think the unknowability claim provides grounds for a more direct answer to this question of truth. To wit, I think it supports the idea that it is *metaphysically indeterminate* which standard theory is true. I base this on the following principles and observations. Firstly,

(P) If it is unknowable (in principle) whether *p*, and this is not due to semantic indeterminacy in '*p*', spatiotemporal inaccessibility, or Fitchian properties of '*p*', then it is metaphysically indeterminate whether *p*.

Secondly, it seems that the unknowability of which standard theory is true, i.e., of which of the relevant inference rules are strictly valid and which are not, does not derive from semantic indeterminacy. That is, we could not make this matter knowable by precisifying vague or otherwise semantically indeterminate terms involved (assuming there are any). And, even more plausibly, the matter is not unknowable due to spatiotemporal inaccessibility (we will get clearer on thus notion below). Finally, I am assuming that no standard theory is "Fitchian", i.e., no such theory is formulated in such a way that the assumption that it is known leads to contradiction. In "Fitch's paradox", we consider a sentence of the form, '*p* but no one knows that *p*'. By simple reasoning, we can see that no such sentence can be known to be true. Hence, a proposition expressed by such a sentence is unknowable (sometimes called "structurally unknowable").

The following might seem like a counter-example to (P): it is unknowable exactly how many molecules there are in my coffee, yet this is not due to any of the conditions mentioned in (P). It cannot, for instance, be due to "spatiotemporal inaccessibility", since it is "here and now" (thanks to XXredacted?? for this objection). This is a fair point, but by looking closer at the possible understandings of the alleged counter-example, it will emerge that (P) withstands the argument, and the notion of spatiotemporal accessibility is clarified along the way.

As stated, the alleged counter-example fails since the matter is unknowable due to semantic indeterminacy. This is because 'my coffee' will have several sums of molecules as admissible precisifications. So let us consider a case where an exact region of space has been defined, along with a precise definition of "being in" a region. We also need to be precise about the time of counting, since many regions will constantly change "inhabitants".

Having fixed all of these parameters, it may seem that we have an unknowable fact not due to any of the conditions mentioned in (P). But I will argue, on the contrary, that it is now clearer that we can write the case off as due to spatiotemporal inaccessibility. Whether a given fact is spatiotemporally accessible or not depends, we might say, on its impact. A huge event, like the Big Bang, will leave traces that survive for billions of years, allowing us now to know that it happened. By contrast, it is unknowable whether I woke up on my back on 22 April 2022, because I can't remember and it leaves no other traces we can detect. Even a fact that is spatiotemporally very close can be spatiotemporally inaccessible. Considering the precisified question about the exact number of molecules in a region, we can see that the difference between its containing n molecules rather than $n + 1$ will not leave any trace allowing us to determine which it is. If we could take a snapshot of the region, to count the molecules pictured there in peace and quiet, then the question is no longer unknowable. Lacking such a snapshot, however, it is unknowable, but now it should emerge more clearly that it is unknowable due to spatiotemporal inaccessibility, a situation only different in degree from my waking position on a certain date. The fact that it is *roughly* "here and now" is not enough when such details as the exact number of molecules is at stake, just as the details about someone's waking position can be rendered unknowable very soon. This shows that the counter-example fails.

Now, given these observations and clarifications, (P), and the unknowability claim, it follows that:

(MI) For every standard theory, it is metaphysically indeterminate whether it is true.

A "standard theory", to repeat, is one saying, for every deductive inference rule and structural rule relevant to the truth-paradoxes, whether it is strictly valid or not. Thus, it follows from (MI) that:

(MI′) It is not the case that for each deductive inference rule relevant to the truth-paradoxes, it is metaphysically determinate whether it is strictly valid or not.

The most questionable assumption of this argument is probably (P). (P) is equivalent to the claim that unknowable propositions are either semantically indeterminate, structurally unknowable, metaphysically indeterminate, or concern spatiotemporally inaccessible matters. As an argument from exclusion, this one is far from conclusive. Such arguments are often dismissed as relying on lack of knowledge or imagination. Like (P2), then, (P) remains a conjecture. But I believe many will agree that it is hard to think of any further alternative ground for unknowability. Further, if a proposition could be unknowable for some other reason than the four mentioned in (P), then this will still only block (MI) if standard theories have this further property, explaining the unknowability of which one is correct. Otherwise, an argument featuring a premise similar to (P) but adding this further property can be given for (MI). Otherwise put, if what makes it unknowable which standard theory is true could not be that the relevant theories have this further property, because they do not *have* this property, then we could still infer that the matter is unknowable because it is metaphysically indeterminate.

Note that in order for the kind of revamping of logic recommended by Hofweber is to be of interest, we only need (C2), not (MI). If no standard theory is knowable, then the alternative conception of validity is motivated, whether (MI) is true or not. Nevertheless, for those who find it hard to understand how it could be unknowable which standard theory is true, (MI) provides an answer and may thereby make the unknowability claim more plausible.

Of course, the notion of metaphysical indeterminacy remains moot, and there is much discussion about how to define it. The most recent literature on the topic deals especially with the question of how to define it compatibly with classical logic (see Greenough (2008) and Williams (2008) for overviews). Of course, we are here only committed to "classical logic" in the sense that we take the classical inference rules to be truth-conducive and defeasibly rational, which is of course compatible with taking cases of metaphysical indeterminacy to be exceptions to classical principles (in particular, LEM). However, it is preferable to take such exceptions to be as few as possible, so I, too, will assume that these cases are no exception to classical logic.

A central tenet of these new accounts of metaphysical indeterminacy is that LEM is determinately true. This commits us to denying that determinacy distributes over disjunction: $\mathsf{D}(p \text{ or } q) \nvDash \mathsf{D}p \text{ or } \mathsf{D}q$. Otherwise, no proposition would be indeterminate. We can now say that it is determinate that either the traditional theory of the paradoxes is true, or the paraconsistent theory is true or the paracomplete theory is true, or … without taking any of these theories to be determinately true. By the same token, we can hold that while it is determinate that some standard theory is true, there is no standard theory that is determinately true. This is to deny that metaphysical determinacy commutes with the existential quantifier: $\mathsf{D}\exists x\varphi \not\models \exists x\,\mathsf{D}\varphi$. The latter denial, then, allows us to say that although some standard theory is true, it is indeterminate which on is. This view may seem congenial to the above reasoning about standard theories' satisfaction of desiderata. For the difficulties of finding a standard solution to the paradoxes may be thought to indicate merely that no such theory can be justified, not that none of them are *true*. The claim that some such theory is true, but it is indeterminate which, may therefore seem like the properly balanced position. This is an optional claim, however. We might also just deny that any standard theory is true, or remain silent on the issue.

The failure of distribution over disjunction and of commuting with existential quantification are common features of other necessity-like operators, including other kinds of determinacy. It is reasonable to think metaphysical determinacy obeys S4 or S5, or, if there is higher-order metaphysical indeterminacy, KT or KTB. (See Williams (2008: §3) for an overview and Barnes and Williams (2011: III) for a more detailed discussion of the logic of metaphysical indeterminacy.)

The main issue discussed in the recent literature, however, concerns the semantics of the metaphysical determinacy operator, which directly involves questions about its metaphysics. In particular, it has been proposed that metaphysical determinacy can be analysed in terms of truth-making. I have nothing to add to this discussion, and will instead venture to answer a rather different question: how did we end up being able to ask questions about logic and truth that simply have no answer? Should we be surprised? I believe an answer to these questions is available, which, however, involves some potentially controversial assumptions about the primary function of logical constants. Logical constants are plausibly introduced via inference rules or categorical principles that define them. This includes the concept of truth, which is plausibly introduced via some equivalence schema or the truth rules. It seems plausible also that we began to use basic logical constants because they served certain *practical, non-representational* functions. (For a recent discussion about non-representational "linguistic functions", see Thomasson (2020, forthcoming 1, 2). See also Price (2011, 2013). For some classical examples of proposed non-representational functions of specific expressions, see Ramsey on universal quantification (1929, 137), Sellars on counterfactuals (1958), Hare on normative terms (1952), and Quine on the truth-predicate (1970: 11). For an extension of Quine's idea (commonly associated with "deflationism about truth") to property-designators, see Båve (2015).)

The idea is that logical constants and the truth-predicate serve purely inferential or expressive properties, consisting in allowing the formation of sentences that enable us to draw conclusions, given new information, that ultimately provides new knowledge about our surroundings, or new directives for specific actions. For instance, one might speculate that the function (or *a* function) of disjunction is to enable the formation of sentences allowing us to infer a categorical sentence '*p*' upon the rejection of another sentence (as per disjunctive syllogism). Similarly, the function of the truth-predicate could be that it allows the formation of sentences with special inferential properties, like 'Everything he said is true', which allows one to infer, for every assertible sentence of the form, 'He said that *p*', the corresponding sentence '*p*'. The logical constants are thus not primarily useful due to such representational properties as "referring to a certain truth function" just as the truth-predicate is not primarily useful in that it allows us to "refer to the property of truth" (although we need not deny that they do refer to those entities). Now, once introduced for these various cognitive/practical purposes, there is no reason to believe that these expressions will end up mapping an independent, determinate reality in such a way that, for any question that can be asked using them, or about them, there is always a determinate answer.

This rather vague talk about mapping reality can be made clearer by speaking of semantic values. Matti Eklund (2002) proposes that the semantic values of logical constants and the truth predicate are determined by way of best fit (following Merricks and Lewis). The "fit" between an expression and a candidate value is determined by several factors, including: (1) the semantic values satisfying sentences containing the relevant expressions that we hold true, (2) the semantic values having properties attributed to them by our best theory, and (3) the values' intrinsic naturalness. The most relevant aspect here is (2). The idea, as applied to the case of the paradoxes, is that if there is a best-by-far theory solving the semantic paradoxes, which, out of several candidate truth properties, takes truth to be T, then the semantic value of 'true' is T. Similarly for logical constants: if the best overall theory (which as such includes an account of the paradoxes) incorporates a certain non-classical logic, then our constants, 'not', 'if', etc., pick out functions that are the semantic values of the constants on the (best) semantics of this logic. (This talk of semantic values of logical constants does not contradict what I said earlier about functions: that a given entity does not have the property P as a function does not entail that it does not have P.)

Eklund is neutral on which theory about the paradoxes is better, and, hence, neutral on what exactly are the semantic values of 'true' and the logical constants. His main aim is to defend his view that the paradoxes show our language to be inconsistent, by containing expressions governed by jointly inconsistent meaning-constitutive principles. Part of the defence involves an account of how the inconsistency is compatible with a principled assignment of semantic values to the relevant expressions. My point is slightly different: our conceptual, linguistic, and inferential activities have developed mainly as a response to practical needs. These activities involve the acceptance of certain sentences and inferences in accordance with certain rules. There is no reason to think that this phenomenon has developed in such a way that a neat and simple assignment of semantic values, following the principles (1)–(3), will be available. In fact, the difficulties with solving the paradoxes in standard ways seem to show that there is no very satisfactory assignment of semantic values to the relevant expressions. (In (2007a: n. 39), Eklund himself actually suggests that it may be indeterminate which are the semantic values of ‘true’ and the logical constants.) If there is no determinately true assignment of semantic values of the relevant expressions, there is no determinately true standard theory, and conversely. Thus, (MI) follows from the claim that there is no determinately true assignment of semantic values. The latter claim is in turn not surprising, given that a plausible story about the function of the logical constants and the truth predicate. Returning to an earlier point, this picture also further motivates the kind of revamping of logic advocated by Hofweber, which, as argued above, is naturally coupled with the view that “there is no solution to the semantic paradoxes”.

7. A better logic?

Let us define “a logic” as a theory that says which of the inference rules involving the usual logical constants and the truth-predicate are valid and which are not. If validity is identified with strict validity, then no logic meets the desiderata on logic well enough for us to be justified in believing it, or so I conjectured. If, on the other hand, we take validity to be truth-conduciveness, then we can devise a logic that meets these desiderata very well. The most important virtue is that it takes both the classical inference rules and the unrestricted truth rules to be valid. Let us call this logic “D-logic”.

It may be thought, however, that D-logic is simply not comparable to standard logics, because they concern different things. But while they operate with different notions of validity, both arguably aspire to capture the properties of inferences that are important for epistemic purposes. Many have probably reacted to these ideas by thinking, D-logic isn’t really a logic, because logic is about validity, and validity is strict validity. But this is purely terminological. It is usually preferable to use a terminology that is theoretically neutral, at least on controversial issues. On a neutral terminology, “validity” is defined as something like “the truth-involving property of good deductive inferences”. (This definition assumes that there is only one such property, which is of course controversial, but let us for the time being consider only the dispute between defenders of strict validity and defenders of validity as truth-conduciveness, and so ignore logical pluralism.) Now, the substantive question becomes whether the important truth-related property of good inferences is strict validity or truth-conduciveness. And the fact that D-logic validates both classical logic and unrestricted truth rules obviously counts in favour of taking it to be the latter.

The reason this is an advantage has to do with the desiderata on logic, especially strength and simplicity. Things are complicated, however, by the fact that D-logic are in some ways incomparable with standard logics. Let me elaborate, beginning with strength. On the one hand, the claim that an inference rule is strictly valid is stronger than the claim that it is truth-conducive. It is therefore not the case that one is stronger than the other in the usual sense. But we can speak of different kinds of strengths: D-logic is stronger in validating more inference rules, while the other logics are stronger in using a stronger notion of validity. Thus, there is a respect in which D-logic fares better with respect to strength than standard theories.

As for simplicity, the advantages of D-logic are more obvious. Firstly, the truth rules and classical inference rules are all relatively simple, and a logic validating both is always simpler, in that respect, than traditional or non-classical solutions to the semantic paradoxes. This holds especially for traditional solutions, which tend to get very complex and still face revenge paradoxes. Secondly, non-classical logicians take pains to recapture classical logic for use in mathematics. With D-logic, this is not needed. This is another respect in which D-logic is simpler than its rivals: it allows for a simpler overall theory.

Objection: even if it is true that recapture in the ordinary sense is not needed for D-logic, the latter surely needs something analogous, since we know that mathematics is safe from paradox while truth-talk isn’t. Reply: granting that mathematics is indeed consistent, the analogous claim we must make as D-logicians is merely that in truth-talk, there are defeaters and exceptions to classical logic and/or truth rules, whereas in mathematics, there aren't. This is merely to restate what is already known, it does not amount to further complexity in our overall theory.

There may seem to be an obvious and serious drawback in D-logic with respect to simplicity, in that truth-conduciveness is more complex than strict validity. However, it is not clear that this particular type of added complexity really speaks against D-logic. This is because the more complex notion might be needed anyway, in epistemology. Consider the following principle:

(Simp.) For any two local theories T1 and T2, if T1 is simpler than T2 and they are alike in all other respects, then T1 is preferable.

Here, a theory's being *local* merely means that it does not include all of one's beliefs, i.e., it is not one's "total theory". As intimated above, (Simp) can be questioned in the following way: local theories, like logics in our sense, cannot be assessed without considering the overall theories they may be part of. It may be that T1 is simpler than T2 because T2 involves a complex notion not found in T1, but where this notion is needed *anyway*, in one's total theory. So if truth-conduciveness is a notion that is needed anyway, e.g., in epistemology, then, arguably, although D-logic is more complex than its rivals, this is not a cost. Alternatively, one may argue, more cautiously, that although (Simp) is true, it is *less* of a cost for a local theory that it uses a complex notion if that notion is needed elsewhere. So, since truth-conduciveness is needed in epistemology, it is less of a cost for D-logic to use it, even if it is still a cost by comparison to its rivals.

There is also an entirely different consideration, which might be a stronger defence of D-logic against this objection. It is no easy question how to partition one's total theory into local ones in a non-arbitrary way, and I cannot begin to answer it here. But we can note that the area in which truth-conduciveness is needed, i.e., epistemology, can reasonably be thought of as containing logic as a part (as argued in §1). This is reasonable since both have to do with norms of belief, and since both, arguably, have to do with truth-related characteristics of belief-forming processes, like truth-preservation. On such a partitioning of local theories, D-logic faces no cost even if (Simp) is true.

Here is a possible desideratum that was not listed above, but which may seem reasonable to consider in relation to logic: *safety*. Let us take the safety of a logic L to consist in the likelihood of coming to believe something false, assuming the truth of the premises, if one reasons in accordance with L. Reasoning in accordance with D-logic simply means drawing only inferences that D-logic takes to be valid (i.e., only with highly truth-conducive ones). Now, there is of course a non-zero probability of inferring in accordance with D-logic and yet go from true premises to a false conclusion. But it is not clear that standard theories fare any better with safety. A *true* standard theory is of course perfectly safe in the trivial sense that, given that it is true, we cannot reason in accordance with its rules and go from true premises to an untrue conclusion. But it is not clear that safety should not be measured this way. Perhaps the likelihood in question should not be taken as relative to the assumption that the logic is true, but to depend rather on the likelihood of the logic itself.

This is not something that can be handled mechanically by an application of a probability calculus. But there is an obvious sense in which it is not certain which standard theory is the right one, whether it is a classical, traditional one, or a non-classical one that retains unrestricted truth rules. Hence, no such theory has a probability of 1. And this uncertainty, one might think, must have some effect on the safety of a logic, on any reasonable interpretation of "safety". Otherwise put, the degree of certainty that one will not go from true premisses to a false conclusion following a given standard theory must depend on one's certainty that one's standard theory is true. Given the paradoxes and other problems making it uncertain which standard logic is the true one, the probability that we go from true premisses to false conclusions cannot be zero for standard approaches either (cf. Hofweber (2007: 5.4.3)). In conclusion: if safety in some probabilistic sense is a desideratum on logics, then, it is not obvious how D-logic fares with respect to it compared with standard theories.

I have argued in this section that D-logic fares well with respect to the desiderata on logic. So although the finding (if it is one) that no standard theory is knowable might at first have appeared as a defeat, we see now that it motivates a different kind of logic that in fact may perform better than standard ones.

8. Defeasibility and the normativity of logic

I have been assuming that logic essentially concerns normative questions about how to infer. This idea was famously criticized by Gilbert Harman, who denied that there are any interesting connections between validity claims and normative claims. For instance, he said, even if the argument from A to B is valid, it does not follow that if you accept A, you should accept B, since it may rather be that once you realize that B follows from A, you should reject A, due to an overriding reason to reject B. It is now commonplace to take Harman's rejection of the normativity of logic to have been premature, since there are many further possible normative claims about valid inference rules than those considered by him. John MacFarlane (2004) is a useful paper surveying some of these possibilities, although it does not mention what I take to be the central normative property of deductive inferences, *defeasible rationality*.

I assume, then, that there is an important question of which is the central normative property of deductive inferences. I this to be the same as the question of how to define the single turnstile, '⊢'. I also assume that '⊢' should ideally be co-extensive with the double turnstile, '⊨', which expresses validity, and which I, of course, take to be truth-conduciveness.

Let us now consider some rival candidates for the "central normative property". Restall (2005) proposes that the claim that an inference from A to B is valid entails that one ought not both accept A and reject B. While this plausibly holds for all good inferences, it is common to think that it cannot be the whole truth about the matter, since there should also be *positive* oughts pertaining to belief or inference: logic not merely forbids beliefs or inferences, but also obligates them. (The normative property considered by Harman, by contrast, does satisfy this desideratum.)

Turning to the semantic paradoxes, we can see that neither Restall's or Harman's property can be had by both the classical rules and the unrestricted truth rules. I will be assuming that in these paradoxes, we rely on premisses, namely, the proposition that (L) = '(L) is true' (in the Liar case) or the proposition that (C) = 'If (C) is true, then *p*' (in Curry's paradox). Now, we clearly cannot say that all the classical rules and the truth rules are such that if one believes the premises, one ought to believe the conclusion, since, given the Liar, it then follows that we ought to believe a contradiction. Given Curry's paradox, it also follows that we ought to believe any proposition whatsoever. A similar fate befalls Greg Restall's proposal. Since we ought to accept the assumption of Cury's paradox, it follows, if we think the relevant rules satisfy Restall's condition, that we may not reject *p*, for any *p*. To give one further example, Hartry Field (2009) holds that if an inference is valid, then if one is justified in believing the premises to degree 1, one is justified in believing the conclusion to degree 1. Since we are presumably justified in believing to degree 1 the assumptions of Liar or Curry derivations, we arrive at similarly unacceptable conclusions.

Let us instead turn to defeasible rationality. I defined this notion above, but we need to refine it somewhat. Firstly, in order to compare it with the above normative claims about belief, it will be useful to redefine it as pertaining explicitly to belief. Accordingly, we could take a basic inference rule as defeasibly rational just in case: if one believes its premises and has no epistemic (as opposed, e.g., to a practical or moral) reason against inferring the conclusion, then one ought to believe the conclusion. (Note that there may be epistemic reasons for not inferring a proposition that is nevertheless not a reason for rejecting it. Such reasons are called "undercutting", as opposed to "undermining"—the distinction goes back at least to Pollock (1986) and is a cornerstone of contemporary epistemology.) A problem with this definition, however, is that the claim that all valid basic inference rules are defeasibly rational in this sense entails that we ought to be logically omniscient (this problem also troubles Harman's notion). It is for this kind of reason that many opt for weaker claims involving epistemic entitlement, rather than outright oughts. For it is at least more plausible to say that we are entitled to believe any proposition that follows logically from what we believe than that we ought to do so.

This problem is best avoided, I believe, by conditionalizing on the thinker's *considering* or *entertaining* some or all of the relevant propositions. As argued by John Broome (2016), this qualification is also natural to add when defining the notion of an inferential disposition: we are not disposed simply to believe the (obvious) conclusions of what we believe, only to believe those that we *consider* (at least if we also consider the beliefs entailing the conclusion). With this qualification, we can also make an "ought claim", rather than merely one about entitlement. Thus, I propose the following definition:

> A rule R is defeasibly rational $=_{df}$ if someone believes the premises of an instance of R, considers the conclusion (and perhaps also the premises), and has no epistemic reason not to infer the conclusion, then they ought to believe the conclusion.

With this definition, the claim that valid rules are defeasibly rational at once:

(1) avoids Harman's "*modus tollens*" objection (thanks to the "no defeater" clause),
(2) avoids the omniscience problem, since it only demands that we believe conclusions we consider, and
(3) issues in a positive ought.

Given these advantages, it is reasonable to take defeasible rationality to be the central normative concept in logic. Together with the ideal mentioned above, that the two turnstiles should be co-extensional, and given our earlier observation that a basic inference rule is defeasibly rational just in case it is truth-conducive, we also have a new argument for taking validity to be truth-conduciveness. This last result shows that the overall account on which truth-conduciveness and defeasible rationality play the central roles of the two turnstiles displays an attractive unity.

9. Strictly invalid but rational

I believe most readers agree with me that, although the naïve truth-rules and classical rules cannot all be strictly valid, they are nevertheless all defeasibly rational. Again, this does not amount to more than saying that defeater-free applications are rational, like the inference from 'Snow is white' to 'That snow is white is true', and similarly for non-pathological instances of classical rules. Hence, there are rational deductive inferences that instantiate strictly invalid rules. For anyone in doubt, however, I would like to adduce the following two arguments:

(a) It is counterintuitive and violates reasonable principles of charity to take truth-conducive basic inferences that are standard among ordinary reasoners to be irrational.

(b) Since normal perceptual inference is rational, and since normal deductive reasoning in the face of paradox are in relevant respects like normal perceptual inference, the former is rational, too.

A quick note before we proceed: the claim to be supported could be expressed by saying that deduction is fallible. However, this claim is often taken as evidenced by the fact we can infer rationally, yet go from true premises to a false conclusion due to *misremembering* earlier steps of our reasoning. This is not what I mean here. I mean rather that a single, simple inference can be rational despite instantiating an non-strictly-valid rule. Memory is irrelevant to this matter.

Here is a way of spelling out the first argument:

(P1) Ordinary reasoners standardly infer in accordance with the truth-rules and classical rules, barring defeaters.

(P2) For all logical concepts C, we ought to say that ordinary reasoners' standard, truth-conducive deductive inferences involving C are rational.

(C) The truth-rules and classical rules are defeasibly rational.

(P1) should be read as meaning, not merely that speakers reason from, say, ‘T(<*p*>)’ to ‘*p*’ *given further premises* that occur crucially in the inference. The additional premise might be, e.g., that ‘*p*’ is non-pathological, non-self-referential, ungrounded, etc. Doing so would not count as inferring in accordance with the unrestricted truth rules. Similarly, they do not reason classically in the sense that they infer along classical rules together with further premises, e.g., to the effect that they are reasoning within a consistent context, etc. This is simply not how people reason. Rather, they go from the *mere* premises of classical inferences to their conclusion, and similarly for the truth-rules. Should we say that ordinary reasoners are therefore irrational, even if the forms of inference are highly truth-conducive? That seems implausible, as (P2) says. (P2) may in turn be motivated by some principle of charity, e.g., one saying that we should make people come out rational as far as possible. Note that even if people often commit such fallacies as “affirming the consequent” (a dubious claim), this has no bearing on the present argument, which concern only truth-conducive inferences.

The second argument appeals to certain parallels between deductive inference and other cases of defeasible justification. These parallels help illustrate the general proposal defended in this paper and also, I think, provide some weak inductive support for our present thesis. Take *perception*. Here, it seems, *normal* inferences (in the broad sense of a belief-forming process) tend to coincide with *rational* inferences. In the case of perception, also, the normal and rational inferences are both *defeasible and simple*. That is, people normally infer *by default* and from *simple* premises. So, when it looks as if *p*, one normally infers that *p*, barring any special reason not to. No further input-beliefs are typically required for the belief to be formed (hence, “simple”). That is, people do not normally use background beliefs about the reliability of perception, etc., when forming perceptual beliefs. On the normative side, similarly, when it looks as if *p*, it is *rational* to infer that *p*, again barring defeaters: no further input-beliefs are necessary for rationality (I am assuming a weak form of externalism here).

Now consider the normal reactions to the paradoxes:

(a) We do not accept the paradoxes' conclusions (contradiction, arbitrary claim, etc.)

(b) We use the concepts the *same way* as before encountering the paradoxes, i.e., inferring in accordance with the classical and naïve truth rules, barring special reasons not to. (This is unsurprising, since we do not know which inference rule to blame.)

(c) We do not, either before or after encountering the paradoxes, require more than the belief, e.g., that <*p*> is true, to infer that *p*, or more than the beliefs that *p* or *q* and that not *p*, to infer that *q*, and so on. *A fortiori*, we do not normally infer only if there is an additional belief to the effect that the particular instance in question is "safe". (This is unsurprising since only very few of us even think about "inference rules", "instances", etc.)

Note also that all of (a)–(c) parallel normal perceptual reasoning. To wit, (a′) we do not accept an absurd belief that *p* even if it looks like *p*; (b′) we go on relying on perception, even knowing we have once misperceived, (c′) even knowing we have once misperceived, we do not require additional premises about the reliability of our perception, to infer that *p* when it looks like *p*. All of this also seems rational.

These considerations provide a kind of inductive argument for the claim that the rules involved in the paradoxes are defeasibly rational. Roughly, the argument is that since normal perceptual inference works like normal deductive inference (in the face of defeaters), the facts about the rationality of deductive inference should match those of perceptual inference. (Or, at least, someone claiming that they come apart must explain why that is so despite the similarity of the respective normal inferences.) Thus, our standard deductive inferences should be seen as defeasibly rational, even though they cannot all be strictly valid. Surely, the two types of inference differ markedly in degree, in that deduction is in general safer than perception. But this of course only makes the argument stronger. If perceptual inference is rational due to its truth-conduciveness, then inferences that instantiate rules that are markedly *more* truth-conducive should be rational, too.

Clearly, the premises of these arguments could be questioned, but I will not here delve into any further discussion about them. I merely wanted to show what kind of commitments befalls anyone who wants to deny that strictly invalid rules can be defeasibly rational. Needless to say, more work is also needed to justify other claims I have made in this paper, particularly the ones I have called "conjectures". Several important and interesting questions also remain about exactly what is rational to believe about the paradoxical derivations, once we have noted their unacceptable conclusions. Why, for instance, should we stop short of accepting the Liar contradiction, rather than taking the Liar paradox to show that there is an exception to the law of non-contradiction? But those are questions of lesser immediate urgency than those I have addressed in this paper.

Acknowledgment

I have benefitted immensely from detailed discussions with ???redactedXXX on this paper.

References

Barnes, E., & Williams, J. R. G. (2011), "A theory of metaphysical indeterminacy", *Oxford Studies in Metaphysics* 6, pp. 103–148.

Båve, A. (2020) "Semantic Dispositionalism without Exceptions", *Philosophical Studies* 17, pp. 1751–1771.

Båve, A. (2018) "Trivial Languages", *Acta Analytica* 33, pp. 1–17.

Båve, A. (2015), "A Deflationist Error Theory of Properties", *Dialectica* 69, pp. 23–59.

Båve, A. (2012) "On Using Inconsistent Expressions", *Erkenntnis* 77, pp. 133–148.

Beall, J.C. (2007) (ed.), *Revenge of the Liar*, Oxford, Oxford University Press.

Bergmann, M. (2006), *Justification without Awareness: A Defense of Epistemic Externalism*, Oxford, Clarendon Press.

Broome, J. (2016), "Synchronic requirements and diachronic permissions", *Canadian Journal of Philosophy* 45, pp. 630–646.

Bueno, O. (2007), "Troubles with Trivialism", *Inquiry: An Interdisciplinary Journal of Philosophy* 50, pp. 655–667.

Cobreros, P., P. Égré, D. Ripley, and R. van Rooij (2013), "Reaching transparent truth", *Mind* 122, pp. 841–66.

Cotnoir, A.J. (2018), "Logical Nihilism", in *Pluralisms in Truth and Logic*, J. Wyatt, N. Pedersen, N. Kellen (eds.), Palgrave Innovations in Philosophy, Cham, pp. 301–329.

Daglish, S. (2020), *Accepting Defeat: A Solution to Semantic Paradox with Defeasible Principles for Truth*, doctoral dissertation, The Ohio State University.

Dretske, F. (2000), "Entitlement: Epistemic Rights without Epistemic Duties?", *Philosophy and Phenomenological Research* 60, pp. 591–606.

Eklund, M. (2002), "Inconsistent Languages", *Philosophy and Phenomenological Research* 64, pp. 251–275.

Eklund, M. (2007a), "The Liar Paradox, Expressibility, Possible Languages", in Beall (2007), pp. 53–77.

Eklund, M. (2007b), "Meaning-constitutivity", *Inquiry* 50, pp. 559–574.

Field, H. (2008), *Saving Truth from Paradox*, Oxford, Oxford University Press.

Field, H. (2009), "What Is The Normative Role of Logic", *Proceedings of the Aristotelian Society Supplementary*, Volume 83, pp. 251–268.

Greenough, P. (2008), "Indeterminate Truth", *Midwest Studies in Philosophy* 32, pp. 213–241.

Grundmann, T. (2011), "Defeasibility Theory", in S. Bernecker and D. Pritchard (eds.), *Routledge Companion to Epistemology*, New York, Routledge, pp. 156–166.

Gupta, A. and N. Belnap (1993), *The Revision Theory of Truth*, Cambridge, MA., MIT Press.

Hare, R. M. (1952), *The Language of Morals*, Oxford, Clarendon Press.

Hofweber, T. (2021), "The Unrevisability of Logic", *Philosophical Perspectives* 35, pp. 251–274.

Hofweber, T. (2009), "Inferential Role and the Ideal of Deductive logic", in B. Armour-Garb, D. Patterson, and J. Woodbridge (eds.), *The Baltic International Yearbook of Cognition, Logic and Communication, Vol 5: Meaning, Understanding, and Knowledge*, New Prairie Press.

Hofweber, T. (2007), "Validity, Paradox, and the Ideal of Deductive Logic", in Beall (2007), pp. 145–158.

Kabay, P. D. (2008), *A defense of trivialism*, PhD thesis, School of Philosophy, Anthropology, and Social Inquiry, The University of Melbourne.

Liggins, D. (2019), "In Defence of Radical Restrictionism", *Philosophy and Phenomenological Research* 98, pp. 3–25.

MacFarlane, J. (2004), "In what sense (if any) is logic normative for thought?", unpublished, 2004.

McGee, V. (1992), "Maximal Consistent Sets of Instances of Tarski's Schema (T)", *Journal of Philosophical Logic* 21, pp. 235–241.

McGee, V. (1985), "A Counterexample to Modus Ponens", *Journal of Philosophy* 82, pp. 462–471.

Murzi, J. and L. Rossi (2020), "Generalized Revenge", *Australasian Journal of Philosophy* 98, 153–177.

Pollock, J. (1986), *Contemporary Theories of Knowledge*, Savage, MD, Rowman and Littlefield.

Price, H. (2011), *Naturalism Without Mirrors*, New York, Oxford University Press.

Price, H. (2013), *Expressivism, Pragmatism and Representationalism*, Cambridge, Cambridge University Press.

Priest, G. (1979), "The Logic of Paradox", *Journal of Philosophical Logic* 8, pp. 219–241.

Ramsey, F. (1929), "General Propositions and Causation", in *Foundations: Essays in Philosophy, Logic, Mathematics and Economics* (1978), Atlantic Highlands: Humanities Press, pp. 133–151.

Restall, G. (2005), "Multiple Conclusions," in P. Hájek, L. Valdés-Villanueva, and D. Westerståhl (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, London: KCL Publications, pp. 189–205.

Ripley, D. (2013), "Paradoxes and Failures of Cut", *Australasian Journal of Philosophy* 91, pp. 139–164.

Ripley, D. (2012), "Conservatively Extending Classical Logic with Transparent Truth", *Review of Symbolic Logic* 5, pp. 354–378.

Rumfitt, I. (2000), "'Yes' and 'No'", *Mind* 109, pp. 781–824.

Russell, G. (2018), "Logical nihilism: could there be no logic?", *Philosophical Issues* 28, *Philosophy of Logic and Inferential Reasoning*, pp. 308–324.

Quine, W.V.O. (1970), *Philosophy of Logic*, Englewood Cliffs, NJ, Prentice Hall.

Sellars, W. (1958), "Counterfactuals, Dispositions and the Causal Modalities", in *Minnesota Studies in Philosophy of Science Volume 2: Concepts, Theories and the Mind-Body Problem*, Herbert Feigl, Michael Scriven, and Grover Maxwell (eds.), Minneapolis, University of Minnesota Press, pp. 225–308.

Thomasson, A. (2020), "A Pragmatic Method for Normative Conceptual Work", in *Conceptual Engineering and Conceptual Ethics*, A. Burgess, D. Plunkett, and H. Cappelen (eds.), Oxford, Oxford University Press, pp. 435–458.

Thomasson, A. (forthcoming 1), "Conceptual engineering: when do we need it? How can we do it?", forthcoming in *Inquiry*, DOI: 10.1080/0020174X.2021.2000118.

Thomasson, A. (forthcoming 2), "How should we think about linguistic function?", forthcoming in *Inquiry*, DOI: 10.1080/0020174X.2022.2074886.

Whittle, B. (2021), "Exceptional Logic", *The Review of Symbolic Logic* 14, pp. 85–111.

Williams, R. (2008), "Ontic Vagueness and Metaphysical Indeterminacy", *Philosophy Compass* 3, pp. 763–788.