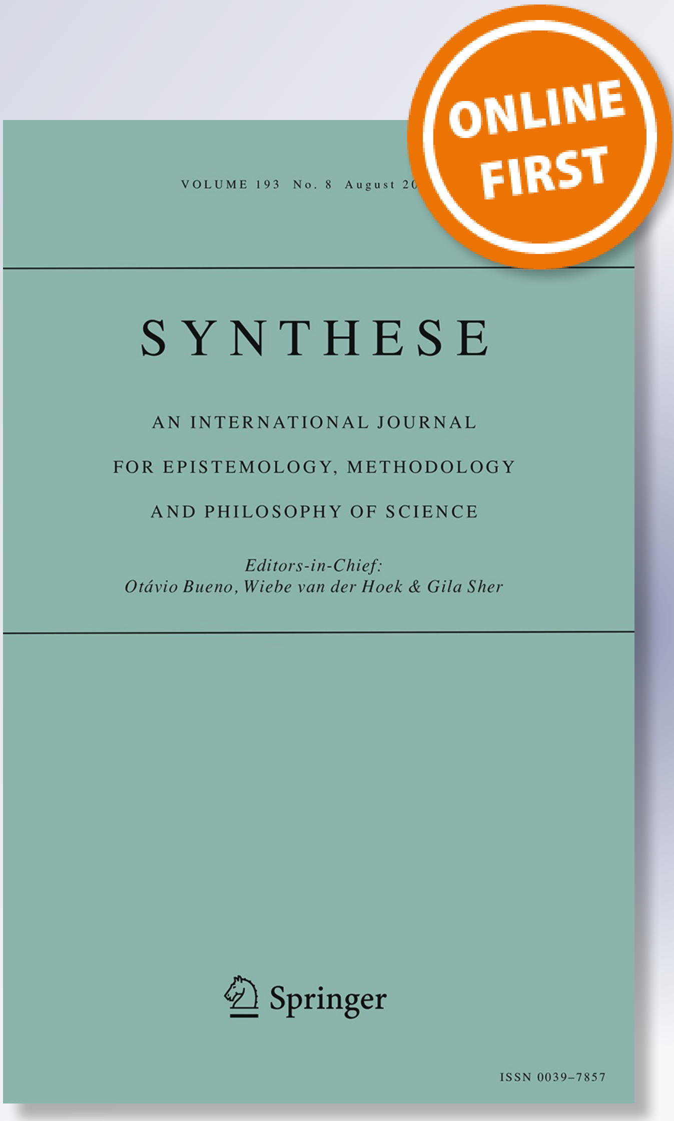
Capturing naive validity in the Cut-free approach

Eduardo Barrio, Lucas Rosenblatt & Diego Tajer

Synthese
An International Journal for Epistemology, Methodology and Philosophy of Science

ISSN 0039-7857

Synthese
DOI 10.1007/s11229-016-1199-5



Springer

Synthese
DOI 10.1007/s11229-016-1199-5

S.I. : SUBSTRUCTURAL APPROACHES TO PARADOX

Capturing naive validity in the Cut-free approach

Eduardo Barrio[1] · **Lucas Rosenblatt**[1] · **Diego Tajer**[1]

Received: 27 July 2015 / Accepted: 13 August 2016

Abstract Rejecting the Cut rule has been proposed as a strategy to avoid both the usual semantic paradoxes and the so-called v-Curry paradox. In this paper we consider if a Cut-free theory is capable of accurately representing its own notion of validity. We claim that the standard rules governing the validity predicate are too weak for this purpose and we show that although it is possible to strengthen these rules, the most obvious way of doing so brings with it a serious problem: an internalized version of Cut can be proved for a Curry-like sentence. We also evaluate a number of possible ways of escaping this difficulty.

Keywords Validity · Paradoxes · Strict-tolerant logic · Substructural logics · Cut

1 Introduction

A naive theory of validity is, roughly, a theory that is capable of talking about its own notion of validity. It should be able to express the validity of its own inferences, which involves internalizing—in a sense to be made precise shortly—its own laws, rules and metarules, using the concept of validity.[1]

[1] It is more or less standard to make a distinction between laws (e.g. the law of Excluded Middle), rules (e.g. the rule of Explosion) and metarules (e.g. Conditional Proof or Reasoning by Cases). A law establishes that a certain sentence is valid. A rule establishes that the argument that goes from certain sentences to another sentence(s) is valid. Finally, a metarule says that if certain arguments are valid, then another argument is valid. These distinctions are not meant to be exhaustive nor exclusive. They are not exhaustive because there might be perfectly legitimate and intelligible principles which are neither laws, nor rules nor metarules. And they are not exclusive because we can understand a law as a 0-premise rule and, similarly, we can understand a rule as a 0-premise metarule. Later on we'll have a chance to see various examples of this.

✉ Eduardo Barrio
eabarrio@gmail.com

[1] Conicet - University of Buenos Aires, Bulnes 642, Zip code: 1176 Ciudad Autónoma de Buenos Aires, Argentina

The whole point of this sort of approach is that it might be, as Fitch (1964) claimed, a formalization of the language we ultimately ought to use. In Fitch's words (p. 397), having a formal language that can adequately deal with its own concepts goes against the traditional idea that

> no formal language seems able to comply with the requirements for serving as an ultimate universal metalanguage within which philosophy, or at least philosophical theories of some depth and insight, could be formalized.

In Kripke (1975) terminology, the naive view is opposed to the orthodox approach, in which the metalanguage is 'essentially richer' than the object language for which the definition of validity is given. The analogy with naive truth theories is hard to miss. As in the case of truth theories, a naive theory of validity tries to accurately represent its own concept of validity without being forced to climb an unending topless ladder of formal metalanguages, following Fitch's metaphor. Nevertheless, like in the case of naive truth theories, naive theories of validity are risky: there are paradoxes of naive validity that need to be addressed.

By now it is familiar how to generate a paradox of this sort. A naive two-place validity predicate $Val(x, y)$[2] for a theory $\mathcal{T}$[3] is usually taken to be a predicate that accurately represents whatever the theory $\mathcal{T}$ declares as valid. Typically, such a predicate is said to satisfy the principles of Validity Detachment (VD) and Validity Proof (VP). Roughly, VD says that ϕ and $Val(\phi, \psi)$[4] imply ψ in $\mathcal{T}$, whereas VP says that if ϕ implies ψ in $\mathcal{T}$, then $\mathcal{T}$ implies $Val(\phi, \psi)$.[5]

$$\text{VD}\ \frac{}{\phi, Val(\phi, \psi) \Rightarrow \psi}$$

$$\text{VP}\ \frac{\phi \Rightarrow \psi}{\Rightarrow Val(\phi, \psi)}$$

[2] To simplify the discussion below we take validity to be a two-place predicate. Needless to say, nothing important depends on this, since in most logical systems multiple premises can be collected into a single conjunction and multiple conclusions (if our theories allow for such a thing) can be collected into a single disjunction. There are some exceptions to this, like the non-contractive theory supported in Lionel Shapiro (2015). However, we'll ignore this possibility here.

[3] It is important to point out that by a theory $\mathcal{T}$ we do not mean a set of formulae closed under some consequence relation. In this context it will be more appropriate to understand a theory as a set of pairs of multisets of formulae closed under certain metarules or, more simply, as a set of arguments closed under certain metarules.

[4] Val is a predicate, so this should be formalized as $Val(\langle\phi\rangle, \langle\psi\rangle)$, where $\langle\rangle$ works as a name forming device. However, to ease the notation we will write $Val(\phi, \psi)$ instead of $Val(\langle\phi\rangle, \langle\psi\rangle)$ throughout the paper.

[5] As it stands, VD can be seen as a rule or as a 0-premise metarule. But sometimes VD is presented as

$$\frac{\Rightarrow Val(\phi, \psi) \qquad \Rightarrow \phi}{\Rightarrow \psi}$$

Zardini (2013) suggests that if there are reasons for thinking that the informal idea behind validity detachment is expressed at least partly by this principle, then those are reasons for thinking that naive validity is actually not faithfully captured by the non-transitive approach. However, this version of VD could be said to be nothing more than a variant of Cut. Since we want to consider a Cut-free approach, we need to stick to the first version.

Most theories couched in a language capable of expressing self-referential sentences and satisfying VD and VP are trivial. In such theories, we can construct a sentence saying of itself that it implies something absurd. This sentence is responsible for what has come to be known as the 'validity paradox' or the 'v-Curry paradox'.[6] Say that $\bot$ is an absurd sentence and that π is the sentence $Val(\pi, \bot)$, i.e., a sentence saying of itself that it implies $\bot$. Then we can reason in the following way:[7]

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\pi, Val(\pi,\bot) \Rightarrow \bot}\, VD}{\pi \Rightarrow \bot}\, LC}{\Rightarrow \pi}\, VP \qquad \dfrac{\vdots}{\pi \Rightarrow \bot}}{\Rightarrow \bot}\, Cut$$

Besides VP, VD and the definition of π, two structural metarules were used in the proof:

$$\text{LC}\ \dfrac{\Gamma, \phi, \phi \Rightarrow \Delta}{\Gamma, \phi \Rightarrow \Delta}$$

$$Cut\ \dfrac{\Gamma \Rightarrow \Delta, \phi \qquad \phi, \Pi \Rightarrow \Sigma}{\Gamma, \Pi \Rightarrow \Delta, \Sigma}$$

What makes this paradox different from, say, the Curry paradox or the liar paradox, is that -at least in some versions of it- no use is made of operational rules. So, for example, the law of Excluded Middle, the rule of Explosion, Modus Ponens, Conditional Proof, and other principles that are sometimes rejected to avoid the usual paradoxes play no role in this paradox. It is clear then that it affects in particular typical paracomplete and paraconsistent approaches [such as those in Field (2008), Beall (2009) and Priest (2006)] that seek to represent naive concepts by weakening the operational rules of the underlying logic.[8]

This suggests that the use of substructural logics looks promising. There are a number of authors that recommend logics without the metarule of Structural Contraction, like Mares and Paoli (2014), Murzi and Shapiro (2015), Priest (2015), Restall (1993), and Zardini (2011), amongst others. In this paper we will mostly ignore Contraction-free theories (although we will say something about them towards the end). The issue of representing validity within a non-contractive theory has been recently explored in Priest and Wansing (2015), Caret and Weber (2015) and Zardini (2014). Instead, we will consider how the v-Curry paradox can be dealt with and how validity can be represented in one Cut-free theory: the Strict-Tolerant (ST) approach developed by Cobreros, Egré, Ripley and van Rooij [see, for instance, Cobreros et al. (2012), Cobreros et al. (2012), Cobreros et al. (2013), Cobreros et al. (2015), Ripley (2012)

[6] For a natural deduction presentation of the paradox, see Beall and Murzi (2013).

[7] Notice that at two points in the proof we are implicitly relying on the identity between π and $Val(\pi, \bot)$.

[8] It is worth remarking that the concept of validity we are discussing is not a purely logical concept, for it can be iterated (i.e., sentences about validity can themselves be valid). In fact, the purely logical notion of validity can actually be captured in any first-order arithmetical theory extending Robinson's arithmetic, as Ketland (2012) and Cook (2014) point out. Also in Field (2008) and Field (forthcoming) it is claimed that under a certain understanding of Val, VD fails, and under another, it is VP that fails. So, the issue is far from being uncontroversial.

and Ripley (2013)].[9] As far as we know this issue hasn't been sufficiently discussed in the literature, though some hints can be found in Ripley (2013). According to him (p. 19):

> Beall and Murzi point out that most non-classical truth theories must treat this paradox differently from the ordinary Curry paradox, despite the apparent similarity. ST, of course, has no such obligation (...) Because this paradox is already accounted for, there is no difficulty in introducing a validity predicate of this sort into a system governed by ST.

We want to cast some doubts on this last claim. For once there is a validity predicate available in the language, we can note something interesting about this approach: we can express, in the object language, not only facts about what follows from what in ST, but also facts about the metarules that hold or fail to hold in ST. For example, we can claim not only that, say, the rule of Modus Ponens is valid using the sentence $Val((\phi \rightarrow \psi) \wedge \phi, \psi)$, but also that (a simplified version of) the metarule of conjunction introduction holds: $Val(\phi, \psi) \wedge Val(\phi, \chi) \rightarrow Val(\phi, \psi \wedge \chi)$.

So a natural question in this context is whether ST plus a validity predicate satisfying VD and VP is capable of truthfully representing the metarules that hold in ST. We will suggest three things. Firstly, in Sect. 2 we will show that ST plus a validity predicate satisfying (generalized versions of) VD and VP does *not* provide a correct characterization of its own notion of validity. The difficulty, in this case, is not that these metarules lead to triviality, but that they are too weak to prove some claims about metarules that we would like to make. More precisely, there are certain metarules that hold in this theory but that cannot be proved to hold, even though we can express them in the language of the theory. Secondly, in Sect. 3 we will claim that this is not enough to completely settle the issue. It turns out that both VD and VP can be strengthened in a very natural way so that those facts about metarules can in fact be represented. Thirdly, in Sect. 4 we will claim that the resulting system faces an apparent problem that has so far been unnoticed in the literature. The problem is that the proposed strengthening brings with it some unwanted consequences. In particular, we will show that the most obvious way to strengthen VD and VP will allow us to prove an internalized version of Cut. We will finish by considering a number of ways of avoiding this difficulty.

2 Internalizing the metarules

Without the rule of Cut, it is impossible to carry out the derivation that leads to the v-Curry paradox. Hence, it is safe to add a validity predicate satisfying VD and VP to ST. In fact, for reasons that will become clear in the next section, we'll add generalized versions of VD and VP to ST . We'll call the resulting system STV.[10]

[9] The Strict-Tolerant approach has also been used to deal with the truth-theoretic paradoxes, the set-theoretic paradoxes and the paradoxes of vagueness.

[10] ST is often presented as a trivalent system with a Strong Kleene matrix. The valid arguments are those in which, if the premises have value 1 (i.e., are strictly true), the conclusion has value 1 or $\frac{1}{2}$ (i.e., is

Definition 1 The system STV contains the following initial sequents and metarules (where Γ and Δ are (finite) multisets):

$$\text{Ax}\ \frac{}{\phi \Rightarrow \phi}$$

$$\text{L}W\ \frac{\Gamma \Rightarrow \Delta}{\Gamma, \phi \Rightarrow \Delta} \qquad \text{R}W\ \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \phi, \Delta}$$

$$\text{L}C\ \frac{\Gamma, \phi, \phi \Rightarrow \Delta}{\Gamma, \phi \Rightarrow \Delta} \qquad \text{R}C\ \frac{\Gamma \Rightarrow \phi, \phi, \Delta}{\Gamma \Rightarrow \phi, \Delta}$$

$$\text{L}\neg\ \frac{\Gamma \Rightarrow \phi, \Delta}{\Gamma, \neg\phi \Rightarrow \Delta} \qquad \text{R}\neg\ \frac{\Gamma, \phi \Rightarrow \Delta}{\Gamma \Rightarrow \neg\phi, \Delta}$$

$$\text{L}\wedge\ \frac{\Gamma, \phi, \psi \Rightarrow \Delta}{\Gamma, \phi \wedge \psi \Rightarrow \Delta} \qquad \text{R}\wedge\ \frac{\Gamma \Rightarrow \phi, \Delta \qquad \Gamma \Rightarrow \psi, \Delta}{\Gamma \Rightarrow \phi \wedge \psi, \Delta}$$

$$\text{L}\vee\ \frac{\Gamma, \phi \Rightarrow \Delta \qquad \Gamma, \psi \Rightarrow \Delta}{\Gamma, \phi \vee \psi \Rightarrow \Delta} \qquad \text{R}\vee\ \frac{\Gamma \Rightarrow \phi, \psi, \Delta}{\Gamma \Rightarrow \phi \vee \psi, \Delta}$$

$$\text{L}\rightarrow\ \frac{\Gamma \Rightarrow \phi, \Delta \qquad \Gamma, \psi \Rightarrow \Delta}{\Gamma, \phi \rightarrow \psi \Rightarrow \Delta} \qquad \text{R}\rightarrow\ \frac{\Gamma, \phi \Rightarrow \psi, \Delta}{\Gamma \Rightarrow \phi \rightarrow \psi, \Delta}$$

$$VD'\ \frac{}{\Gamma, Val(\bigwedge \Gamma, \bigvee \Delta) \Rightarrow \Delta} \qquad VP'\ \frac{\Gamma \Rightarrow \Delta}{\Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta)}$$

It seems clear that in this system we can express (and prove) facts about what follows from what in STV. For example, we can express laws and rules[11] like:

- $Val(\phi \wedge \psi, \psi)$ *(Conj Elim)*
- $Val(\top, \phi \vee \neg\phi)$[12] *(Excluded Middle)*
- $Val(\phi \wedge (\phi \rightarrow \psi), \psi)$ *(Modus Ponens)*
- $Val(\neg\neg\phi, \phi)$ *(Double Neg)*

Note that these formulae are internalized versions of *laws* and *rules* that hold in STV. But, interestingly, the presence of a validity predicate in the object language makes it possible to express *metarules* as well. In light of this, it seems appropriate to say that a validity predicate is naive if it satisfies not only its corresponding rules, but also its corresponding metarules. For example, we can express metarules like:

- $Val(\phi, \psi) \wedge Val(\phi, \chi) \rightarrow Val(\phi, \psi \wedge \chi)$ $(R\wedge)$
- $Val(\phi, \psi) \rightarrow Val(\phi, \chi \vee \psi)$ (RW)
- $Val(\phi \wedge \psi, \chi) \rightarrow Val(\phi, \psi \rightarrow \chi)$ $(R \rightarrow)$

Footnote 10 continued
tolerantly true). However, since we don't know if this semantic characterization is supposed to apply also to extensions of ST such as STV, our presentation of STV is proof-theoretic. More specifically, STV will be presented by means of a (multiple conclusioned) sequent calculus. Also, to simplify things we'll focus on the quantifier-free part of STV.

[11] These are simplified versions of the rules. Strictly speaking, the object language renderings of the rules should mention Γ and Δ, but we omit them for readability. The same applies to the metarules below.

[12] Whenever necessary, we'll assume that the language contains a truth constant $\top$ such that $\Rightarrow \top$ is an initial sequent of our system. This comes in handy for stating in the object language sentences expressing the validity of *laws*.

– $Val(\phi, \psi \vee \chi) \rightarrow Val(\phi \wedge \neg\chi, \psi)$ $(L\neg)$

The issue we want to address is if STV is capable of *proving* these formulae. As for the things in the first bunch, it's clear that every provable sequent is provably valid in STV, in the sense that if $\Gamma \Rightarrow \Delta$ has a proof in STV, then by VP' there is also a proof of $\Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta)$ in STV.[13]

But what about the metarules? Is STV capable of internalizing its own metarules? The answer is 'no'. To explain why, let us be a bit more precise now.

Definition 2 We say that a theory $\mathcal{T}$ *internalizes* a metarule $\mathcal{R}$ of the form

$$\mathcal{R}\ \frac{\Gamma_1 \Rightarrow \Delta_1 \quad ..... \quad \Gamma_n \Rightarrow \Delta_n}{\Gamma \Rightarrow \Delta}$$

if $\mathcal{T}$ proves every instance of [14]

$$\Rightarrow Val(\bigwedge \Gamma_1, \bigvee \Delta_1) \wedge .... \wedge Val(\bigwedge \Gamma_n, \bigvee \Delta_n) \rightarrow Val(\bigwedge \Gamma, \bigvee \Delta).$$

The idea behind this definition is clear enough.[15,16] STV should be able to internalize all (and only) those metarules that hold according to its own standard of validity. For instance, STV should prove internal versions of $R\neg$, $L\wedge$, LW, and so on, but it should not prove an internalized version of Cut.[17]

Unfortunately, the following can be proved:

---

13 Of course, there are *invalid* sequents that are not provably invalid in STV, but that is another matter.

14 Of course, this definition is meant to apply to one-premise metarules as well. So we say that a theory $\mathcal{T}$ *internalizes* a metarule $\mathcal{R}$ of the form

$$\mathcal{R}\ \frac{\Gamma \Rightarrow \Delta}{\Pi \Rightarrow \Sigma}$$

if $\mathcal{T}$ proves every instance of

$$\Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta) \rightarrow Val(\bigwedge \Pi, \bigvee \Sigma)$$

15 One thing we should point out is that the definition we've given for internalizing a metarule seems to be sensible to the order in which the premises of the metarule occur. So for example the sequent $\Rightarrow Val(\psi_1, \psi_2) \wedge Val(\phi_1, \phi_2) \rightarrow Val(\chi_1, \chi_2)$ does not strictly count as the internalization of a metarule with left premise $\phi_1 \Rightarrow \phi_2$, right premise $\psi_1 \Rightarrow \psi_2$ and conclusion $\chi_1 \Rightarrow \chi_2$. Of course, this is harmless. The system we are considering deals with multisets and so the exchange metarules are built in. As a consequence $\wedge$ is a commutative connective and this means that if $\Rightarrow Val(\psi_1, \psi_2) \wedge Val(\phi_1, \phi_2) \rightarrow Val(\chi_1, \chi_2)$ has a proof, $\Rightarrow Val(\phi_1, \phi_2) \wedge Val(\psi_1, \psi_2) \rightarrow Val(\chi_1, \chi_2)$ has a proof as well.

16 We could have instead demanded that for a certain metarule to be internalized the following sequent should have a proof:

$$Val(\bigwedge \Gamma_1, \bigvee \Delta_1), ..., Val(\bigwedge \Gamma_n, \bigvee \Delta_n) \Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta)$$

And indeed, if we have this in STV we can obviously reach the other sequent by L$\wedge$ and R$\rightarrow$.

17 In Priest and Wansing (2015), the authors also introduce a notion of internalization. But, unlike us, they work with a language that only contains a validity operator and their goal is to show that a variation of the v-Curry paradox that uses external validity (roughly, preservation of theoremhood) is not forthcoming without the use of the appropriate form of Contraction, just like the usual v-Curry paradox. So their purpose is quite different from ours.

Proposition 1 *STV cannot internalize some of its metarules.*

An example should be enough to convince the reader that this is so. Consider again (a simplified version of) $L\neg$:

$$\Rightarrow Val(\phi, \psi \vee \chi) \rightarrow Val(\phi \wedge \neg\chi, \psi).$$

Since Cut is not available in STV, we only need to look at its left and right introduction metarules. The most obvious way to derive this sequent is as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{Val(\phi, \psi \vee \chi), \phi \Rightarrow \psi, \chi}\ VD'}{Val(\phi, \psi \vee \chi), \phi, \neg\chi \Rightarrow \psi}\ L\neg}{Val(\phi, \psi \vee \chi), \phi \wedge \neg\chi \Rightarrow \psi}\ L\wedge}{Val(\phi, \psi \vee \chi) \Rightarrow Val(\phi \wedge \neg\chi, \psi)}\ ?$$

But of course the last step is unjustified. The problem is that if the only metarules to introduce Val are VD' and VP', then no derivation of this sequent can be constructed. In particular -as the reader can check for herself- a stronger version of VP' is needed. More precisely, VP' cannot be used to derive this sequent because it is a context-free metarule. This is so for a good reason of course. It being context-free avoids proofs of unwanted sequents such as $\phi \rightarrow \psi \Rightarrow Val(\phi, \psi)$, which intuitively states that if a conditional holds, then the argument from the antecedent to the consequent is valid. A non-context-free version of VP' would look like this:

$$\dfrac{\Gamma, \Phi \Rightarrow \Psi, \Delta}{\Gamma \Rightarrow Val(\bigwedge \Phi, \bigvee \Psi), \Delta}$$

But then we can construct the following derivation:

$$\dfrac{\dfrac{\dfrac{\phi \Rightarrow \phi}{\phi \Rightarrow \phi, \psi}\ RW \quad \dfrac{\psi \Rightarrow \psi}{\phi, \psi \Rightarrow \psi}\ LW}{\phi \rightarrow \psi, \phi \Rightarrow \psi}\ L\rightarrow}{\phi \rightarrow \psi \Rightarrow Val(\phi, \psi)}$$

However, there are a number of ways of avoiding this sort of sequents without demanding that VP' be context-free, as the next section shows.

3 Strengthening the metarules

As we anticipated, it is possible to strengthen the validity metarules of STV. Moreover, it is not hard to see that the strengthening can be carried out in several different ways. Let's start with *validity detachment*. Although for most of our purposes VD' is enough, it will be useful to consider a stronger version of it which we borrow from Zardini (2014):[18]

[18] Instead of using this multiplicative or non-context sharing version of VD^+, we could have used, as one anonymous referee suggests, an additive or context-sharing version of this metarule. This would have the benefits of making the Contraction metarules admissible and of maintaining the validity metarules uniform with the logical metarules of ST, which are additive when they involve two premises and multiplicative when they involve one premise. Of course, given the presence of Weakening and Contraction both metarules are interderivable. In what follows we stick to the non-context sharing version only because it makes some of the proofs below less cumbersome.

$$VD^{+}\ \frac{\Gamma_1 \Rightarrow \phi_1, \Delta_1 \quad ... \quad \Gamma_n \Rightarrow \phi_n, \Delta_n \qquad \Pi_1, \psi_1 \Rightarrow \Sigma_1 \quad ... \quad \Pi_m, \psi_m \Rightarrow \Sigma_m}{\Gamma, \Pi, Val(\bigwedge \Phi, \bigvee \Psi) \Rightarrow \Delta, \Sigma}$$

(where $\bigwedge \Phi$ is the conjunction of $\phi_1, ..., \phi_n$, $\bigvee \Psi$ is the disjunction of $\psi_1, ..., \psi_m$, Γ is $\Gamma_1, ..., \Gamma_n$ and similarly for Π, Δ and Σ). This is nothing more than a generalized version of the metarule for introducing a conditional on the left of a sequent. Omitting the context, it intuitively says that if we accept all the ϕs and reject all the ψs, we should reject the claim that the argument from the ϕs to the ψs is valid.

For *validity proof* things are far more interesting. Taking a cue from sequent calculus presentations of modal logics, we can investigate in a more or less systematic way how to strengthen this metarule. First, we offer a metarule that is reminiscent of the modal logic K:[19]

$$VP^{K}\ \frac{\Gamma, \phi \Rightarrow \psi}{Val(\Gamma) \Rightarrow Val(\phi, \psi)}$$

(where $Val(\Gamma)$ stands for the multiset of formulae obtained by replacing every formula of the form $\chi_1 \rightarrow \chi_2$ in Γ with a formula of the form $Val(\chi_1, \chi_2)$ and any other formula χ with a formula of the form $Val(\top, \chi)$). This metarule is something we expect to hold in a naive theory of validity and it can be showed that a system with this new metarule proves certain sequents involving a validity predicate that have no proof in STV. For instance, the sequent $Val(\phi \wedge \psi, \chi), Val(\top, \phi) \Rightarrow Val(\psi, \chi)$ is unprovable in STV but is provable if the metarule VP^{K} is available, as the reader can easily check. However, the issue with it is that although we can internalize every logical metarule using it, it cannot internalize itself, as there is no proof of the sequent $\Rightarrow Val(\bigwedge \Gamma \wedge \phi, \psi) \rightarrow Val(Val(\Gamma), Val(\phi, \psi))$. Hence, VP^{K} is ultimately too weak for our purposes.[20]

There are different and -we think- better ways to strengthen VP'. The most obvious one is as follows:[21]

$$\text{VP}^{+}\ \frac{Val(\bigwedge \Gamma_1, \bigvee \Delta_1), ..., Val(\bigwedge \Gamma_n, \bigvee \Delta_n), \Gamma \Rightarrow \Delta}{Val(\bigwedge \Gamma_1, \bigvee \Delta_1), ..., Val(\bigwedge \Gamma_n, \bigvee \Delta_n) \Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta)}$$

[19] In the system like the one we are considering, the standard modal metarule corresponding to K is

$$K\ \frac{\Gamma \Rightarrow \phi}{\Box\Gamma \Rightarrow \Box\phi}$$

where $\Box\Gamma$ stands for the multiset $\Box\gamma$ for each $\gamma \in \Gamma$.

[20] Thanks to an anonymous referee for suggesting this very interesting alternative.

[21] This metarule should ring bells for anyone acquainted with sequent calculus presentations of the modal logic $S4$, where the following metarule is given:

$$S4\ \frac{\Box\Gamma \Rightarrow \phi}{\Box\Gamma \Rightarrow \Box\phi}$$

As far as we know this way of presenting $S4$ actually originates with Prawitz' (1965, p. 74) natural-deduction system for $S4$ (thanks to Elia Zardini for the reference). A more recent presentation can be found in Negri (2011).

The reason for having many validity predications on the left is that we want to be able to internalize metarules containing any finite number of premises.[22]

The reader might wonder why in considering how to strengthen VP' we have only focused on the modal logics K and $S4$, thus ignoring, among other modal logics, T and $S5$. The reason is different in each case. We are not considering T because in this setting the most natural candidate for it is, we reckon, the following metarule:[23]

$$\frac{\Gamma, \phi \Rightarrow \psi}{\Gamma, Val(\phi) \Rightarrow \psi}$$

(where $Val(\phi)$ stands for the formula $Val(\chi_1, \chi_2)$ if ϕ is of the form $\chi_1 \rightarrow \chi_2$ and for $Val(\top, \phi)$ if ϕ is any other kind of formula). However, notice that unlike VP^K and VP^+ this metarule only introduces the validity predicate on the left of sequents, not on the right. In fact, it can be proved that it is derivable from VD^+, so it is not very interesting to consider it in this context (not mentioning that it is useless to internalize metarules). Also, observe that already with VD^+ we can prove the sequent corresponding to the modal principle T: $Val(\top, \phi) \Rightarrow \phi$.[24]

As for $S5$, it is well-known that there is an issue with the admissibility of Cut.[25] We could follow Zardini (2014) and introduce the following metarule:

$$VP^Z \ \frac{\Gamma, \Phi \Rightarrow \Psi, \Delta}{\Gamma \Rightarrow Val(\bigwedge \Phi, \bigvee \Psi), \Delta} \ \text{where } \Gamma \text{ and } \Delta \text{ are Val-logical}$$

We say that a formula is *Val-logical* if it is of the form $Val(\bigwedge \Phi, \bigvee \Psi)$ or it is formed by such formulas using the standard logical operations (of course, this can be defined recursively). And obviously, a multiset of formulas Γ is Val-logical if all its formulas are Val-logical.

It is interesting to observe that the issue with the admissibility of Cut for $S5$ reemerges in this context. The following is a proof of the sequent corresponding to the modal principle B:[26]

---

22 Actually, the metarule we've given works properly because metarules only contain one conclusion. But it might also be interesting to consider a version of this metarule with multiple validity predications on the right as well. That is:

$$\frac{Val(\bigwedge \Gamma_1, \bigvee \Delta_1), ..., Val(\bigwedge \Gamma_n, \bigvee \Delta_n), \Gamma \Rightarrow \Delta, Val(\bigwedge \Pi_1, \bigvee \Sigma_1), ..., Val(\bigwedge \Pi_m, \bigvee \Sigma_m)}{Val(\bigwedge \Gamma_1, \bigvee \Delta_1), ..., Val(\bigwedge \Gamma_n, \bigvee \Delta_n) \Rightarrow Val(\bigwedge \Gamma, \bigvee \Delta), Val(\bigwedge \Pi_1, \bigvee \Sigma_1), ..., Val(\bigwedge \Pi_m, \bigvee \Sigma_m)}$$

We will come back to this metarule below in the discussion of the issue of the admissibility of Cut. But for most of our purposes, we can ignore it.

23 In a modal sequent calculus, the metarule corresponding to T is as follows:

$$T \ \frac{\Gamma, \phi \Rightarrow \psi}{\Gamma, \Box\phi \Rightarrow \psi}$$

24 By the way, this was one of the reasons to use VD^+ instead of VD'. This sequent cannot be proved from VD' if Cut is not available.

25 For an overview of the current situation and of the relevant literature, see Negri (2011).

26 In modal logic, this principle can be formulated as follows: $\phi \rightarrow \Box\Diamond\phi$.

$$
\dfrac{
  \dfrac{\dfrac{p \Rightarrow p}{p, \neg p \Rightarrow}\,\text{L}\neg \qquad \Rightarrow \top}{Val(\top, \neg p), p \Rightarrow}\,VD^{+}
  \qquad
  \dfrac{\dfrac{\dfrac{\vdots}{Val(\top, \neg p), \top \Rightarrow Val(\top, \neg p)}}{\top \Rightarrow Val(\top, \neg p), \neg Val(\top, \neg p)}\,\text{R}\neg}{\Rightarrow Val(\top, \neg Val(\top, \neg p)), Val(\top, \neg p)}\,VP^{Z}
}{p \Rightarrow Val(\top, \neg Val(\top, \neg p))}\,\text{Cut}
$$

This sequent cannot be proved without an application of Cut, so Cut is not admissible in the corresponding system. In this proof we use Zardini's VP^Z but in fact it is enough to use the version of VP^+ that allows for validity predications on the right (see footnote 22). In fact, that version of VP^+ (and a fortiori VP^Z) give us, together with VD^+, a proof of the sequent corresponding to the modal principle usually associated with $S5$, namely, the sequent $\neg Val(\top, \neg\phi) \Rightarrow Val(\top, \neg Val(\top, \neg\phi))$. So it looks like we already have an $S5$ish logic with these metarules. That's what creates a problem with the admissibility of Cut.[27]

Of course, this problem might be downplayed by pointing out that once there are paradoxical sentences around, Cut is no longer admissible. But, even so, the Cut-free theorist surely wants to say that the cases where Cut admissibility breaks down are only those which involve paradoxical sentences. Hence, we can say that the issue of Cut admissibility is specially serious in this setting, for if there are sequents that are provable only with Cut, this is not only a matter of lacking 'nice' proofs, but of lacking proofs simpliciter. So perhaps the Cut-free theorist might be better off employing a system where no such thing happens.

In any case, even ignoring the issue with the admissibility of Cut, for our purposes it makes sense to focus on VP^+ instead of the problematic VP^Z because it is enough for the internalization of the metarules, as we'll soon see. So, going back to our target metarules VD^+ and VP^+, it is easy to see that VP^+ implies VP' and that (if the consequence relation is reflexive) VD^+ implies VD'. We will use the name STV^+ for the system that can be obtained from STV by replacing VD' by VD^+ and VP' by VP^+. STV^+ is clearly stronger than STV. It is quite easy to show that STV^+ can internalize its own pritimive[28] metarules, even VD^+ and VP^+ themselves.

Proposition 2 *STV^+ internalizes all its primitive metarules.*

Proof sketch We do one example of an operational metarule (R$\wedge$) and another of a validity metarule (VP^+). The proof of (a simplified version of[29]) the internalization of $R\wedge$ is the following:

[27] We are extremely grateful to an anonymous referee for pointing this out to us.

[28] By a *primitive* metarule we mean a metarule that is an explicit part of the definition of STV^+. At this point it seems necessary to recall a very familiar distinction between two ways in which a metarule might be said to hold. It is one thing to say that a metarule holds if there is a proof from the top sequent to the bottom sequent and it is another thing to say that a metarule holds if the bottom sequent is provable whenever the top sequent is provable. A metarule holding in the first sense is sometimes said to be derivable, while a metarule holding in the second sense is usually called admissible. Obviously, a derivable metarule is also admissible, but the converse might fail. Thanks to an anonymous referee and to Dave Ripley for urging us to clarify this matter.

[29] For matters of readability we prefer to show a simplified version of the proof where Γ is ϕ and Δ is empty. Strictly speaking, the internalization of R$\wedge$ should say $\Rightarrow Val(\bigwedge\Gamma, \phi \vee \bigvee\Delta) \wedge Val(\bigwedge\Gamma, \psi \vee \bigvee\Delta) \rightarrow Val(\bigwedge\Gamma, (\phi \wedge \psi) \vee \bigvee\Delta))$. We make a similar simplification for VP^+.

$$
\dfrac{\dfrac{\dfrac{\phi \Rightarrow \phi \qquad \psi \Rightarrow \psi}{Val(\phi,\psi), \phi \Rightarrow \psi}\, VD^+}{Val(\phi,\psi), Val(\phi,\chi), \phi \Rightarrow \psi}\, LW \qquad \dfrac{\dfrac{\phi \Rightarrow \phi \qquad \chi \Rightarrow \chi}{Val(\phi,\chi), \phi \Rightarrow \chi}\, VD^+}{Val(\phi,\psi), Val(\phi,\chi), \phi \Rightarrow \chi}\, LW}{\dfrac{\dfrac{\dfrac{Val(\phi,\psi), Val(\phi,\chi), \phi \Rightarrow \psi \wedge \chi}{Val(\phi,\psi), Val(\phi,\chi) \Rightarrow Val(\phi, \psi \wedge \chi)}\, VP^+}{Val(\phi,\psi) \wedge Val(\phi,\chi) \Rightarrow Val(\phi, \psi \wedge \chi)}\, L\wedge}{\Rightarrow Val(\phi,\psi) \wedge Val(\phi,\chi) \rightarrow Val(\phi, \psi \wedge \chi)}\, R\rightarrow}\, R\wedge
$$

As for VP^+ (simple version):

$$
\dfrac{\dfrac{\dfrac{\dfrac{Val(\phi,\psi) \Rightarrow Val(\phi,\psi) \qquad \chi \Rightarrow \chi \qquad \zeta \Rightarrow \zeta}{Val(Val(\phi,\psi) \wedge \chi, \zeta), Val(\phi,\psi), \chi \Rightarrow \zeta}\, VD^+}{Val(Val(\phi,\psi) \wedge \chi, \zeta), Val(\phi,\psi) \Rightarrow Val(\chi,\zeta)}\, VP^+}{Val(Val(\phi,\psi) \wedge \chi, \zeta) \Rightarrow Val(Val(\phi,\psi), Val(\chi,\zeta))}\, VP^+}{\Rightarrow Val(Val(\phi,\psi) \wedge \chi, \zeta) \rightarrow Val(Val(\phi,\psi), Val(\chi,\zeta))}\, R\rightarrow
$$

□

Moreover, there is a stronger result available for STV^+ (from which Proposition 2 follows), namely that all its *derivable* metarules can be internalized as well.

Proposition 3 *STV^+ internalizes its derivable* metarules. *That is, if $\Pi \Rightarrow \Sigma$ is derivable from $\Gamma_1 \Rightarrow \Delta_1, \ldots, \Gamma_n \Rightarrow \Delta_n$, then $\Rightarrow Val(\bigwedge \Gamma_1, \bigvee \Delta_1) \wedge \ldots. \wedge Val(\bigwedge \Gamma_n, \bigvee \Delta_n) \rightarrow Val(\bigwedge \Pi, \bigvee \Sigma)$ has a proof.*

Proof sketch We'll prove this result for two-premise metarules only, but it clearly generalizes to n-premise metarules. The reason for not taking the simpler case of one-premise metarules is that Weakening is only required for metarules with two or more premises (except for the case of the one-premise metarules of Weakening, which do require an application of Weakening). First we assume that $\Pi \Rightarrow \Sigma$ follows from $\Gamma_1 \Rightarrow \Delta_1$ and $\Gamma_2 \Rightarrow \Delta_2$. As a consequence, there is a proof of the sequent $\Pi \Rightarrow \Sigma$ in the system that results from adding both $\Gamma_1 \Rightarrow \Delta_1$ and $\Gamma_2 \Rightarrow \Delta_2$ as initial sequents to STV^+. Take that proof and add to each node the formulas $Val(\bigwedge \Gamma_1, \bigvee \Delta_1)$ and $Val(\bigwedge \Gamma_2, \bigvee \Delta_2)$ on the left. The resulting object is a proof of $Val(\bigwedge \Gamma_1, \bigvee \Delta_1), Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Pi \Rightarrow \Sigma$ in STV^+, since adding validity claims does not make any of the steps in the proof incorrect. This means that

(*) $Val(\bigwedge \Gamma_1, \bigvee \Delta_1), Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Pi \Rightarrow \Sigma$ is derivable from
$Val(\bigwedge \Gamma_1, \bigvee \Delta_1), Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Gamma_1 \Rightarrow \Delta_1$ and
$Val(\bigwedge \Gamma_1, \bigvee \Delta_1), Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Gamma_2 \Rightarrow \Delta_2$

Now we reason as follows. First, we can use Reflexivity and VD^+ to obtain $Val(\bigwedge \Gamma_1, \bigvee \Delta_1), \Gamma_1 \Rightarrow \Delta_1$ and $Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Gamma_2 \Rightarrow \Delta_2$

From this we get

$Val(\bigwedge \Gamma_2, \bigvee \Delta_2), Val(\bigwedge \Gamma_1, \bigvee \Delta_1), \Gamma_1 \Rightarrow \Delta_1$ and

$Val(\bigwedge \Gamma_2, \bigvee \Delta_2), Val(\bigwedge \Gamma_1, \bigvee \Delta_1), \Gamma_2 \Rightarrow \Delta_2$
by Weakening in both cases. In virtue of our previous result (*), we infer
$Val(\bigwedge \Gamma_1, \bigvee \Delta_1), Val(\bigwedge \Gamma_2, \bigvee \Delta_2), \Pi \Rightarrow \Sigma$.
Then by VP^+, L$\wedge$ and R$\rightarrow$ we get
$\Rightarrow Val(\bigwedge \Gamma_1, \bigvee \Delta_1) \wedge Val(\bigwedge \Gamma_2, \bigvee \Delta_2) \rightarrow Val(\bigwedge \Pi, \bigvee \Sigma)$. □

This is a very nice result for those interested in internalizing validity within a theory. Moreover, VD^+ and VP^+ seem plausible in themselves, that is, independently of their role in the internalization of metarules. In fact, it could be argued that anyone interested in theories capable of representing the naive concept of validity should be sympathetic to the idea of having metarules like these available.[30]

4 Internalizing cut

Although this approach seems promising, it turns out that VP^+ together with some of the other metarules of ST are enough to prove an internalized version of an unwanted instance of Cut. Consider again the sentence π and recall that as long as VD and LC are available, we can prove that $Val(\pi, \bot) \Rightarrow \bot$:

$$
\begin{array}{rc}
VD & \overline{} \\
 & Val(\pi, \bot), \pi \Rightarrow \bot \\
\text{L}C & \overline{} \\
 & Val(\pi, \bot) \Rightarrow \bot \\
\text{L}W & \overline{} \\
 & Val(\top, \pi), Val(\pi, \bot) \Rightarrow \bot \\
\text{L}W & \overline{} \\
 & Val(\top, \pi), Val(\pi, \bot), \top \Rightarrow \bot \\
VP^+ & \overline{} \\
 & Val(\top, \pi), Val(\pi, \bot) \Rightarrow Val(\top, \bot) \\
L\wedge & \overline{} \\
 & Val(\top, \pi) \wedge Val(\pi, \bot) \Rightarrow Val(\top, \bot) \\
R \rightarrow & \overline{} \\
 & \Rightarrow Val(\top, \pi) \wedge Val(\pi, \bot) \rightarrow Val(\top, \bot)
\end{array}
$$

What the previous derivation shows is that some instance of Cut can be internalized. But observe that it is precisely the instance of Cut involving π. So at least prima facie this is a problem for the Cut-free approach.[31,32]

Moreover, it is not hard to see that if VD^+ is used instead of VD, Cut can be internalized in full generality (that is, not only for the instance including π):

[30] Notice that the proof above only shows that derivable metarules can be internalized in STV^+ and although we suspect that *admissible* metarules can be internalized as well, a different kind of proof is needed for that.

[31] A similar problem occurs if a naive truth predicate is available in the language. It has been noted in Ripley (2013) that STV proves that valid arguments preserve truth:

$$\Rightarrow Val(\phi, \psi) \rightarrow (Tr(\phi) \rightarrow Tr(\psi)).$$

This seems to be a problem, at least to the extent that ST's notion of validity is essentially non-truth-preserving.

[32] Curiously, we can also 'internalize' the claim that Cut does *not* hold for the instance involving π, in the sense that $\Rightarrow Val(\top, \pi) \wedge Val(\pi, \bot) \rightarrow \neg Val(\top, \bot)$ is provable.

$$VP^{+}\frac{VD^{+}\dfrac{VD^{+}\dfrac{\Gamma\Rightarrow\Gamma\qquad\phi\Rightarrow\phi\qquad\Delta\Rightarrow\Delta}{\Gamma, Val(\bigwedge\Gamma,\phi\vee\bigvee\Delta)\Rightarrow\phi,\Delta}\qquad\Pi\Rightarrow\Pi\qquad\Sigma\Rightarrow\Sigma}{\Gamma,\Pi, Val(\bigwedge\Gamma,\phi\vee\bigvee\Delta), Val(\bigwedge\Pi\wedge\phi,\bigvee\Sigma)\Rightarrow\Delta,\Sigma}}{Val(\bigwedge\Gamma,\phi\vee\bigvee\Delta), Val(\bigwedge\Pi\wedge\phi,\bigvee\Sigma)\Rightarrow Val(\bigwedge\Gamma\wedge\bigwedge\Pi,\bigvee\Delta\vee\bigvee\Sigma)}$$

(for readability in the proof above, we take $\Gamma \Rightarrow \Gamma$ as shorthand for the collection of sequents $\gamma_1 \Rightarrow \gamma_1, ..., \gamma_n \Rightarrow \gamma_n$, for each $\gamma \in \Gamma$. We do something similar for $\Delta \Rightarrow \Delta$, etc.)

Given this last sequent, an application of L$\wedge$ and an application of R$\rightarrow$ deliver the official internalized version of Cut, i.e.,

$$\Rightarrow Val(\bigwedge\Gamma,\phi\vee\bigvee\Delta)\wedge Val(\bigwedge\Pi\wedge\phi,\bigvee\Sigma)\rightarrow Val(\bigwedge\Gamma\wedge\bigwedge\Pi,\bigvee\Delta\vee\bigvee\Sigma)$$

How to respond? In what's left of the paper we will consider three different options, although we do not intend them to be exhaustive.

The first one is to simply go back to STV. After all, if only $VP^{'}$ and $VD^{'}$ are available, Cut cannot be internalized in full generality.

Proposition 4 *Cut cannot be internalized in STV.*

Proof sketch On the one hand, it is clear that the Contraction metarules are admissible in STV because of the metarules we've chosen for the connectives (see Negri & von Plato Negri and von Plato, p. 53 for a proof of this claim for a system without a validity predicate). On the other hand, the Weakening metarules are not admissible because $VP^{'}$ is a context free metarule. We can prove that Cut cannot be internalized in full generality by doing a very quick root-first proof search. For consider the following instance of it: $\Rightarrow Val(p, q) \wedge Val(q, r) \rightarrow Val(p, r)$. If this is provable, it is provable from the operational metarules of ST plus $VD^{'}$, $VP^{'}$ and Weakening. So clearly, if this is provable, it comes from the sequent $Val(p, q), Val(q, r) \Rightarrow Val(p, r)$. But it is not hard to see that this sequent has no proof, for it is neither an initial sequent, nor an instance of $VD^{'}$, nor can it be obtained from provable sequents by some other metarule. □

Moreover, there are two stronger results available. First, Cut cannot be internalized in full generality in ST plus $VD^{'}$ and VP^{+}, and second, Cut cannot be internalized in full generality in ST plus VD^{+} and $VP^{'}$. The proofs of these facts are similar to the proof offered for the proposition above, so we omit them. In light of this, STV doesn't appear to be so bad. However, this is a case where appearances are deceiving. In fact, STV is problematic in two different respects. First notice that if $VP^{'}$ is used in place of VP^{+} this amounts to giving up the entire project of internalizing metarules. As we've seen, Proposition 1 shows that something stronger than $VP^{'}$ is needed to carry out the project of representing validity inside our theories. Of course, perhaps there is a version of validity proof that is weaker than VP^{+} (and stronger than $VP^{'}$ and VP^{K}) that can be used to internalize metarules while avoiding the particular instance of internalized Cut that causes trouble. After all, all we've shown is that the most obvious ways of strengthening $VP^{'}$ are problematic.

But even if this is true, there is another problem lurking in the wings. Already in STV it is possible to prove things that are quite problematic. Consider the following sequent, which is similar to an internalized version of Cut:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{Val(\pi,\bot)\Rightarrow\bot}}{Val(\pi,\bot)\Rightarrow\bot, Val(\top,\bot)}\,\mathrm{RW}}{Val(\top,\pi), Val(\pi,\bot)\Rightarrow\bot, Val(\top,\bot)}\,\mathrm{LW}}{Val(\top,\pi)\wedge Val(\pi,\bot)\Rightarrow\bot, Val(\top,\bot)}\,L\wedge}{Val(\top,\pi)\wedge Val(\pi,\bot)\Rightarrow\bot\vee Val(\top,\bot)}\,R\vee}{\Rightarrow Val(\top,\pi)\wedge Val(\pi,\bot)\rightarrow\bot\vee Val(\top,\bot)}\,R\rightarrow$$

Intuitively, this sequent expresses that if the argument from $\top$ to π is valid and the argument from π to $\bot$ is valid (i.e., the premises of Cut), then either the argument from $\top$ to $\bot$ is valid or $\bot$ follows. In other words, either Cut can be internalized (in STV) or something absurd follows. Or, as one anonymous referee puts it, either Cut is admissible or the premise-sequents are not derivable in which case we do not have a failure of Cut after all.

Interestingly, if this is a problem, it is a much more general problem. For consider the following very simple metarule:

$$\frac{\pi\Rightarrow\bot}{\top\Rightarrow\bot}$$

Clearly, this metarule is unsound in STV and in any of its non-trivial extensions. Its premise sequent holds but, of course, its conclusion sequent fails. However, a very simple derivation shows that the conditional $\Rightarrow Val(\pi,\bot)\rightarrow Val(\top,\bot)$ is provable from VD^+ alone (i.e., even if VP^+ is not available), so that means that the corresponding system internalizes an unacceptable metarule. Moreover, if we are allowed to cut on the formula $\bot$ -which we arguably are even in the Cut-free approach- $Val(\pi,\bot)\Rightarrow Val(\top,\bot)$ is provable in STV. But observe that using Weakening on the left, L$\wedge$ and R$\rightarrow$ we can obtain $\Rightarrow Val(\top,\pi)\wedge Val(\pi,\bot)\rightarrow Val(\top,\bot)$, which implies that this instance of Cut can even be internalized in STV.[33] So, it doesn't seem that simply going back to STV is a very good idea.

The second option is to bite the bullet. We can argue that the derivation of the instance of internalized Cut is not really problematic and that VP^+ is in fact correct. Since the conditional is not detachable, we cannot move from

$\Rightarrow Val(\top,\pi)\wedge Val(\pi,\bot)\rightarrow Val(\top,\bot)$

and

$\Rightarrow Val(\top,\pi)\wedge Val(\pi,\bot)$

to

$\Rightarrow Val(\top,\bot)$.

This is unsurprising in the context of STV^+, because sentences such as π are precisely the reason why the conditional fails to be detachable.[34] However, it is hard to see how this kind of response addresses the real problem, which, simply put, is that

33 Thanks to Elia Zardini for suggesting this.

34 Of course, the conditional *is* detachable in the sense that the rule of Modus Ponens ($\phi, \phi\rightarrow\psi\Rightarrow\psi$) is provable, but the metarule of Modus Ponens (if $\Rightarrow\phi\rightarrow\psi$ and $\Rightarrow\phi$, then $\Rightarrow\psi$) does not hold. In fact, it was proved in Barrio et al. (2015) that the metarules that hold in ST are (under a certain translation) the same as the *rules* that hold in the logic LP (Priest's Logic of Paradox), so the fact that the rule of Modus

the predicate Val does not seem to express the concept of validity endorsed by the theory. While the theory is Cut-free in the sense that Cut fails, the theory also proves the claim stating that Cut holds. Moreover, there would be no problem if the instance of Cut that is being internalized is one that is not paradoxical, but the point is that the instance is precisely π, the sentence for which Cut should not hold in the first place.

A third and perhaps more radical option is to weaken the logic even more, by rejecting not only Cut but also the metarules of Structural Weakening and/or the metarules of Structural Contraction. This would avoid the internalization of the problematic instance of Cut in STV^+ and the similar derivation we've presented for STV.

The cost is nevertheless very high. On the one hand, rejecting Contraction or Weakening would most likely mean that the deductive power of the resulting system would be substantially diminished. It will no longer be the case that every classically valid inference is valid, which was one of the main motivations behind the Cut-free approach.

On the other hand, and more importantly, there will be an issue with the internalization of some metarules. Once Contraction and/or Weakening are given up, it is possible to define two kinds of junctions, the so-called additive junctions and the so-called multiplicative junctions. This leaves open the possibility of defining internalization in different ways. In a context where we have two conjunctions available, we have to choose which of the conjunctions features in the definition of internalization. In this respect, one interesting idea is to use the additive conjunction to internalize the additive metarules and to use the multiplicative conjunction to internalize the multiplicative metarules.[35] Interestingly, if we do this we can internalize all primitive metarules without using Weakening or Contraction as long as we use VP^Z instead of VP^+.[36]

Unfortunately, there are still two problems with this approach. First, that the internalization of Cut still goes through without using Weakening or Contraction. Second, that having two ways of internalizing metarules amounts to admitting that there are two ways bunching premises *of metarules* together. And that seems at least odd, given that when these metarules are presented, the premise-sequents can only be bunched together in one way. In fact, at the level of metarules we use sets rather than multisets, so we do not keep track of occurrences of sequents.

5 Conclusion

To finish, we'll summarize what we've done and then we'll try to provide answers to a couple potential objections to the general framework we've adopted. We analyzed the possibility of representing the notion of validity in a dimension that has so far been mostly ignored in the literature: the metarules, i.e., statements to the effect that if certain arguments are valid, then another argument is valid as well. We have shown that

Footnote 34 continued
Ponens fails to hold in LP is enough to infer that the metarule of Modus Ponens fails to hold in ST. See also Zardini (2013) and Fjellstad (2016) for some discussion of this aspect of ST.

35 Thanks to Dave Ripley for this clever suggestion.

36 We need VP^Z because to additively internalize the context-sharing metarules R$\wedge$, L$\vee$ and L$\rightarrow$ without using Contraction, VP^+ is not enough.

STV, which can adequately deal with the v-Curry paradox, is unable to internalize its own logical metarules. We proposed an extension of STV -STV^+- which internalizes all its correct (primitive and derivable) metarules. However, STV^+ also internalizes some incorrect metarules, such as Cut. We explored three alternatives to deal with this problem. First, it is possible to go back to STV and thus reject the project of internalizing the metarules. Second, one can downplay the difficulty using the fact that STV^+'s conditional is not as strong as the classical conditional. Finally, one can restrict the logic even more, by rejecting the Weakening or the Contraction metarules. We leave the issue of which of these alternatives is better -and the possibility of finding different alternatives- for another occasion.

Before finishing we would like to address two objections raised by an anonymous referee. The first is that there are certain 'mixed' sequents involving validity that do not represent laws, nor rules, nor metarules. In fact, it could be argued that it is not clear what the content of these mixed claims is. As examples the referee mentions the sequents $p \rightarrow q \Rightarrow Val(p, q)$ and $\Rightarrow Val(p, Val(\top, p))$ (the first of which we've already considered).

We don't think that this constitutes a major problem. Even though these sequents are unprovable, we believe that they are perfectly intelligible and so that they should be expressible. What might draw someone to the idea of banning these claims -specially, the second one- is understanding Val as purely logical validity. But we are doing no such thing. As we suggested in footnote 8, Val is meant to capture a much broader relation of validity. Hence, in our approach, iterations of validity should be no more strange than iterations of truth predications or iterations of the provability predicate in an arithmetical theory.

To put the point a bit differently, there is a perfectly good sense in which the second sequent is capturing a rule. The rule that goes from p to $Val(\top, p)$. It is just that the conclusion of the rule is in turn capturing another rule. As for the first sequent, if it is argued that its content is unclear, it should follow that the content of its converse $Val(p, q) \Rightarrow p \rightarrow q$ is also unclear. But its converse is just the standard way of saying that validity implies truth preservation.

The second objection we'll look at is that we've only considered a couple of ways of strengthening VP and VD and so we do not have a conclusive argument to the effect that the Cut-free approach cannot capture its own concept of validity. But then, it might be concluded, the general significance of our results is unclear.

Now, admittedly, we haven't covered all the possible options, but the way in which we have provided different strengthened versions of VP was, we dare to say, quite systematic, at least to the extent that we followed well-known developments in modal logic. So we are pretty confident that we've considered the best possible candidates for the project of internalizing metarules.

The last thing we'll do is to mention an obvious option we find worth exploring, but whose thorough analysis lies for obvious reasons beyond the scope of this paper. Given their prominent role in the recent literature on semantic paradoxes, it would be interesting to consider if other substructural theories of truth and validity can internalize their metarules. Of particular interest is the question of whether non-contractive theories (where Cut is valid) fare better than ST in this respect. Unlike ST, it is known that some non-contractive theories do not internalize metarules that are unsound according to their own standards. In particular, Zardini (2014) proved that the Contraction metarules cannot be internalized in his non-contractive system, because to internalize Contraction we would need to apply the metarules of Contraction. So, there seems to be a stark opposition with approaches that reject Cut, whose internalization does not require the use of Cut.

Acknowledgments We are very grateful to Elia Zardini, Dave Ripley and two anonymous referees for extremely helpful comments on previous versions of this paper. Some of this material was presented at conferences in Campinas (CLE) and Buenos Aires (Buenos Aires Logic Group). We also owe thanks to the members of these audiences for their valuable feedback.

References

Barrio, E., Rosenblatt, L., & Tajer, D. (2015). The logics of strict-tolerant logic. *Journal of Philosophical Logic, 44*(5), 551–571.
Beall, J. C. (2009). *Spandrels of truth*. Oxford: Oxford University Press.
Beall, J. C., & Murzi, J. (2013). Two flavors of Curry's paradox. *The Journal of Philosophy*, *110*, 143–165.
Caret, C., & Weber, Z. (2015). A note on Contraction-free logic for validity. *Topoi*, *31*(1), 63–74.
Cook, R. (2014). There is no paradox of logical validity. *Logica Universalis*, *8*(3–4), 447–467.
Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2012a). Tolerance and mixed consequence in the S'valuationist setting. *Studia Logica*, *100*(4), 855–877.
Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2012b). Tolerant, classical, strict. *Journal of Philosophical Logic*, *41*(2), 347–385.
Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2015). Vagueness, truth and permissive consequence. In T. Achourioti, H. Galinon, K. Fujimoto, & J. Martínez-Fernández (Eds.), *Unifying the philosophy of truth*. Dordrecht: Springer.
Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind*, *122*(488), 841–866.
Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.
Field, H., Disarming a paradox of validity. *Notre Dame Journal of Formal Logic*, forthcoming.
Fitch, F. (1964). Universal metalanguages for philosophy. *The Review of Metaphysics*, *17*(3), 396–402.
Fjellstad, A. (2016). Naive modus ponens and failure of transitivity. *Journal of Philosophical Logic*, *45*(1), 65–72.
Ketland, J. (2012). Validity as a primitive. *Analysis*, *72*(3), 421–430.
Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy*, *72*, 690–716.
Mares, E., & Paoli, F. (2014). Logical consequence and the paradoxes. *Journal of Philosophical Logic*, *43*(2–3), 439–469.
Murzi, J., & Shapiro, L. (2015). Validity and truth-preservation. In H. Achourioti, F. Fujimoto, & J. Martínez-Fernández (Eds.), *Unifying the philosophy of truth*. Dordrecht: Springer.
Negri, S. (2011). Proof theory for modal logic. *Philosophy Compass*, *6*(8), 523–538.
Negri, S., & von Plato, J. (2001). *Structural proof theory*. New York: Cambridge University Press.
Prawitz, D. (1974). *Natural deduction. A proof-theoretical study*. Stockholm: Almquist and Wiksell.
Priest, G., & Wansing, H. (2015). External curries. *Journal of Philosophical Logic*, *44*(4), 453–471.
Priest, G. (2015). Fusion and confusion. *Topoi*, *34*, 55–61.
Priest, G. (2006). *In contradiction* (2nd ed.). Oxford: Oxford University Press.
Restall, G. (1993). *On logics without contraction*, Ph.D. Thesis, University of Queensland.
Ripley, D. (2012). Conservatively extending classical logic with transparent truth. *The Review of Symbolic Logic*, *5*, 354–378.
Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy*, *91*(1), 139–164.
Shapiro, L. (2015). Naive structure, contraction and paradox. *Topoi*, *34*(1), 75–87.
Zardini, E. (2011). Truth without contra(dic)ction. *The Review of Symbolic Logic*, *4*, 498–535.
Zardini, E. (2013). Naive modus ponens. *Journal of Philosophical Logic*, *42*(4), 575–593.
Zardini, E. (2014). Naive truth and naive logical properties. *The Review of Symbolic Logic*, *7*(2), 351–384.