Logical Predictivism

Ben Martin[1] · **Ole Hjortland**[1]

Abstract

Motivated by weaknesses with traditional accounts of logical epistemology, consider-able attention has been paid recently to the view, known as anti-exceptionalism about logic (AEL), that the subject matter and epistemology of logic may not be so different from that of the recognised sciences. One of the most prevalent claims made by advocates of AEL is that theory choice within logic is significantly similar to that within the sciences. This connection with scientific methodology highlights a considerable challenge for the anti-exceptionalist, as two uncontentious claims about scientific theories are that they attempt to *explain* a target phenomenon and (at least partially) prove their worth through successful *predictions*. Thus, if this methodological AEL is to be viable, the anti-exceptionalist will need a reasonable account of what phenomena logics are attempting to explain, how they can explain, and in what sense they can be said to issue predictions. This paper makes sense of the anti-exceptionalist proposal with a new account of logical theory choice, *logical predictivism*, according to which logics are engaged in both a process of prediction and explanation.

Keywords Anti-Exceptionalism about logic · Logical methodology ·
Theory choice · Logical abductivism · Explanation · Prediction

1 Anti-Exceptionalism About Logic

Anti-exceptionalism about logic (AEL) is the thesis that the subject matter and epistemology of logic is, contrary to philosophical orthodoxy, similar in significant regards to the recognised sciences [25]. Examples of prominent claims made by anti-exceptionalists include that logical laws are simply generalisations on the most fundamental properties of the world [64, 65], logical laws provide no more norms

✉ Ben Martin
    benjamin.martin@uib.no

    Ole Hjortland
    ole.hjortland@uib.no

1   Department of Philosophy, University of Bergen, Bergen, Norway

for thought than laws from other areas of research [52], and that evidence for logical laws is not wholly *a priori* [47, 48].

Our interest in this paper is with one particularly popular claim made by anti-exceptionalists:

Methodological AEL: Theory choice within logic is similar in important respects to that of the recognised sciences.[1]

Methodological AEL explicitly rejects that logic is *epistemically foundational*, and thus that logical propositions are known immediately through non-inferential means. Proponents of methodological AEL often identify this shared methodology between logic and the sciences as abductivist in nature:

Given any theory, in science, metaphysics, ethics, logic, or anything else, we choose the theory which best meets those criteria which determine a good theory. Principal amongst these is adequacy to the data for which the theory is meant to account. In the present case, these are those particular inferences that strike us as correct or incorrect... Adequacy to the data is only one criterion, however. Others that are frequently invoked are: simplicity, non-(ad hocness), unifying power, fruitfulness. (Priest [44], 217)

Thus, rather than having some direct access to the truth of a logical claim *p*, such as "all instances of modus ponens are valid", as prominent accounts of logical knowledge based upon rational insight [6] or epistemic analyticity [2, 5] suppose, we come to be justified in believing *p* in virtue of being justified in believing a logical theory *L* containing *p*. Further, we come to be justified in believing a certain logical theory *L* because it better accommodates the relevant data, and possesses the relevant theoretical virtues to a greater extent, than competing theories. Call this account of logical epistemology *logical abductivism*.

However, there is good reason to think that *logical abductivism* is inadequate for the anti-exceptionalist's purposes. Firstly, it fails to take adequate account of important features of scientific methodology, and explain how these features occur within logical methodology. By committing oneself to methodological AEL, one takes on the burden to show how exactly logical methodology is similar to the sciences, which further demands that one recognises uncontroversial claims about scientific methodology and demonstrates how exactly these same claims can be translated over meaningfully to logic. Particularly important for the methodological anti-exceptionalist are the uncontroversial facts that:

- Scientific theories attempt to explain certain phenomena.
- Scientific theories (at least partially) prove their worth by making successful predictions.

While methodological anti-exceptionalists have been happy to speak so far about logical theories "capturing" or "accommodating" data, this talk is undeniably vague and falls far short of showing in what respect logics can be said to explain certain

---

[1]Vocal advocates of methodological AEL include Williamson [65], Priest [44, 45], Russell [51] and Hjortland [25].

phenomena, let alone propose predictions which can be tested. If methodological AEL is to be successful, we need a far more detailed account of how these features of scientific methodology are replicated within logic.

Secondly, given that the available data which any adequate theory must "fit", according to abductivism, will necessarily underdetermine the correct theory, abductivism is forced to require that we choose our preferred theory on the basis of further theoretical virtues, such as simplicity, deductive strength, and unifying power. However, no viable account has yet been provided of why logicians should use these virtues to dictate theory choice (that is, why such reliance upon these virtues is rational), nor even what sense we can make of these virtues within the context of logic.[2] Without such an analysis (and justification) these putative theoretical virtues can be given no more weight than a mere list of subjective prompts for the practitioner, leading to the logical abductivist's account of logical methodology being essentially incomplete. If such virtues are to be appealed to in an account of logical methodology, more detail on their identity and particular role within theory choice is required.

This paper aims to rectify both of these weaknesses of current anti-exceptionalist accounts of logical methodology, by proposing a novel *predictivist* account of logical methodology. According to *logical predictivism*, logical theories aim to explain a certain phenomenon, validity, and are at least partially evaluated on their ability to make successful predictions. Further, while logical predictivism admits that ultimately successful predictions will underdetermine the logician's theory choice, unlike in the case of current discussions of logical abductivism, we specify here only those further theoretical virtues which are clearly appealed to in actual logical practice to justify logical theories. Happily, for the anti-exceptionalist's proposal, it just happens that these theoretical virtues have a strong resemblance to those theoretical virtues often discussed in connection with the empirical sciences [32, 39].

This account will have the benefit of not only substantiating the anti-exceptionalist's claim that logical methodology is similar to scientific methodology, for we take it that the role of explanation and prediction within science are uncontroversial claims about scientific methodology, but allow the anti-exceptionalist to stop speaking in the unhelpfully vague terms of logics "capturing" or "accommodating" data. We can replace this talk with a more precise account of logics making predictions and these predictions being tested. Overall, according to the picture of logical methodology proposed, the success of a logical theory should be judged not on its ability to simply "accommodate" data, but to make successful predictions.

In emphasising the importance of successful predictions over mere accommodation of data, we are further aligning our position here with the predictivist proposal in the philosophy of science—that successful novel predictions are theoretically more valuable than mere accommodations of some data [24]. However, given that multiple versions of the position are found in the philosophy of science literature, two points

---

[2]Indeed, the only virtue which has been clarified with any precision, *deductive strength*, once clarified, clearly is not a theoretical virtue (cf. Hjortland [25], Russell [54]). Of course, this is not just a problem for theory choice within logic. There is now a well-established literature questioning both the identity and value of the purported theoretical virtues within the empirical sciences. See, for example, Hunt [26] and Schindler [55].

of clarification on our commitment here are necessary. Firstly, unlike advocates of *strong* predictivism, we do not propose that prediction is intrinsically superior to accommodation, rather that theories which are predictively successful are often better supported eventually than those which merely accommodate existing data, due to the latter theories' tendency to overcompensate for recalcitrant data in an ad hoc fashion.[3]

Secondly, we do not require the data against which a prediction is tested to be *unknown at the time* of the theory's formulation for the prediction to be novel; as the so-called *temporal* interpretation of predictive novelty requires. Instead, we only require that the theory was *not* constructed specifically to fit that data, which is known as the *heuristic* interpretation of novelty (a point we shall return to later). We will not argue for either of these two claims here, as they have been defended at length elsewhere (see Hitchcock and Sober [24]; Lipton [34]: Ch. 10; Worrall [67]). Instead, our case for logical predictivism will focus on the clarity and detail it offers on the mechanism of logical theory-choice (which its competitor logical abductivism fails to), and how it is able to make good sense of actual logical practice.

The rest of the paper runs as follows. Section 2 clarifies what we mean by logical theories, and outlines our general method. In Sections 3 and 4 we show how two cases of logical evidence can plausibly be interpreted in terms of talk of explanation and prediction, and in Section 5 we detail how further theoretical virtues can impact logical theory choice. Finally, in Section 6 we outline certain important upshots of *logical predictivism* for AEL.

2 Explanation and Prediction in Logic

In what follows, we tell an idealised story about how explanation and prediction work within logic, informed by real logical arguments and practice. In so appealing to this practice, we aim to show that:

i)   There are phenomena which logical theories are attempting to explain.
ii) The success of these explanations are (at least partially) judged by how successful the predictions, made on the basis of these explanations, are.

Logics then have phenomena they attempt to explain, and use successful predictions as a criterion to judge the fruitfulness of these explanations.[4] If our predictivist account is ultimately successful in reflecting how logicians go about supporting their theories, this will go a considerable way to substantiating the methodological anti-exceptionalist's claims.

We will build our case for how logics explain and predict by highlighting three distinct types of data which logicians use to support their theories. We begin with classical logic's attempt to capture the validity of steps within informal mathematical

---

[3]For more on the value of prediction over accommodation, see Hitchcock and Sober [24] and Lipton [34]: Ch. 10.

[4]Although, as will be suggested in the final section, it's highly likely further considerations are involved in judging a logical theory's explanatory success.

proofs. We then move onto logical theories' attempts to explain why certain steps in vernacular arguments are valid, and lastly look at how our other theoretical commitments can contribute to logical theory choice, and thus form part of a logic's evidential base (which Kuhn ([32]: 321-2) calls "external consistency").

Before we move onto our three cases, we need to say more about what we mean by logical *theories*. After all, it seems very strange indeed to say that formal logical *systems* themselves can explain or make predictions—they are merely calculi. However, by logical theories we do not simply mean logical systems, such as *Strong Kleene*, *Logic of Paradox* and *First Degree Entailment*, which are uninterpreted calculi. Rather, logical theories are theories *of validity*. They are what logicians ultimately defend when arguing over validity. While logical systems can be used to model many different phenomena, whether this be electrical gates or information states, logical theories so conceived are aimed at validity. To this extent, logical theories contain propositions, just like other theories do. While these theories contain propositions expressing the formal system, they also contain other elements such as:

- *Representation rules*: rules informing us how to translate between natural language sentences and the formal language contained in the theory.
- *Semantics*: an informative semantics for the elements of the calculus (such as how to interpret the functions and their arguments/outputs).
- *Account of consequence*: an account of what the consequence relation is.
- *Underlying philosophical assumptions*: claims presupposed by other elements of the theory, such as how to conceive of truth given the applied semantics.

Thus, while logicians do disagree over which formal logical system we ought to use when assessing the validity of particular arguments, they also potentially disagree over how to understand the concept of *logical consequence*, the identity of the designated values preserved by the consequence relation, and how many truth-values these systems commit us to.

This list is in no way intended to be exhaustive. Identifying all of the components of a logical theory is far beyond this paper's scope, and would merely detract us from its main business. Instead, the list is intended to demonstrate that we mean more than simply formal logical systems by our talk of logical theories. In effect, by advocating a logical system within an applied setting, one takes on these other commitments which constitute the overall logical theory. It is this theory as a whole, and not just the formal system, which according to our account engages in the business of explanation and prediction. Concrete examples of what are contained within a theory will be outlined through our cases. With this in mind, let us move onto our cases for thinking that logical theories explain and predict. We begin with the example of classical logic's success in regimenting mathematical proofs.

3 Capturing Mathematical Proofs

Classical logic enjoys a privileged position among logics. It is commonly taught as *the* logic in introductory courses on logic, both within philosophy and mathematics departments, and is the subject of the majority of introductory and intermediate

logic textbooks. While there are numerous reasons why classical logic holds this position, not least the versatility of the formal tools it provides, one significant reason undoubtedly is its successful application within mathematics. The significance of this application is only emphasised by the fact that certain of the founding figures of modern formal logic (notably, Frege and Russell) at the turn of the twentieth century designed classical logic (both its first- and second-order variants) in order to regiment mathematical proofs, a task which its predecessor, syllogistic logic, was not well suited to.[5]

Classical logic regiments mathematical proofs by taking informal versions of these proofs and attempting to explicate their underlying structural features by formalizing ubiquitous steps within the informal proofs. In other words, it makes transparent some important features of the proofs' forms. The question, of course, is what is achieved by so regimenting these informal proofs? According to this account, it is because the logical theory aims to *explain* why the particular steps used within informal proofs—and not others—are valid. In other words, the logical theory has some phenomenon to explain: the validity of those steps found within informal proofs.[6]

The question then comes, how does the theory go about successfully explaining the validity of these steps? To see how this is possible, we must start with the data which initially motivates the theory—informal proofs. Here are some:

Definition 1 Some integer $n \in \mathbb{Z}$ is called *odd* iff $n = 2k + 1$ for some integer k; *even* iff $n = 2k$ for some k.

Theorem 1 *Assume $x \in \mathbb{Z}$. If $x^2 - 4x + 7$ is even, then x is odd.*

Proof We prove our result indirectly. Suppose x is even, and let $x = 2k$ for some $k \in \mathbb{Z}$, so $x^2 - 4x + 7 = (2k)^2 - 4(2k) + 7$. Then, $(2k)^2 - 4(2k) + 7 = 4k^2 - 8k + 7 = 2(2k^2 - 4k + 3) - 1$, and so $x^2 - 4x + 7$ is odd. Thus, assuming x is even, $x^2 - 4x + 7$ is odd. $\square$

Theorem 2 *For all $n \in \mathbb{Z}$, if $3n + 2$ is odd, then n is odd.*

Proof We prove our result indirectly. Assume n is even, and so $n = 2k$ for some $k \in \mathbb{Z}$. Consequently, $3n + 2 = 3(2k) + 2 = 6k + 2 = 2(3k + 1)$. But, then $3n + 2$

---

[5]We won't take a stand here on what these individuals took themselves to be demonstrating, or aiming to demonstrate, with this regimentation. In the case of Frege, for example, interpretation of his logicism is an industry unto itself. See Jeshion [27], Kitcher [29] and Weiner [62]. Though, in some cases, such as that of Gentzen ([22]: 183), it's clear logicians were primarily motivated to model the rules of inference mathematicians used—a project continued by some contemporary logicians, such as Tennant ([61]).

[6]All steps within an informal proof? This is an interesting question. *Certainly not* is the answer, for there are certain steps made within informal proofs which are manipulations in the meaning of introduced mathematical terms, and logical theories have no interest in these. A somewhat tentative answer is that it aims to explain the validity of those steps which are the most general, and found in proofs across most, if not all, areas of mathematics. That is, those steps not reliant upon the particular mathematical axioms or definitions. This would certainly fit with the historical presumption that logical laws ought to hold with the utmost *generality*, on which see Section 4.

is even, as $2(3k + 1) = 2j$ for some $j \in \mathbb{Z}$, where $j = 3k + 1$. So, if n is even, then $3n + 2$ is even. □

And one further from set theory:

Definition 2 **Union** ($\cup$): $A \cup B = \{x \| x \in A \text{ or } x \in B\}$.

Theorem 3 *For all sets X, Y, Z, if $X \subseteq Y$, then $X \cup Z \subseteq Y \cup Z$.*

Proof We prove by contraposition. Let A, B, C be sets such that $A \cup C \nsubseteq B \cup C$. Consequently, for some a, $a \in A \cup C$ but $a \notin B \cup C$, which ensures that $a \in A$ or $a \in C$ but $a \notin B$ and $a \notin C$. Now, given that $a \notin C$, it must be the case that $a \in A$. Thus, $a \in A$ and $a \notin B$, and so $A \nsubseteq B$. Consequently, if $A \cup C \nsubseteq B \cup C$ then $A \nsubseteq B$. □

Now, while all of these proofs contain their own particular moves, whether this be the manipulation of equations in Theorem 1 and 2, or the explication of set-theoretic notions within Theorem 3, the logician may well see a pattern here in the *form* of the proofs. Particularly, that all of the proofs claim to establish that *if* some claim φ holds *then* another claim ψ also holds by demonstrating that *if ψ fails* to holds, *then φ* also *fails* to hold. In other words, that these three acceptable informal proofs for the mathematician all contain an instance of the following proof step:

$$\frac{\text{If not } \psi \text{ then not } \varphi}{\text{If } \varphi \text{ then } \psi}$$

What we have here in effect are the logician's first two working hypotheses—that, i) the three proofs above are of fundamentally the same basic form, and ii) that *this* shared form is the one given above.

Next, as the logician is concerned with providing an account of which steps within a proof are valid, rather than simply making a claim about these three proofs, she forms a general hypothesis about arguments of this form, motivated by the fact that she's putatively found instances of mathematicians accepting proofs of this form (namely, the proofs for Theorems 1-3):

Hypothesis 1

All arguments of the form

$$\frac{\text{If not } \psi \text{ then not } \varphi}{\text{If } \varphi \text{ then } \psi}$$

are valid.

Note that all we have so far is a generalisation, albeit one that can be falsified. This generalisation does *not* constitute an explanation of validity, any more than the generalisation that "all swans are white" is an explanation of why swans are white. In order to *explain* validity, we need to build a theory which shows why arguments of this form are valid (if they really are, that is.)

What does this theory look like? Simply, it's constituted of definitions and laws which when combined aim to underwrite the (hypothesised) true generalisation given

above. Let's consider a proto-classical theory introduced just to deal with the validity of arguments of the above form. This theory would need *representation rules* to formalise the English language indicative conditional as a two-place connective $\rightarrow$, and the English language negation as an unary operator $\neg$. Further, the theory would postulate that an implication, $\varphi \rightarrow \psi$, is true if and only if φ is false or ψ is true (i.e. the material implication is assigned the Boolean semantics), and that a negated proposition $\neg\varphi$ is true if and only if φ is false. In addition, the theory would postulate that an argument from premises Γ to conclusion φ is valid just in case, for every valuation, if each member of Γ is true, so is φ (i.e. an account of the consequence relation), and that any sentence is either true or false, and not both (i.e. bivalence and contravalence). Thus, such a theory would look like the following:

Theory A

- **Definition 1**: Let $\varphi \rightarrow \psi$ be Boolean material implication.
- **Definition 2**: Let $\neg\varphi$ be Boolean negation.
- **Representation Rule 1**: $\ulcorner$ if φ then $\psi \urcorner = \ulcorner\varphi \rightarrow \psi\urcorner$.[7]
- **Representation Rule 2**: $\ulcorner$ not $\varphi\urcorner = \ulcorner\neg\varphi\urcorner$.
- **Law 1**: For every valuation, all sentences are either true or false, and not both.
- **Law 2**: An argument is valid iff, for every valuation v, if every premise is true in v, the conclusion is true in v.[8]

All of these components of the theory contribute towards an explanation of why Hypothesis 1 is true, and thus why instances of contraposition are valid. In sum, the theory says that instances of the generalisation are valid *because*: i) their underlying form ensures that whenever the premises are true the conclusion is also true, while providing a further explanation of *why* this is the case for arguments of this form, through the theory's definitions, representation rules and Law 1; and, ii) the results from part i) of the explanation ensure that the argument is valid, due to Law 2.[9]

---

[7]Note that we are oversimplifying these representation rules for ease of exposition here. It's quite clear that logicians do not simply treat all occurrences of "if...then..." claims within English as expressing the material conditional, and similarly for the other connectives. Instead, they only treat certain cases of these vernacular claims as related to the formal connectives. Involved here is a process of idealisation, discussed in Section 6.

[8]Law 2 is a standard account of validity, but by no means the only one. Other theories of validity give proof-theoretic accounts or information-theoretic accounts, and even the alethic accounts of validity differ on the details. Each is an attempt to provide a technical explication of an intuitive (perhaps, modal) account of logical consequence, and each will differ in the ultimate explanation they provide of validity. We'll come back to this final point in Section 6.

[9]As we might expect, given the similarities which logical predictivism proposes between logical and scientific methodology, some of the same concerns raised over laws within the sciences will also arise when it comes to interpreting these logical laws. Are they, as Humeans propose, for example, mere generalisations of instances, or should such laws be understood as having some modal content? For a review of the options, see Carroll [14]. The proposal here, happily, seems consistent with both. Ultimately, the best account of logical laws will need to take into account both our best theory of scientific laws and the peculiarities of logic as a subject matter and its practice—something it is beyond the scope of this paper to comment on. Consequently, we reserve judgement at present on how best to understand these laws. Many thanks to a referee for pushing us on this point.

While this toy theory "saves the data", in that it allows for the generalisation above to be true, this is equally true of numerous other theories. It is not difficult to build a theory that accommodates this particular generalisation, and provides a potential explanation of its truth. Consequently, at present, we do not have any reason to commit ourselves to one of these theories (and, explanations) over its competitors.

How then does the logician show that her theory fares positively in comparison to competitors which also fit the hypothesised generalisations? By making *predictions*. This is possible because those postulates within her theory which, when combined, fit the current generalisation also ensure that other inferential steps are valid. Further, because underlying the logician's attempt to explain the validity of steps within informal proofs is the assumption that mathematicians' judgements over the (un)acceptability of these putative proofs are a reliable guide to their validity, the judgements of mathematicians can be used to test these predictions. Thus, if the predictions are correct, the logician ought to find instances of these forms of argument within informal proofs—they ought to be found to be acceptable moves by mathematicians.

In so testing the theory with predictions we actually have three stages. Firstly, one must draw out the consequences of the theory's postulates.[10] Within the context of *Theory A*, such consequences would include:

- **Consequence 1**:

    All arguments of the form

    $$\frac{\varphi \quad \varphi \rightarrow \psi}{\psi}$$

    are valid.

- **Consequence 2**:

    All arguments of the form

    $$\frac{\neg\neg\varphi}{\varphi}$$

    are valid.

- **Consequence 3**:

    All arguments of the form

    $$\frac{\varphi \rightarrow \psi \quad \varphi \rightarrow \neg\psi}{\neg\varphi}$$

    are valid.

---

[10]One will notice we are talking of the *consequences* of the theory's postulates here, which presupposes there are some rules of inference the logician can rely upon to draw out these consequences. Does this mean the logician is required to rely upon her own theory in testing its adequacy? We come back to this problem in the final section.

- **Consequence 4**:

    Not all arguments of the form

    $$\frac{\varphi \qquad \psi \to \varphi}{\psi}$$

    are valid.

These consequences of the theory are then operationalised into concrete predictions to be tested. For example, Consequence 1 would be operationalised into the prediction:

- **Prediction 1**:

    Steps within informal proofs of the form

    $$\frac{\varphi \qquad \text{If } \varphi \text{ then } \psi}{\psi}$$

    are found acceptable by mathematicians.

Likewise for the other consequences, such as Consequences 2-3, which propose that all arguments of a certain form are valid. In comparison, those consequences of the theory which propose that not all arguments of a certain form are valid, such as Consequence 4, will be operationalised into predictions such as:

- **Prediction 2**:

    Steps within informal proofs of the form

    $$\frac{\varphi \qquad \text{If } \psi \text{ then } \varphi}{\psi}$$

    are not found acceptable by mathematicians.

The final stage is then of course to test these predictions against further informal proofs which haven't yet been relied upon to motivate the logician's theory.[11] This requires "collecting" informal proofs with the expectation of finding instances of the forms of argument given above. Additionally, in order to test Prediction 2 and those similar, the logician will need to look at instances of "pseudo-proofs", where mathematicians judge that inferential mistakes are being made. If the result of this search finds instances that fit the predictions, then the theory finds itself further supported,

---

[11]Note that, as we clarified in Section 1, what is required here are not necessarily *temporally* "novel" predictions made about data as of yet uncollected, but merely that the theory be tested by further data in addition to that already used to motivate the initial hypotheses. Thus, the account here is totally accepting of so called "retrodictions". For more on the exaggerated role of temporally novel predictions within scientific theory choice, see Brush [11, 12].

and, inversely, if the result of this search consistently finds instances that contradict the predictions, then the theory faces problems.[12]

If our logician were to engage in this process, she may find herself pleasantly surprised, for there are indeed clear cases of informal proofs of the form given in Consequence 3, well-known as *proofs by contradiction*. Further, by considering the pedagogical literature, she finds that mathematicians preclude their students from using inferences of the form detailed in Consequence 4 within proofs (thus, corroborating Prediction 2). She could cite, for example, such an example of a faulty proof:

Theorem 4 *Assume $x, y \in \mathbb{Z}$. If xy is divisible by 3, then x is divisible by 3 or y is divisible by 3.*

Pseudo-proof If xy is divisible by 3, then $xy = 3k$ for some $k \in \mathbb{Z}$. Thus, as $x = 3l$ for some $l \in \mathbb{Z}$ or $y = 3l$ for some $l \in \mathbb{Z}$, x is divisible by 3 or y is divisible by 3. $\square$

The more successful the predictions made, the more successful the theory is deemed. Further, these successes will be weighed against failures which cannot be subsequently explained away. However, as always, successes count for more than failures, even if they are left unexplained for a time—a predictively successful theory is able to live with a certain mass of anomalies.

Theory A only provides a partial picture, of course. For example, it only offers a bare-boned account of the logical connectives, and gives no logic for quantifiers, both of which are required for a satisfactory account of mathematical proofs. Even more importantly, a complete theory has to include hypothetical forms of proof steps. This complication is worth briefly mentioning, since it has direct bearing on the theory of validity. Consider, for example, the following informal proofs:

Theorem 5 *For all $n \in \mathbb{Z}$, if n is odd then n^2 is odd too.*

Proof As n is odd, $n = 2k + 1$ for some k. So, $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. But, then n^2 is just of the form $2j + 1$, and so odd too. $\square$

Theorem 6 *Assume $x \in \mathbb{Z}$. If $5x + 7$ is even, then x is odd.*

Proof Assume $5x + 7$ is even. So, $5x + 7 = 2k$ for some integer k. Thus, by subtraction we get $x = 2k - 4x - 7$. So, $x = 2k - 4x - 7 = 2k - 4x - 8 + 1 = 2(k - 2x - 4) + 1y$ But, this is just $x = 2j + 1$, and so x is odd. $\square$

---

[12]Of course, just as within the sciences, we wouldn't expect a certain number of unsuccessful predictions to be the death of a theory if it's otherwise successful. What is endorsed here, therefore, is not a form of naïve falsificationism, according to which any unsuccessful prediction falsifies the theory—there are many reasons for why a prediction may be (apparently) unsuccessful. How exactly logical theories can deal with unsuccessful predictions is a matter we will come to in the next section.

In these examples, the theorem is established directly by conditional proof. The logician might then form the following hypothesis:

Hypothesis 2

All arguments of the form

$$\frac{\begin{array}{c}[\varphi]\\ \vdots\\ \psi\end{array}}{\text{If }\varphi\text{ then }\psi}$$

are valid.

However, the supposed validity of this essential proof step cannot be accounted for in *Theory A*. The reason is that the theory's law about validity (Law 2) only tells us when an argument from sentence-premises to a sentence-conclusion is valid. But conditional proof, and other hypothetical proof steps, involve arguments with assumptions as premises. Put differently, they are meta-arguments: arguments from arguments to arguments.[13]

Consequently, the theory will have to be generalized to account for these more complicated argument forms. Of course, this is standard in the semantics of both classical and nonclassical logics. For our proto-classical theory we can give a definition of satisfaction (for an argument), and use that to introduce an account of validity for meta-arguments:

Theory B

- **Definition 1**: Let $\varphi \to \psi$ be Boolean material implication.
- **Definition 2**: Let $\neg\varphi$ be Boolean negation.
- **Representation Rule 1**: $\ulcorner$ if φ then ψ $\urcorner = \ulcorner\varphi \to \psi\urcorner$.
- **Representation Rule 2**: $\ulcorner$ not $\varphi\urcorner = \ulcorner\neg\varphi\urcorner$.
- **Law 1**: For every valuation, all sentences are either true or false, and not both.
- **Law 2′**: An argument is satisfied by a valuation v iff either some premise is false in v or the conclusion is true in v.
- **Law 3**: A meta-argument is valid iff, for every valuation v, whenever the premise-arguments are satisfied, the conclusion-argument is satisfied.

Admittedly, *Theory B* sacrifices the more intuitive description of validity in favour of a less intuitive generalization in the form of Law 3. But, the introduction of the new laws has advantages: First, *Theory B* not only accommodates the conditional proof data, it also allows predictions of other common hypothetical steps in mathematical proofs (e.g., *reductio ad absurdum* and proof by cases).[14] Second, the generalization

---

[13]This is most easily seen for conditional proof when it is expressed in its sequent calculus form:
$$\frac{\Gamma, \varphi \Rightarrow \psi}{\Gamma \Rightarrow \varphi \to \psi}$$

[14]Note that the *reductio* proof step discussed above, as Consequence 3, was formulated with conditionals so as to avoid the complication of suppositions.

in Law 3 in fact subsumes Law 2 of *Theory A*. Since an ordinary argument is a special case of a meta-argument with zero premises, such an argument will be valid in *Theory B* under precisely the same circumstances as those stated in *Theory A*. In what follows, as nothing relevant to our case hangs upon the distinction, we will bracket those complications arising from meta-arguments and stick with the simpler cases.

The account of evidence given so far for a logical theory is clearly incomplete, for logic is traditionally thought to provide a *general* account of validity, rather than simply an account of the validity of mathematical proofs. While we will move onto consider these further types of evidence in the following sections, concentrating on this more restricted case initially is informative, and has highlighted several important features of the proposed predictivist account.

Firstly, logical theories are deemed successful in virtue of predictive successes, and given that different logics provide different predictions, we are able to assess their relative success. Further, as there is no absolute value against which theories' predictive successes are judged, the level of predictive success, and thus the strength of evidence for a logical theory, is judged by comparison to that of other theories. Consequently, as with theories in other areas of inquiry, logical theories are assessed on the basis of their success *relative to competitors*. Of course, there may be some minimal criteria, or basic level of predictive success that any theory must meet in order to even be part of the conversation of theory choice, but what these requirements would be exactly are at present unclear.

Secondly, not only can distinct logical theories deliver divergent predictions about a set of argument forms, but these logics can provide varied explanations for why instances of the same argument form is valid. After all, classical and intuitionist logicians agree that all instances of modus ponens are valid, but disagree on why. That it is possible for two logics to deliver the same predictions in a set number of cases, while having different explanations for these cases of validity, is due solely down to the fact that there are many theoretical routes to the same predictions. In other words, there are many available laws that can underwrite the same predictions. We take it to be a strength of the current account that it is able to explain how different logical theories could agree in terms of certain predictions while disagreeing on *why* instances of these argument forms are valid.[15]

Thirdly, for predictions to be used to test a theory, we need some data to test the predictions. At least in the case of classical logic's attempt to capture the validity of informal proofs, it's clear that what's taken as a reliable indicator of validity, and thus suitable data to test the consequences of the theory, are the judgements of mathematicians regarding acceptable informal proofs. Whether we can broaden the data to encompass other types of judgements is an important question, which we will

---

[15] Whether it's possible to have two theories which are predictively *equivalent* while diverging in their explanatory content is an interesting question. For example, there are various non-standard interpretations of both classical and non-classical logics, which deliver the same theorems and rules of inference as the standard interpretations, but with different semantics. In these cases, we couldn't choose between the theories on the basis of their comparative predictive successes, but would have to rely on other measures, such as *explanatory power*. Saying anything of any great detail about these cases is beyond the scope of the present paper, but we will touch on the matter in the final section.

consider in the next section. However, what is clear is *how* judgements regarding what are *acceptable* inferences could serve as data for a logical theory.

Finally, the restricted case has given us a clear sense of where errors can occur within the process of supporting a logical theory if the current predictivist model is correct:

i) With the initial assumption that the putative proofs (arguments) first considered are of a similar form *F*.

ii) With the generalisation that all putative proofs (arguments) of this form *F* are valid.

iii) With the assumption that the data used is reliable (which, in this case, would be assuming that mathematicians are reliable at spotting invalid putative proofs).

iv) With the proposed postulates and laws in the theory itself, reflected in unsuccessful predictions.

Errors of the type i) and iii) arise not only at the initial stage of motivating the logical theory, but at the stage of testing its predictions. These predictions may seem a lot more successful than they ought to if the data is unreliable or misinterpreted.

The plausibility of the current predictivist model of logical methodology can be tested on the basis of these expectations. If indeed we do find logical theories criticised on the basis of these four errors, then this will provide prima facie support for the present account. We'll look at some examples of such criticisms in more detail in the following sections, and how the model ought to predict theories protect themselves from these concerns. However, it's worth noting here that, even restricting ourselves to logic's attempt to explain validity within informal proofs, the current account makes perfect sense of a historically significant challenge to classical logic—that of intuitionism.

Intuitionistic logicians criticised the classical logician for relying upon classical mathematics in order to justify their theory (see, for example, Brouwer [7]: Ch. 3; [8]; [9]). Instead, they proposed, we ought to be building our logic based upon (a certain flavour of) constructivist mathematics.[16] The current account can make perfect sense of this challenge. According to the intuitionist, the classical logician is using unreliable data—the judgements of classical mathematicians. Instead, they should be relying upon the judgements of constructivist mathematicians for which putative proofs are acceptable. Further, given that intuitionistic-minded constructivists reject the use of contraposition and double-negation elimination, ultimately the classical logician's reliance upon the judgement of non-constructivist mathematicians leads to a false logical theory. We take it that the model's ability to make sense of a significant dispute within the history of logic gives the account at least some initial plausibility.

We have begun with logic's challenge of regimenting mathematical proofs for a reason. A recognised strength of classical logic (at least, for most) is that it can provide an account of why these important steps within proofs are valid, and why others not generally accepted by mathematicians are invalid. Further, it is a strength that its predecessors did not possess. While we may now take this strength for granted,

---

[16]We say a *certain flavour* of constructivist mathematics for, as is well known, intuitionistic mathematics is not the only form of constructivist mathematics.

the ability of classical logic to provide an account of (in)valid steps within a putative mathematical proof is a significant theoretical achievement. It is no surprise that first-order classical logic is taught in discrete mathematics courses. A theory of logical epistemology should fully embrace this achievement. However, as we shall now see, it's clear that mathematical proofs cannot serve as the sole arbiters of a successful theory of logic.

4 General Theories of Validity

So far we have been assuming that logical theories are only concerned with explaining the validity of informal mathematical proofs. While this was useful in order to provide a simple outline for the proposed predictivist model of logical methodology, we well know that logical theories are not only concerned with the validity of informal mathematical proofs. They are concerned with validity *tout court*. Indeed, traditionally the *generality* of logical laws have been considered a constitutive part of what makes them *logical*:

> [The logical laws] are the most general laws, which prescribe universally the way in which one ought to think if one is to think at all. (Frege [20]: xv)

> Thought is in essentials the same everywhere: it is not true that there are different kinds of laws of thought to suit the different kinds of objects thought about. (Frege [20]: iii)

> [L]ogic is the science of the most general laws of truth. (Frege [21]: 128)

> [General logic] contains the absolutely necessary rules of thought without which there can be no employment whatsoever of the understanding. (Kant [28]: A52/B76)

So, logical theories attempt to account for validity *tout court*, not only in mathematics.[17] This ensures that relying upon mathematicians' judgements regarding informal proofs will be an insufficient source of data to justify a fully worked out logical theory. But, what further data can the logician appeal to in order to motivate, and ultimately justify, her theory? One available option is to extend our willingness to admit the judgements of mathematicians regarding the (un)acceptability of steps within putative informal proofs as reliable data for logical theories, to the

---

[17]It is an interesting question to what extent this quest for generality within logic is associated with the theoretical virtue of *scope* within the sciences (Kuhn [32]: 335). One may either interpret the proposed generality of logical theories as a heuristic, so that the more generally applicable the logic is, the better. In which case, the desired generality of logic would be a theoretical virtue, and related to the virtue of scope in the sciences. Or, as the property is sometimes interpreted within the philosophical community (see MacFarlane [35]), it is a pure stipulation that the correct logical theory must be generally applicable. In this case, any theory that failed as a general account of validity would not only be a deficient theory of logic, but *not* a logical theory. How exactly we ought to understand the proposed generality of logic, and its relationship to scope within the sciences, is an important topic, but unfortunately far beyond this paper's scope.

judgements of others regarding the (un)acceptability of normal vernacular arguments. This is certainly an option which has been entertained previously in the literature:

> [W]hat counts as data? It is clear enough what provides the data in the case of an empirical science: observation and experiment. What plays this role in logic? The answer, I take it, is our intuitions about the validity or otherwise of vernacular inferences. (Priest [45]: 9)

In what follows, we'll see how much sense can be made of this proposal within the current model, and to what extent it fits logical practice. Firstly, however, some points of clarification are required.

It is important to recognise that the judgements which serve as data for logical theories are not about argument *forms*, but about argument instances. Argument forms are schematic generalisations, a theoretical construct, postulated by logical theories as part of the process of explaining the validity of particular arguments. In comparison, particular arguments are not a product of logical theories, and judgements about *these* serve as data to inform the theory.

Secondly, while the judgements may ultimately be *about* the (in)validity of the argument, as it is the property of validity which an argument actually possesses (or, fails to), the content of an individual's judgement will *not* be directly about validity. There will be no "validity-talk" within the expression of the judgements. Rather, these judgements are the expressions of what the reasoner finds acceptable, or deems to follow. Validity is a technical term introduced by the logical community, with the attempt of discovering some substantive property of arguments which can be explained. In so introducing this concept, the community is hypothesising that there is some genuine phenomenon to be explained behind the everyday talk of some claims "following from" others. In other words, the *validity* of arguments is being treated as the phenomenon which logicians are attempting to explain. In contrast, the judgements of individuals over the correctness of arguments, or over whether some conclusion "follows from" some premises, are treated as data and taken to be *prima facie* reliable indicators *of* validity. To interpret the content of the judgements as judgements about the *validity* of arguments would be to mistake the data with the phenomenon.[18]

Further, just as mathematicians' judgements regarding putative informal proofs were only treated as viable data because such judgements were taken to be reliable guides to the (in)validity of steps within informal proofs, so any judgements on the acceptability or correctness of everyday vernacular arguments can only be treated as viable data if they are taken to be reliable guides to (in)validity. This, of course, does not require the logician to take these judgements as infallible. As with observational data, instances of these judgements can be justifiably found to be erroneous. It is only required that such judgements offer a reliable guide.

Here clearly is a pressing concern for the present account, as a potential disanalogy between the mathematician's judgements regarding informal proofs and general

---

[18]On the importance of distinguishing between a theory's target phenomenon and its data, see Bogen and Woodward ([4]).

judgements regarding which conclusions follow from which premises becomes apparent. While we may justifiably deem mathematicians to be experts in recognising when a putative proof is indeed a proof, why should we deem others to be experts, or otherwise reliable, in spotting when some proposition follows from others? After all, we are well aware from cognitive psychology of the unreliability of individuals' logical reasoning under certain conditions (see, for an introduction, Evans [19]). Thus, it seems either we must admit that the proposed data is unreliable, or we need to pre-identify certain agents as reliable judges of which propositions follow from others in particular arguments.

We highlight this problem here not because we have a ready solution to it, but because it is a challenge that will eventually need to be met—what justifies the presumption that the proposed data is reliable, when we have good reasons from empirical findings to believe it isn't? In what follows, we only aim to show what sense can be made of a logical methodology that treats such judgements as reliable, and further that logical practice does indeed suggest that logicians rely upon such data. As with the previously considered case of putative proofs within mathematics, general theories of validity seek to accommodate our judgements about actual arguments in natural language, ultimately providing postulates and laws in order to explain generalisations over valid forms of argument. It may be that logicians do indeed only take into account the judgements of perceived "reliable reasoners", whether this be logicians themselves, philosophers as a whole, or members of professions required to engage in detailed reasoning within their working lives, such as lawyers and scientists. This would certainly explain why logicians do not go in much for empirical studies. The question of whether this stance is justified or not is a conversation for elsewhere. Our claim here is only that a presumption of reliability for such judgements (from certain agents) is a prerequisite for the current methodology to make sense, and further that logicians do at least appear regularly to appeal to such judgements, whether these are personal judgements or those of a community.

Let's begin, as with we did in the previous section, with the initial motivating data for a theory of validity. This time, in the form of some natural language arguments, in which the conclusions will be judged to follow from the premises:

Argument 1

The UK is going to leave the EU without a deal. But the pound will crash if the UK leaves without a deal. So the pound will crash.

Argument 2

No doubt, we will make it to Rome on time. For we will make it unless the strike has started, and it hasn't.

Standard vernacular arguments are less regimented than those often found within informal proofs, and so our logician faces an extra complication when building a theory of validity from these cases. She must first engage in a process of regimentation in order to identify any possible shared structural features in the arguments. Already here we have a general underlying assumption of logic—that arguments can share

structural features, and that identifying these structural features can tell us something fruitful about validity. As would be the norm in an introductory logic class, we can expect our logician to hypothesise that the two arguments should be regimented so as to both exhibit an argument with two premises, one of which is a conditional:

> If the UK leaves without a deal, then the pound will crash.
> The UK leaves without a deal.
> _____
> The pound will crash.

> If the strike hasn't started, then we will make it to Rome on time.
> The strike hasn't started.
> _____
> We will make it to Rome on time.

Once the arguments are regimented, and the hypothesised shared structure is made more explicit, the proposed identified structure can then be used to produce an initial schematization to express a hypothesised generalisation about valid arguments:

Hypothesis 3

All arguments of the form
$$\text{If } \varphi, \text{ then } \psi$$
$$\frac{\varphi}{\psi}$$
are valid.

Again, all we have here is a hypothesised generalisation about which arguments are valid. We do not yet have an explanation of why arguments of this form are valid. For this, we need a theory. Let's take *Theory A* from above, with a disjunction and conjunction added:

Theory C

- **Definition 1**: Let $\varphi \rightarrow \psi$ be Boolean material implication.
- **Definition 2**: Let $\neg\varphi$ be Boolean negation.
- **Definition 3**: Let $\varphi \vee \psi$ be Boolean disjunction.
- **Definition 4**: Let $\varphi \wedge \psi$ be Boolean conjunction.
- **Representation Rule 1**: $\ulcorner \text{if } \varphi \text{ then } \psi \urcorner = \ulcorner \varphi \rightarrow \psi \urcorner$.
- **Representation Rule 2**: $\ulcorner \text{not } \varphi \urcorner = \ulcorner \neg\varphi \urcorner$.
- **Representation Rule 3**: $\ulcorner \varphi \text{ or } \psi \urcorner = \ulcorner \varphi \vee \psi \urcorner$.
- **Representation Rule 4**: $\ulcorner \varphi \text{ and } \psi \urcorner = \ulcorner \varphi \wedge \psi \urcorner$.
- **Law 1**: For every valuation, all sentences are either true or false, and not both.
- **Law 2**: An argument is valid iff, for every valuation v, if every premise is true in v, the conclusion is true in v.

The same points hold as previously. While this theory accommodates the hypothesised generalisation—it "fits the data" in the form of our judgements over Arguments 1 and 2—and provides a potential explanation of its putative truth, so do many others.

We need reasons to prefer it over competitors. So, consequences are drawn from the theory. Here are some examples:

- **Consequence 5**:

  All arguments of the form

  $$\frac{\neg\psi \qquad \varphi \to \psi}{\neg\varphi}$$

  are valid.

- **Consequence 6**:

  All arguments of the form

  $$\frac{(\varphi \to \psi) \to \varphi}{\varphi}$$

  are valid.

- **Consequence 7**:

  All arguments of the form

  $$\frac{\psi}{\varphi \to \psi}$$

  are valid.

- **Consequence 8**:

  Not all arguments of the form

  $$\frac{\varphi \qquad \psi \to \varphi}{\psi}$$

  are valid.

- **Consequence 9**:

  All arguments of the form

  $$\frac{(\varphi \wedge \psi) \to \chi}{(\varphi \to \chi) \vee (\psi \to \chi)}$$

  are valid.

- **Consequence 10**:

  All arguments of the form

  $$\frac{\varphi \to \psi}{\varphi \to (\psi \wedge \varphi)}$$

  are valid.

As before, to test her theory, the logician must operationalise the consequences as predictions. These are in turn put to the test, by looking to actual arguments of

(putatively) the same form. If our judgements regarding these instances cohere with the predictions, then the theory is supported by this new data. If instead, however, we find examples that clash with a prediction, then the logician has some answers to give.

Some of these predictions will look initially promising, for example one based on Consequence 5:

- **Prediction 3**:

    All arguments of the form
    $$\frac{\text{Not } \psi \qquad \text{If } \varphi \text{ then } \psi}{\text{Not } \varphi}$$
    are judged acceptable.

Other predictions will be more surprising, perhaps because they speak about argument forms rarely found in natural language arguments (e.g. Consequence 6, Peirce's Rule):

- **Prediction 4**:

    All arguments of the form
    $$\frac{\text{If, if } \varphi \text{ then } \psi, \text{ then } \varphi}{\varphi}$$
    are judged acceptable.

In such cases we might not have strong judgements about the acceptability of its instances. As such, it's a surprising prediction of the theory, but one that's not so easily testable, and thus unlikely to generate further evidence in support of the theory.

Yet, further predictions might appear problematic because they run into putative counterexamples. This is arguably how some non-classical logicians react to the predictions based on Consequence 7 (the positive paradox of the material implication), Consequence 9 and 10:

- **Prediction 5**:

    All arguments of the form
    $$\frac{\psi}{\text{If } \varphi \text{ then } \psi}$$
    are judged acceptable.

- **Prediction 6**:

    All arguments of the form
    $$\frac{\text{If } \varphi \text{ and } \psi, \text{ then } \chi}{\text{It is the case that either if } \varphi \text{ then } \chi, \text{ or that if } \psi \text{ then } \chi}$$
    are judged acceptable.

- **Prediction 7**:

  > All arguments of the form
  >
  > $$\frac{\text{If } \varphi \text{ then } \psi}{\text{If } \varphi \text{ then } \psi \text{ and } \varphi}$$
  >
  > are judged acceptable.

Take, for example, a putative counterexample to Prediction 5, similar to those which motivated relevant logicians:

> $$\frac{\text{John's going skiing this weekend.}}{\text{If John breaks his legs, he's going skiing this weekend}}$$

Similarly, Prediction 6 and Prediction 7 have been met with their own putative counterexamples:[19]

> If you close switch x and switch y the light will go on.
> ―――――――――――――――――――――――――――――
> It is the case either that if you close switch x the light will go on,
> or that if you close switch y the light will go on.

> If a piece of wood makes one bed, it makes four chairs.
> ――――――――――――――――――――――――――――
> If a piece of wood makes one bed, it makes four chairs
> and makes one bed

Faced with such alleged counterexamples, the onus is on the advocate of *Theory C* to provide a response; even if, ultimately, this response is merely a denial that the cases constitute genuine counterexamples or should be given any importance. After all, again, our account of logical methodology here is not committed to a form of naïve falsificationism, such that any unsuccessful prediction is a significant problem for a theory. All that is required is that such unsuccessful predictions prompt some kind of response from the advocate of the theory. What kind of response though, exactly? According to the present model, we should expect the logician to either:

a) Deny that the argument is an instance of the form given in the prediction. This would require either reformulating one of the theory's representation rules, or an account of how the *real* form of the argument is somehow hidden, and what it is exactly.

b) Propose that although the argument is indeed an instance of the form given in the prediction, the argument is acceptable, contrary to the judgements of others. This would be simply to admit the unreliability of the data appealed to in the counterexample.

c) Admit that the case does indeed seem troublesome, but that given it is only one example, the correct course of action would be simply to bracket it off as an anomaly yet to be explained. This response makes most sense in those cases in which the theory has so far been predictively successful, and no clear solution to the anomalous case is forthcoming. After all, it would be irrational to revise an otherwise successful theory just because of a single anomaly.

---

[19]See Priest ([43]: 14) and Zardini [70], respectively, for variants of these counterexamples.

d) Admit that the case is indeed a counterexample to the theory, and thus at least part of the theory needs amending. For example, one could reject validity as solely truth-preservation (Law 2), or reject bivalence (Law 1). There is, of course, no decision procedure for determining which parts of the theory need to go. There is always an initial stage of trial and error when anomalies are encountered.

It's clear that some of these options are found within the literature when logicians are faced with potential counterexamples. For example, motivated by putative counterexamples to Prediction 3 such as that above, relevant logicians have taken option (d), rejecting Law 2 by proposing that validity must be understood in terms of some more intimate connection between the premises and conclusion of an argument than truth-preservation (Anderson and Belnap [1]). Similarly, in response to his own putative counterexample to modus tollens:

> If the marble is big, then it's likely red.
> The marble is not likely red.
> _____
> The marble is not big.

Seth Yalcin [69] rejects Law 2 and provides two alternative logics that define consequence in terms of the preservation of properties of information states. Further, Yalcin's ([69], 1003-8) discussion highlights how logicians may attempt to explain away the putative counterexample, by showing it to *not* be an instance of modus tollens, contrary to appearances (option (a) above). In the current case, this could be achieved by interpreting the probability operator in the first premise to have the widest possible scope, thereby operating over the whole conditional, or by proposing that "likely" equivocates across the premises, so that the second premise is not the negation of the conditional's consequent.

If logical predictivism is correct, then these are the responses to putative counterexamples that we would expect to find in the literature. Consequently, logical predictivism can again be tested against logical practice in order to evaluate its viability. If we fail to find just such replies to putative counterexamples within the literature, this should count against the model of logical methodology proposed here. We leave the question of to what extent each of these types of reply are found, and whether they exhaust how logicians do indeed respond to troublesome cases, to elsewhere. Testing this consequence of the model will require looking at how logicians have replied to famous putative counterexamples, such as McGee's [38] counterexample to modus ponens, and Kolodny and MacFarlane's [30] counterexample to modus tollens, in far more detail than we have space for here. Such a testing of the predictivist model against logical practice should be a task for future work. What is important, however, is that the logical predictivist model does open itself up to the tribunal of actual logical practice, as all accounts of logical methodology should (Martin [36]).

While a detailed consideration of these putative counterexamples, and replies to them, must wait for another occasion, it's clear however that the mere existence of these famous putative counterexamples substantiates the claim that judgements over whether a particular argument is acceptable or not *are* used as pieces of evidence in arguing for or against a logical theory. The whole reason that these cases are so

famous is that our (putative) judgement that the arguments are unacceptable, that the conclusion does not follow from the premises, is taken to be at least *prima facie* evidence against logics that are (putatively) committed to the argument being valid. In this regard, then, the current model fits the practice of logicians well.

So far we have been speaking as though the only types of evidence that one can have for or against a logical theory are direct judgements about the acceptability of particular arguments, which can either support or falsify the theory's predictions. However, this is not the case. In fact, the types of evidence recognised within the contemporary literature are far more varied than this, and it's important that our model reflects this. In the next section we consider the wider types of evidence appealed to by logicians when engaged in theory choice.

5 Indirect Evidence

Not all logical evidence takes the form of judgements about instances of a particular argument form. There are other, more *indirect* means to support the (in)validity of certain argument forms, and subsequently to justify revising one's logical theory. Three such types of indirect evidence are prevalent in the literature: i) *bad company*, ii) post hoc rejections, and iii) clashes with other theoretical commitments.

5.1 Bad Company

Sometimes we don't possess a counterexample to a form of argument *F*. Instead, we have good reason to reject it because admitting the validity of all its instances would require admitting the validity of all the instances of *another* form of argument *F′* which we do have independent reasons to reject the validity of, due to counterexamples. In this case, we reject the validity of *F* on account of its keeping *bad company* with *F′*.

More often than not, a form of argument does not in itself keep bad company with another form. Rather, in combination with other forms, it requires one to admit the troublesome form into one's theory. In this case, one must choose where to lay the blame.

A famous example of *bad company* comes from relevant logic's traditional rejection of the disjunctive syllogism as an axiom or rule. However, the argument form itself,

$$\frac{\varphi \lor \psi \quad \neg\varphi}{\psi}$$

is typically rejected neither because there are clear counterexamples to it (in the form of real natural-language arguments), nor because it directly contradicts some fundamental principle of relevant theories, such as variable sharing. Indeed, although rejected, it is often recognised that a relevant analogue of the disjunctive syllogism is needed in order to capture moves deemed acceptable by mathematicians within proofs (Burgess [13]). Instead, the disjunctive syllogism is rejected as valid because,

in combination with addition, it allows for the validity of explosion, as demonstrated by the infamous Lewis and Langford ([33]: 252) proof:

$$\frac{\dfrac{\varphi}{\varphi \lor \psi} \ (\lor I) \qquad \neg\varphi}{\psi} \ (DS)$$

Now, given that explosion itself,

$$\frac{\varphi \quad \neg\varphi}{\psi}$$

does according to the relevant logician have clear counterexamples, and the disjunctive syllogism and addition are sufficient to ensure that every instance of explosion is valid, one of these two must be rejected. Which one, of course, is a decision for the logician, and is likely to be dependent on the *cost* of denying the validity of all instances of a particular form. In the current case, the cost of denying the validity of all instances of the disjunctive syllogism is considered less damaging theoretically than doing the same for addition.

Consequently, the possibility of *bad company* ensures that in virtue of having direct evidence against the validity of argument form *F*, due to judgements of unacceptability about instances of *F*, we can also have good reasons to reject other argument forms which require us to accept the validity of all instances of *F*. In such cases, the logician is then obviously required to make the necessary adjustments within her theory in order to invalidate the argument form.

Importantly, in practice, bad company arguments will rarely provide evidence for a particular theory over competitors. This is because there will be numerous theoretical avenues through which one can ensure that not all instances of *F* are valid. The arguments only serve to remove certain candidates from the table—namely, those that commit the logician to the validity of *F'* via *F*.[20] In this sense, bad company arguments serve a similar function as *internal consistency constraints* on empirical theories (see McMullin [39]). In order to find discriminating support for the remaining candidates, new consequences must be drawn from them and predictions tested.

5.2 Post Hoc Rejections

Post hoc rejections arise when a particular form of argument clashes with fundamental elements of the theory, and thus suitable adjustments must be made to ensure that the form is not sanctioned by the theory. Thus, in such cases, the invalidity of a form of argument *F* is not determined on the basis of judgements regarding putative instances of the form, but because *F* does not meet certain requirements laid down by the theory's laws. Again, a suitable example of this can be found in relevant logic literature. While the axiom,

(**A**) $\neg(\varphi \to \varphi) \to (\psi \to \psi)$

---

[20]Of course, those theories which deem all instances of *F'* valid need not be off the table forever. This simply depends on whether the community continues to consider the evidence against *F'* to be definitive or not.

is not included within relevant logics, this not generally because logicians have in mind counterexamples to it. Instead, it contravenes a law of such logics, that an argument is only valid if it adheres to variable sharing,[21] which itself was introduced to *explain* why certain other relevantly invalid argument forms, such as explosion, which we *do* putatively have direct counterexamples against, are indeed invalid.

Such post-hoc rejections do not provide additional evidence for a theory, as they are motivated solely by postulates and laws already included within the theory. However, they do give the logician who is already committed to including these postulates and laws within her theory a reason to ensure the validity of such arguments are excluded by the theory, if indeed they do clash with certain postulates or laws.[22] Of course, in order to be justified in being committed to a theory containing these postulates and laws in the first place, the theory must have shown itself to be predictively successful.

There are also bad company cases for post hoc rejections, where an argument form F is not rejected because it directly clashes with a (perceived) well-confirmed law L within a theory, but because, in combination with other argument forms, F entails an argument form F' which does clash with L. Thus, in order to block the inclusion of F', F (or, another form of argument) must also be excluded. For an example, we can look again to relevant logicians. While the *mingle axiom*,

$$(\mathbf{M})\ \varphi \to (\varphi \to \varphi)$$

adheres to the relevant logician's variable sharing requirement, by adding the axiom to the logic R, one produces the logic RM which includes ($\mathbf{A}$) as a theorem. Given that, as we have already noted, ($\mathbf{A}$) contravenes the variable sharing property, this gives the relevant logician good reason not to endorse a theory which includes ($\mathbf{M}$), in combination with certain other forms of argument, much to the surprise of Meyer [40].

Again, the same points as before hold. Such arguments do not provide any new evidence for a theory. They merely map out the viable theories for those who are already committed to a certain combination of elements of a theory (perhaps for very good reasons). These perceived viable theories must still then be tested further for their predictive success.

5.3 Clashes with Other Theoretical Commitments

So far, both types of indirect evidence outlined provide no new positive evidence for a logical theory. They merely show which candidates are viable or not given a certain set of (potentially well-confirmed) commitments. Our final type of indirect evidence for a logical theory, however, can provide further positive evidence for a theory. This is because it requires making predictions about the compatibility of the theory with our wider theoretical commitments. This motivation for theory revision is closely

---

[21] An argument pair $\langle \Sigma, \beta \rangle$ possesses the *variable sharing property* if and only if for every $\alpha \in \Sigma$, α and β have some propositional variable in common (Belnap [3]: 144).

[22] Again, this constraint on logical theory choice is related to the requirement within the empirical sciences that theories ought to be internally consistent.

related to what is often called 'external consistency' in discussions of theoretical virtues in the sciences (Kuhn [32]: 321-2).

We are well aware of examples from the history of logic where a reevaluation of our logical theory has been motivated by our wider theoretical commitments, such as a particular scientific theory [46], a particular theory of meaning [18], the existence of vague predicates [63] and our theory of truth [42]. To see how these wider commitments could justifiably lead us to revising parts of our logical theory, we'll concentrate here on an example based upon truth and the self-referential paradoxes.

Take as our starting point a classical theory of validity. Further, assume that so far we've found significant reason to accept the theory, due to its success in predicting those informal proofs/arguments we judge acceptable and those we don't. In so tentatively accepting the theory, we place upon it the additional constraint that we expect it to cohere with our other theoretical commitments.

Further, let's assume that we have been convinced that we ought to take on two further commitments, independent of our motivation for accepting classical logic: the transparency of the truth predicate, and the semantic closure of natural languages. While commitment to the latter may come directly from what linguists tell us about natural languages, based upon empirical evidence, we may take on the former commitment because it allows us to make blind belief ascriptions to others [31].

Now, for a while we may be perfectly happy that all three of our commitments—classical logic, a transparent truth predicate, and semantic closure—are completely compatible with one another. However, then one day a particularly clever associate [17] points out that given semantic closure, we can express tricky self-referential sentences, such as,

(*C*) If *C* is true, then 0=1

and further, that given our other commitments to a transparent truth predicate and classical logic, *C* allows us to derive the truth of 0=1 (or any falsehood for that matter).

Given that we have very good reasons to reject that 0=1, and further realise that many sentences similar to (*C*) can be constructed committing us to claims we don't wish to be, we conclude that we somehow need to block the conclusion of these tricky arguments. This requires either rejecting the meaningfulness of (*C*) and other similar sentences, which subsequently requires restricting the semantic closure of English, denying the transparency of the truth predicate, or rejecting a classical inference rule. In other words, (*C*) has demonstrated that our wider commitments are incompatible with one another.

Now, let's assume that our reasons for accepting the semantic closure of English and the transparency of the truth predicate are so strong that blocking the argument through these means simply isn't viable. In this case, we're left with no other option than to revise our logical theory.[23] Particularly, I need to revise my theory to one in which the argument is blocked. Here is such a theory:

---

[23]This, in effect, is the dialetheist's argument for why it is classical logic that must go, rather than semantic closure or the unrestricted T-schema, when faced with the self-referential paradoxes. See Priest [42].

Theory D

- **Definition 1**: Let $\varphi \to \psi$ be **LP** material implication.
- **Definition 2**: Let $\neg\varphi$ be **LP** negation.
- **Definition 3**: Let $\varphi \vee \psi$ be **LP** disjunction.
- **Definition 4**: Let $\varphi \wedge \psi$ be **LP** conjunction.
- **Representation Rule 1**: $\ulcorner$ if φ then ψ $\urcorner = \ulcorner\varphi \to \psi\urcorner$.
- **Representation Rule 2**: $\ulcorner$ not $\varphi\urcorner = \ulcorner\neg\varphi\urcorner$.
- **Representation Rule 3**: $\ulcorner\varphi$ or $\psi\urcorner = \ulcorner\varphi \vee \psi\urcorner$.
- **Representation Rule 4**: $\ulcorner\varphi$ and $\psi\urcorner = \ulcorner\varphi \wedge \psi\urcorner$.
- **Law 1**: For every valuation, all sentences are either true or false, or both.
- **Law 2**: An argument is valid iff, for every valuation v, if every premise is true in v, the conclusion is true in v.

The theory blocks the absurd consequences that follow from the Curry sentence because the material conditional does not detach. In this sense, then, *Theory D* "fits the data" by allowing us to keep our commitments to semantic closure and the transparency of the truth predicate, while not committing us to the absurd consequence which did follow from (*C*) under classical logic.

In fitting the data, however, *Theory D* is not unique. There are multiple other theories, such as paracomplete and substructural theories, which do the same. Again, then, the theory must be tested against competitors via the success of its predictions (including which proof steps/arguments are to be judged acceptable).[24]

6 Upshots

According to the *predictivist* account of logical methodology presented here, logical theories are ultimately chosen for their ability to make successful predictions. Such predictions include claims about which informal proofs mathematicians will find acceptable, and which arguments certain "reliable reasoners" will judge acceptable. Additionally, as we argued in Section 5, logical theories are valued for both their *internal consistency*, and their compatibility with our other theoretical commitments (often known as their *external consistency*).

Further, we've highlighted at significant points how this account accords well with the actual practice of logicians. Thus, while undoubtedly a departure from traditional accounts of logical methodology, we have urged its fit with what logicians do when engaged in justifying their theories. To this extent, the theory does not recommend a revision of logical practice, but aims to respect it.

---

[24]Of course, just because a theory is deemed to be compatible with our further theoretical commitments, and thus successful in this regard, does not ensure it is predictively successful when it comes to those proof steps/arguments which are judged acceptable. For example, while dialetheists (Priest [42]) have argued that glutty paraconsistent logics such as *Theory D* respect our commitments to semantic closure and the unrestricted T-schema, it's been argued that they fail to successfully recognise as valid important inference forms used within mathematical proofs, such as the disjunctive syllogism (Tennant [60]).

However, the case for predictivism is at present undoubtedly incomplete. As with any account of logical methodology, logical predictivism will ultimately prove successful or fail in light of how well it fits the wide range of logical practice (Martin [36]). Thus, the theory must now be strenuously tested in detail against paradigm instances of logical practice, and evaluated on the basis of how the theory can make sense of this practice. In highlighting on several occasions what we would expect to occur within practice if predictivism is correct, we hope we have made the evaluation of the theory in light of practice that bit easier for others.

Leaving aside for now the job of further testing predictivism's adequacy against logical practice, it will instructive to end by highlighting certain important consequences of the theory, if it does indeed ultimately turn out to fit the practice well.

Upshot 1: Explanatory Power Predictivism differs from abductivism in stressing the importance of theories not only "fitting" or "capturing" the existent data, but making predictions to be tested by further data. Additionally, while we have drawn attention to the apparent desire for both internally and externally consistent theories within the logical community, we have resisted the temptation, unlike abductivism, to appeal to a whole host of purported theoretical virtues, such as simplicity and deductive strength, without a clear sense of the exact role they play within logical theory choice. There is, however, one further theoretical virtue which predictivism explicitly admits but that we have not provided an account of—explanatory power.

As we have seen, predictivism admits that logical theories are in the business of explaining validity. Further, predictive success constitutes part of the story of what it means for a logical theory to successfully explain validity. It cannot be the whole story, however, if only because one can have two predictively identical theories that have varying levels of explanatory power. A clear example here, as mentioned above, would be two classical theories, one of which includes a model-theoretic account of validity, and another a proof-theoretic account. Both theories are predictively equivalent, yet logicians find reasons to prefer one over the other. On the current account, this must be because one of these explanations of validity is deemed to be more fruitful than the other.

The predictivist is under a burden, therefore, to provide a suitable account of logical explanation. While we have made clear how such explanations are accommodated overall within the predictivist model, with the components of the logical theory combining to give an account of why particular arguments are valid, we are still owed a detailed story of logical explanations. For example, we require an answer to what makes an explanation "fruitful" or "insightful" according to the logical community, and therefore what needs to be included within a explanatorily powerful logical theory.

That is a task for another day. Let us just highlight here that practice lends credence to the idea that logical theories should have some kind of explanatory power. Copeland [16], for example, famously criticised the Routley-Meyer [49, 50] star semantics for being inadequate, due to its formal semantics lacking a suitable philosophically informative interpretation which made transparent *why* certain forms of argument were valid and others invalid. Such criticisms within the literature can only

be made sense of by recognising that explanatory power (somehow conceived) plays a role within logical theory choice. How to understand such explanatory power, and what role it plays exactly in choosing between theories, is a topic for elsewhere.

Upshot 2: Re-evaluating Anti-Exceptionalism The current proposal takes seriously the idea that logical theory choice is akin to that within the sciences, with both scientific and logical theories engaged in a process of providing explanations for a given phenomenon, and (at least partially) demonstrating their worth through successful predictions. In this sense, logical predictivism supports methodological AEL. However, the position deviates from the anti-exceptionalist's overall commitments as they are sometimes put in the literature:

> Logic isn't special. Its theories are continuous with science; its method continuous with scientific method. Logic isn't a priori, nor are its truths analytic truths. Logical theories are revisable, and if they are revised, they are revised on the same grounds as scientific theories. These are the tenets of anti-exceptionalism about logical theories. (Hjortland [25]: 632)

According to *logical predictivism*, while the mechanisms of theory choice for logical theories are no different from those of the sciences, this does not mean that logical theories are revised *on the same grounds as* scientific theories. After all, the *forms of evidence* one's theory appeals to may be very distinctive, even though a theory's success relative to this available evidence may still be explained in terms of predictive success. Particularly, in suggesting that logical theories appeal to *judgements* regarding arguments, the current proposal opens up the possibility that *a priori* evidence does indeed play a role within logical theory choice, contrary to the quote above. Whether *a priori* evidence plays a role within theory choice in the sciences is a moot point, but the current proposal at least diverges from the anti-exceptionalist creed as it is often presented—that logical evidence is *not a priori*.

The upshot of logical predictivism for AEL, therefore, is that not all of its claims about logic need be true together. Some may well turn out to be more plausible than others. Which are, exactly, is a matter for future work, and dependent upon logical practice.

Upshot 3: The Background Logic Problem As was highlighted in Section 3, in order for logical theories to be tested, according to logical predictivism, *consequences* must first be drawn from its postulates. Drawing such consequences, however, requires relying upon certain rules of inference, and here a potentially significant problem for the anti-exceptionalist arises. For, to accurately test a theory according to logical predictivism, we must have good reason to be able to rely upon the rules of inference we use to draw out the theory's consequences. But, which rules of inference can we rely upon?

In short, we have the choice when testing a theory *T* either to rely upon only those rules of inference sanctioned by *T*, or to allow some rules of inference rejected as invalid by *T*. Both options, however, are paved with problems. Firstly, relying solely upon those rules of inference which the theory sanctions leaves the theory open to accusations of begging the question from advocates of competing theories. For, if the

theory is found to be well-confirmed through successful predictions, on the back of drawing these consequences, the opponent may well rightly be concerned that the predictions were the result of background assumptions (in the form of certain rules of inference) which theory T did not have a right to presume. Consequently, there is no reason to think that theory T is actually supported, because we have no non-question begging reasons to rely upon the rules of inference through which the predictions were drawn. Yet, requiring T to rely only upon those rules of inference not sanctioned by the theory will not be viable either. For, if one's justification for endorsing a theory comes from its successful predictions, but those predictions were only formed on the basis of relying upon forms of inference admitted to be invalid by the theory, then the theory undercuts its own possible justification. Therefore, neither relying upon those rules sanctioned by the theory T in order to test it, nor those rules rejected by the theory, seem viable, and consequently there is no way to reliably test a logical theory.

This is known as the *background logic* (or, *centrality problem*) [57, 66, 68], and impacts *any* account of logical methodology which proposes that we come to be justified in believing a logical theory by appealing to some putative non-immediate evidence, including logical abductivism [36, 66]. For, under such a methodology, we will always need to appeal to rules of inference in order to substantiate the claim that the available evidence is (in)consistent with the relevant logical theory.

If logical predictivism is indeed the best candidate for an anti-exceptionalist model of logical methodology, then the anti-exceptionalist will need a solution to the background logic problem, contrary to what has been suggested recently in the literature [41]. A failure to solve the problem will be tantamount to admitting that logical theories cannot be reliably tested. However, in this respect logical predictivism is at no disadvantage with comparison to logical abductivism, as both require solutions to the problem. The anti-exceptionalist will need a solution to the problem either way.[25]

Upshot 4: Idealization and natural language According to logical predictivism, logical theories offer generalizations about which natural language arguments are valid. However, some authors have warned that natural languages—and our judgements about their validity—are too irregular to allow for the sweeping generalizations contained in most logical theories (cf. Glanzberg [23]). Strawson ([59]: 344) famously summed up the view with the claim that "ordinary language has no exact logic", and similar concerns underlie Russell's [53] recent discussion of *logical nihilism*, the view according to which there are no universally valid forms of argument.

The issue, in a nutshell, is that even the most uncontroversial argument forms have purported counterexamples. For example, the commutativity of the conjunction has been met with the objection that 'and' sometimes does not commute in English language arguments: "I got up in the morning and I brushed my teeth. Therefore, I brushed my teeth and I got up in the morning". The argument presumably won't strike many as acceptable, and for good reasons. The classicist will likely object, however, that their theory, with its commutative conjunction, isn't intended to

---

[25]For a recent sketch of the possible solutions to the background logic problem available to the anti-exceptionalist, see Martin and Hjortland ([37]).

accommodate arguments with tensed propositions. Similarly, some classicists will insist that their theory isn't intended for arguments with vague expressions, indexicals, or self-reference.

In short, logicians are often willing to sacrifice universality in order to formulate true generalizations about restricted forms of arguments. As a result, logical theories *idealize* away from natural languages in a number of respects, just like scientific theories—be it physics or economics—idealize away from certain phenomena. That logical theories involve idealization has been pointed out by a number of other authors (such as Cook [15] and Shapiro ([56, 58]: 46-54)). Particularly, Glanzberg ([23]) has warned against supposing that natural language *has an embedded logic*, while nonetheless maintaining that logical theories can capture theoretically important properties of natural language arguments by idealizing away other (perceived) unimportant properties. This perspective of the relationship between natural languages and logic is fully compatible with the predictivist outlook, but suggests that even general theories of validity typically need to restrict their scope in order to produce generalizations that deliver correct predictions. How logical theories go about achieving this is a question for another occasion.

There is then still much work to do, both in terms of evaluating logical predictivism in accordance with logical practice, and demonstrating how predicitivism can answer the challenge posed by the background logic problem. Further, more ultimately needs to be said about how predictively equivalent logics can be theoretically differentiated on the basis of their explanatory power, and how logicians idealise from natural language arguments to ensure their theories are not consistently falsified. However, of course, this paper never intended to be the final word on the viability of logical predictivism. Rather, it's goal was to present a novel theory of logical methodology which properly respected both the claims of methodological AEL and important features of logical practice.

7 Conclusion

In this paper we have presented what we take to be a novel theory of logical methodology, which puts predictions centre stage. If shown to be an accurate picture of logical practice, methodological AEL will have been vindicated. Just as with the empirical sciences, logical theories are engaged in a process of explanation and prediction. We have already gestured at how the proposed *predictivist* account fits the practice of logicians. However, what is now needed in future work is a full evaluation of the theory on the basis of detailed consideration of actual logical practice.

References

1. Anderson, A.R., & Belnap, N.D. (1975). *Entailment: The Logic of Relevance and Necessity volume 1*. Princeton: Princeton University Press.
2. Ayer, A.J. (1936). *Language truth and logic*. New York: Dover.
3. Belnap, N.D. (1960). Entailment and relevance. *The Journal of Symbolic Logic*, *25*(2), 144–146.
4. Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, *97*(3), 303–352.
5. Boghossian, P. (2000). In Paul Boghossian, C.P. (Ed.) *Knowledge of logic*, (pp. 229–254). Oxford: Clarendon Press.
6. BonJour, L. (1998). *In defense of pure reason: a rationalist account of a priori justification*. Cambridge: Cambridge University Press.
7. Brouwer, L.E.J. (1907). Over de grondslagen der wiskunde. PhD thesis, Amsterdam. Extracts reprinted in Brouwer (1975), pp. 11-101.
8. Brouwer, L.E.J. (1908). De Onbetrouwbaarheid der logische principes. *Tijdschrift voor wijsbegeerte*, *2*, 152–158. Reprinted in Brouwer (1975), pp. 107-111.
9. Brouwer, L.E.J. (1948). Consciousness, philosophy and mathematics. In Beth, E.W., Pos, H.J., Hollak, J.H.A., Beth, E.W., et al. (Eds.) *Proceedings of the 10th international congress of philosophy* (pp. 1235–1249). Amsterdam: North Holland. North-Holland, Amsterdam. Reprinted in Brouwer (1975), pp. 480–494.
10. Brouwer, L.E.J. (1975). *Collected works: philosophy and foundation of mathematics* Vol. 1. Amsterdam: North-Holland. Edited by A. Heyting.
11. Brush, S.G. (1989). Prediction and theory evaluation: the case of light bending. *Science*, *246*(4934), 1124–1129.
12. Brush, S.G. (1994). Dynamics of theory change: The role of predictions. In *PSA: proceedings of the biennial meeting of the philosophy of science association* (pp. 133–145). Philosophy of Science Association.
13. Burgess, J.P. (1983). Common sense and "relevance". *Notre Dame Journal of Formal Logic*, *24*(1), 41–53.
14. Carroll, J. (1994). *Laws of nature*. Cambridge: Cambridge University Press.
15. Cook, R. (2002). Vagueness and mathematical precision. *Mind*, *111*(442), 225–48.
16. Copeland, B.J. (1979). On when a semantics is not a semantics: some reasons for disliking the routley-meyer semantics for relevance logic. *Journal of Philosophical Logic*, *8*(1), 399–413.
17. Curry, H. (1942). The inconsistency of certain formal logics. *Journal of Symbolic Logic*, *7*(3), 115–117.
18. Dummett, M. (2000). *Elements of intuitionism*, 2nd edn. Oxford: Oxford University Press.
19. Evans, J. (2013). *The psychology of deductive reasoning psychology revivals*. London & New York: Psychology Press.
20. Frege, G. (1959). *The foundations of arithmetic*. Oxford: Basil Blackwell. Translated by J.L. Austin.
21. Frege, G. (1979). *Posthumous Writings*. Oxford: Blackwell. Translated by P. Long and R. White, and edited by H. Hermes, F. Kambartel and F. Kaulbach.

22. Gentzen, G. (1969). Untersuchungen Über das logische Schliessen, (Vol. 39. Translated in The Collected Papers of Gerhard Gentzen, ed. by M.E. Szabo, Nort Holland.
23. Glanzberg, M. (2015). Logical consequence and natural language. In Caret, C., & Hjortland, O. (Eds.) *Foundations of logical consequence* (pp. 71–120). Oxford: Oxford University Press.
24. Hitchcock, C., & Sober, E. (2004). Prediction verses accommodation and the risk of overfitting. *The British Journal for the Philosophy of Science*, *55*(1), 1–34.
25. Hjortland, O.T. (2017). Anti-exceptionalism about logic. *Philosophical Studies*, *174*(3), 631–658.
26. Hunt, J.C. (2012). On ad hoc hypotheses. *Philosophy of Science*, *79*(1), 1–14.
27. Jeshion, R. (2001). Frege's notions of self-evidence. *Mind*, *110*(440), 937–976.
28. Kant, I. (1998). *Critique of pure reason*. Cambridge: Cambridge University Press. Translated and edited by Paul Guyer & Allen W. Wood.
29. Kitcher, P. (1979). Frege's epistemology. *The Philosophical Review*, *88*(2), 235–262.
30. Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts. *The Journal of Philosophy*, *107*(3), 115–143.
31. Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, *72*(19), 690–716.
32. Kuhn, T. (1977). *The essential tension university of chicago press*. Chicago: IL.
33. Lewis, C.L., & Langford, C.H. (1932). *Symbolic logic*. London: Dover Publications.
34. Lipton, P. (1991). *Inference to the best explanation*. London: Routledge.
35. MacFarlane, J. (2000). *What Does it Mean to Say that Logic is Formal? PhD thesis*. Pittsburgh: University of Pittsburgh.
36. Martin, B. (2020). Identifying logical evidence. Synthese. Online first: https://doi.org/10.1007/s11229-020-02618-y.
37. Martin, B., & Hjortland, O. (202X). Evidence in logic. In *Routledge handbook of philosophy of evidence*. London: Routledge.
38. McGee, V. (1985). A counterexample to modus ponens. *Journal of Philosophy*, *82*(9), 462–471.
39. McMullin, E. (2008). The virtues of a good theory. In Psillos, S., & Curd, M. (Eds.) *The routledge companion to philosophy of science* (pp. 498–508). London: Routledge.
40. Meyer, R.K. (1971). R-mingle and relevant disjunction. *Journal of Symbolic Logic*, *36*(2), 366.
41. Payette, G., & Wyatt, N. (2018). How do logics explain? *Australasian Journal of Philosophy*, *96*(1), 157–167.
42. Priest, G. (2006). *In contradiction*, 2nd edn. Oxford: Oxford University Press.
43. Priest, G. (2008). *An introduction to Non-Classical logic*, 2nd edn. Cambridge: Cambridge University Press.
44. Priest, G. (2014). Revising logic. In Rush, P. (Ed.) *The Metaphysics of logic, chapter 12* (pp. 211–223). Cambridge: CUP.
45. Priest, G. (2016). Logical disputes and the a priori. *Logique et Analyse*, *59*(236), 347–66.
46. Putnam, H. (1969). Is logic empirical?. In Cohen, R.S., & Wartofsky, M.W. (Eds.) *Boston studies in the philosophy of science: proceedings of the Boston colloquium for the philosophy of science 1966/1968* (pp. 216–241). Dordrecht: Springer.
47. Quine, W. (1981). What price bivalence? *The Journal of Philosophy*, *78*(2), 90–95.
48. Quine, W.V. (1951). Two dogmas of empiricism. *Philosophical Review*, *60*(1), 20–43.
49. Routley, R., & Meyer, R.K. (1972a). The semantics of entailment ii. *Journal of Philosophical Logic*, *1*(1), 53–73.
50. Routley, R., & Meyer, R.K. (1972b). The semantics of entailment iii. *Journal of Philosophical Logic*, *1*(2), 192–208.
51. Russell, G. (2015). The justification of the basic laws of logic. *Journal of Philosophical Logic*, *44*(6), 793–803.
52. Russell, G. (2017). Logic isn't normative. Inquiry. Online first: https://www.tandfonline.com/doi/full/10.1080/0020174X.2017.1372305.
53. Russell, G. (2018). Logical nihilism: Could there be no logic? *Philosophical Issues*, *28*(1), 308–324.
54. Russell, G. (2019). Deviance and vice: Strength as a theoretical virtue in the epistemology of logic. *Philosophy and Phenomenological Research*, *99*, 548–63. Online first: https://onlinelibrary.wiley.com/doi/abs/10.1111/phpr.12498.
55. Schindler, S. (2018). *Theoretical virtues in science: uncovering reality through theory*. Cambridge: Cambridge University Press.
56. Shapiro, S. (1998). Logical consequence: Models and modality. In Schirn, M. (Ed.) *Philosophy of mathematics today: proceedings of an international conference in munich* (pp. 131–156). Oxford: Oxford University Press.

57. Shapiro, S. (2000). The status of logic. In Peacocke, C., & Boghossian, P.A. (Eds.) *New essays on the a priori* (pp. 333–366). Oxford: Oxford University Press.
58. Shapiro, S. (2006). *Vagueness in context*. Oxford: Oxford University Press.
59. Strawson, P.F. (1950). On referring. *Mind*, *59*(235), 320–344.
60. Tennant, N. (2004). An anti-realist critique of dialetheism. In Priest, G., Beall, J., Armour-Garb, B. (Eds.) *The law of non-contradiction: new philosophical essays* (pp. 355–384). Oxford: Clarendon Press.
61. Tennant, N. (2017). *Core logic*. Oxford: Oxford University Press.
62. Weiner, J. (1990). *Frege in perspective*. Ithaca: Cornell university press.
63. Williamson, T. (1994). *Vagueness*. London & New York: Routledge.
64. Williamson, T. (2013). *Modal logic as metaphysics*. Oxford: Oxford University Press.
65. Williamson, T. (2017). Semantic paradoxes and abductive methdology. In Armour-Garb, B. (Ed.) *The relevance of the liar*. Oxford: Oxford University Press.
66. Woods, J. (2019). Logical partisanhood. *Philosophical Studies*, *176*(5), 1203–1224.
67. Worrall, J. (1989). Fresnel, poisson and the white spot: The role of successful predictions in the acceptance of scientific theories. In Gooding, D., Pinch, T., Schaffer, S. (Eds.) *The uses of experiment: studies in the natural sciences* (pp. 135–157). Cambridge: Cambridge University Press.
68. Wright, C. (1986). Inventing logical necessity. In Butterfield, J. (Ed.) *Language, mind, and logic* (pp. 187–209). Cambridge: Cambridge University Press.
69. Yalcin, S. (2012). A counterexample to modus tollens. *Journal of Philosophical Logic*, *41*(6), 1001–1024.
70. Zardini, E. (2016). Restriction by noncontraction. *Notre Dame Journal of Formal Logic*, *57*(2), 287–327.