

On Logical and Scientific Strength

Luca Incurvati & Carlo Nicolai

24th July 2020

Abstract

The notion of strength has featured prominently in recent debates about abductivism in the epistemology of logic. Following Timothy Williamson and Gillian Russell, we distinguish between logical and scientific strength and discuss the limits of the characterizations they employ. We then suggest understanding logical strength in terms of interpretability strength and scientific strength as a special case of logical strength. We present applications of the resulting notions to comparisons between logics in the traditional sense and mathematical theories.

1 Introduction

Scientific theories are standardly thought to be selected on the basis of adequacy to the data and how well they fare with respect to a number of theoretical virtues (van Fraassen 1980; Lipton 2004; Keas 2017). One such virtue is strength, which has been discussed in the philosophy of science. This paper provides an account of the notions of logical and scientific strength. Our focus will be on logical and mathematical theories. However, our account is sufficiently flexible to be applicable also in more general contexts, such as scientific theories.

Our study is prompted by the recent interest in logical abductivism. This is the view that logical theories should be selected in the same way as scientific theories. Logical abductivism was famously advocated by Quine (1951), Goodman (1955), and Putnam (1968). It has received much attention in the recent literature as a way to navigate the wide array of non-classical solutions to the logical, set-theoretic and semantic paradoxes (Priest 2005; 2016; Williamson 2013; 2017). Logical abductivism promises to provide a way of resolving in a principled manner disputes between rival logics which would otherwise appear hard to settle. Abductivism, so the story goes, replaces clashes of intuition with

appeal to criteria for theory choice that are accepted by the broader scientific community. For instance, rather than debating the status of paradoxical sentences, one would determine which semantical theory scores better with respect to those criteria.

According to the logical abductivist, then, theory choice in logic is no different from theory choice in the natural sciences. But the recent revival of interest in abductivism has been associated with the idea that logic is similar to the natural sciences in other respects. This is known as *anti-exceptionalism* about logic. The anti-exceptionalist may hold, for instance, that logical principles are not analytic or (metaphysically) necessary or a priori (Hjortland 2017). As Gillian Russell (2018) and Stephen Read (2018) have pointed out, however, some form of exceptionalism is compatible with abductivism. Although our focus is on abductivism, our discussion is clearly relevant for any anti-exceptionalist position which embraces an abductive methodology.

Abductive methodology has been employed also for theory choice in mathematics. Bertrand Russell (1973) advocated the adoption of the regressive method to justify mathematical axioms. An abductivist-friendly account is famously given by Gödel (1947), who suggested that set-theoretic axioms may be extrinsically justified. More recently, Priest (2006) defended naïve set theory against iterative set theory on the grounds of alleged greater simplicity. A thoroughgoing abductivist approach to the philosophy of set theory has been advanced by Quine (1990: 95), who argues that considerations of simplicity, economy and naturalness sanction the Axiom of Constructibility. Against this, Maddy (1997) uses the maxims Unify and Maximize to instead reject Constructibility as a candidate axiom for a foundation of mathematics.

In the philosophy of science, van Fraassen (1980: 67–68) distinguishes between *logical* and *empirical* strength. A similar distinction is made by Williamson (2017) and Russell (2018) under the labels of *logical* and *scientific* strength. Roughly speaking, the notion of logical strength of a theory takes into account only its deductive power, whereas the notion of scientific strength has mostly to do with its informational content.

There has been some controversy about the status of the criterion of strength in the recent abductivist literature. Williamson thinks that logical and scientific strength are both virtues and that the former entails the latter. Russell accepts that scientific strength is a virtue but criticizes the view that logical strength should be regarded as one. A more radical position, adumbrated by Hjortland

(2017), holds that logical *weakness* – and therefore the capability of a logic of drawing more distinctions – is a virtue in a theory.

We examine these accounts of the notions of logical and scientific strength and find them wanting. We suggest understanding logical strength in terms of interpretability strength and scientific strength as a special case of logical strength. The emerging picture contrasts with Russell’s analysis in that logical and scientific strength may still be considered to be theoretical virtues, and with Williamson’s, in that scientific strength is a special case of logical strength.

2 Logical Strength

The aim of this section is to offer a novel account of logical strength. To clear the ground for our account, we first rebut arguments against the status of logical strength as a theoretical virtue and identify problems with extant accounts of logical strength.

2.1 Logical strength as a theoretical virtue

Williamson characterizes logical strength in terms of deductive power. On his account, a theory T is *logically stronger* than a theory T^* just in case every theorem of T^* is a theorem of T but not vice versa. This can be extended to consequence relations by saying that a consequence relation \vdash is stronger than a consequence relation \vdash^* just in case whenever \vdash^* holds so does \vdash but not vice versa.

Williamson’s characterization of logical strength makes it sound as if the comparison of logical theories is a metalinguistic affair. However, Williamson aims to vindicate the idea that it is not. To this end, he considers two strategies for comparing logical theories in a non-metalinguistic way. The first strategy consists in comparing logics by encoding a logic’s consequence relation as a special set of the logic’s theorems. First one reduces logically valid arguments to logical truths by replacing the entailment sign by a suitable conditional. Next, one replaces all its non-logical constants with variables of the corresponding type and universally binds them with quantifiers of that type. With this reduction in place, comparing logical theories is tantamount to comparing the sets of universal generalizations corresponding to their logical consequence relations.

The reduction of logical validity to logical truth makes use of the standard structural rules and of the standard rules for implication (conditional proof and

modus ponens). However, when evaluating logical theories, we want to consider substructural logics or logics which do not have a sufficiently strong conditional. Not to prejudge any issue against the non-classical logician, Williamson therefore considers a second strategy for comparing logical theories. According to this strategy, we compare logics by encoding their consequence relation via an operator which takes a set of premises as argument and returns the set of its consequences. Thus, if Γ is a set of sentences, $\text{Cn}(\Gamma)$ is $\{\varphi \mid \Gamma \vdash \varphi\}$.¹ Comparison of logical theories then proceeds by comparing the different $\text{Cn}(\Gamma)$ s to which the logical theories give rise for different choices of well-confirmed Γ . The second strategy, however, appears to make theory comparison a metalinguistic affair again, *contra* Williamson's intentions. For the set $\text{Cn}(\Gamma)$ is individuated via the relation $\Gamma \vdash \varphi$, which is metalinguistic: the elements of Γ and φ are mentioned rather than used.²

Williamson claims that logical strength is a theoretical virtue and that this, together with the fact that simplicity too is a virtue, amounts to a *prima facie* case for classical logic:

Once we assess logics abductively, it is obvious that classical logic has a head start on its rivals, none of which can match its combination of simplicity and strength. Its strength is particularly clear in propositional logic, since PC is Post-complete, in the sense that the only consequence relation properly extending the classical one is trivial (everything follows from anything).

Recently, Gillian Russell (2018) has challenged Williamson's claims. She agrees with Williamson's characterization of logical strength but argues that logical strength is neither a theoretical virtue nor a theoretical vice. According to her, if logical strength were a virtue, then, *ceteris paribus*, if theory T is logically stronger than theory S , T is better than S . Similarly, if logical strength were a vice, then, *ceteris paribus*, if theory T is logically stronger than theory S , T is worse than S . But, she continues, is plainly not the case that, *ceteris paribus*, a theory is always better off (worse off) by having more (less) of logical strength: a theory can have too much or too little logical strength. **Triv**, the trivial logic in which any sentence follows from any set of premisses, is too strong: *snow is white* just does not

¹In the current context it does not matter whether we characterize Cn in terms of logical consequence or derivability. Clearly, this matters when one considers logics that are not complete.

²We thank Tim Button for raising this interesting point.

entail *grass is purple*. **Ni**, the empty logic in which nothing follows from any set of premisses, is too weak: *snow is white and grass is green* do entail *snow is white*.

This argument will not persuade the defender of logical strength as a theoretical virtue. She can happily grant that if a theory is logically stronger than another theory then, *ceteris paribus*, it is better; but she will insist that in the case considered by Russell *ceteris* are not *paribus*. In particular, *Triv* is plainly not adequate to the data: by entailing everything, the theory sanctions entailments which contradict our intuitions about, say, *grass is green* not following from *snow is white*. Thus, this is just a case, where logical strength is trumped by the fact that the theory is not adequate to the data. As Williamson (2017: 335) puts it: ‘First comes fit with the evidence’. A similar response is available to the defender of logical strength as a *vice*: by entailing nothing, **Ni** fails to be adequate to the data.

Russell presents an analogy (which she credits to Dan Marshall). She suggests that saying that logical strength always makes a theory better would be like saying that theories of love on which more people love each other are always better than ones on which fewer people do. But even granting that this analogy works, the defender of strength can accept that such a theory of love will be better other things being equal. However, she will deny that in several specific cases things *are* equal: the hard reality of romantic life tells us, for instance, that a theory of love on which everybody loves everybody else is hardly adequate to the data.

Similar considerations apply to Read’s response to Williamson abductivist argument for classical logic. Read begins by observing that classical logic is not the only logic to be Post-complete, as witnessed by the case of Abelian logic. He then writes:

A good argument would still ask which logic was the right one: information is not everything, if some of that information is wrong. In the case of Abelian logic, some is indeed wrong: e.g.

$$((p \rightarrow q) \rightarrow q) \rightarrow p \quad (**)$$

is valid in Abelian logic, but is simply false (as an account of conditionals).

But we take it that Williamson would agree with much of this: logical strength is not everything and the case for classical logic is to be understood with the

proviso that the logic we want ought to also be data adequate. And classical logic's fit with the evidence can and has been challenged, e.g. by relevant logicians such as Read. One may consider logical strength a virtue whilst taking fit with evidence as another criterion for theory of choice.

Indeed, considering logical strength as a virtue is compatible with thinking that this virtue is always trumped by adequacy to the data. In the mathematical context, Maddy comes close to claiming as much. She is arguing in favour of the maxim *Maximize*, which tells us that we should strive for set theories which are as generous as possible. Maddy is very clear that subscribing to *Maximize* as a maxim in no way commits one to choosing the most generous of theories—the trivial theory. For, she says, this maxim can be trumped or at least curtailed by other maxims. In particular, she says, 'consistency is an overriding maxim' (Maddy 1997: p. 216).

Thus, the idea that logical strength is a virtue remains unscathed. Nonetheless, we do not stake a stance on whether logical and also scientific strength should ultimately be considered as virtues, vices or neither. Instead, our aim is to provide a framework for comparing the strength of theories which can be used by all parties in this dispute.

Even so, there are a number of issues with Williamson's characterization of logical strength as inclusion between sets of consequences. First, Williamson's characterization is not immediately applicable to cases in which one deals with different languages. In general, on Williamson's characterization, all we can say about the relative strength of two logics featuring disjoint sets of logical constants – such as intuitionistic propositional logic and **S4** – is that they are incomparable. On the other hand, our proposal will take into account translations between languages. This will enable us, for instance, to say something informative about the relation between intuitionistic propositional logic and **S4**, thanks to the so-called Gödel-McKinsey-Tarski translation.

Another issue with Williamson's characterization of logical strength concerns its use of the notion of a well-confirmed sentence. The idea is that we can assess a logic by considering $Cn(\Gamma)$ where Γ is a set of well-confirmed sentences, such as well-established principles of physics. However, in typical cases, whether the members of Γ are well-confirmed or not depends on the background logic of the relevant theory. For instance, whether certain principles of physics can be taken to be well-confirmed depends on whether their consequences fit with the data. But what these consequences are, in turn, may de-

pend on the background logic. Thus, it is not clear that we can find adequate Γ s which we can take to be well-confirmed independently of the background logic.

Finally, Williamson claims that logical strength entails a ‘looser notion’ of scientific strength, but he does not provide a detailed account of scientific strength and of why such an entailment should obtain. In fact, in what follows we will provide a detailed account of scientific strength and of its relationship with logical strength in which such an entailment will fail. More specifically, we will offer a characterization of logical strength based on the notion of translation. This notion will apply to theories formalized in different signatures, and so will be more encompassing than Williamson’s characterization in terms of inclusion. Second, our characterization will extend more naturally so as to apply beyond the purely logical part of a theory. Finally, our characterization will form the basis of a detailed account of scientific strength.

2.2 Characterizing logical strength

In our view, the strength of a theory has to do solely with its deductive power, and in particular with the structure of its derivations. A theory is as strong as another if the former can mimic the inferential structure of the latter. As mentioned, translations allow us to compare theories in different languages. We propose to compare the strength of theories in terms of the existence of suitable translations between them. In order to be suitable, translations need to preserve the structure of derivations.

In comparing theories with respect with their logical strength, we allow logical and non-logical primitives to be reinterpreted as long as the basic structure of derivations is preserved. Any other notion of strength which demands preservation of information in the logical or non-logical component of theories would not count as *logical* strength, because it would not abstract away as much as possible from specific content. For instance, our characterization entails that Peano Arithmetic (PA) and $\text{PA} + \neg\text{Con}(\text{PA})$ – where $\text{Con}(T)$ is a canonical consistency statement for T – have equal logical strength. This is essentially because $\neg\text{Con}(\text{PA})$ can be translated in PA in a way that preserves its role in derivations while re-interpreting the notion of provability involved in the consistency statement.³ As we will see later on, what distinguishes logical and scientific strength

³For the relevant facts concerning interpretations of inconsistency in reasonable theories containing a modicum of arithmetic, we refer to (Lindström 2003: Ch 7), and in particular (Lindström 2003: Thm. 8).

of theories is how specific information contained in their logical or non-logical primitives is handled. In particular, in comparing the scientific strength of theories we will impose stricter conditions on how specific information contained in the primitive concepts of theories is preserved under suitable translations.

We now implement these ideas into a formal framework for logical strength. We take mathematical theories to be axioms closed under some rules of inference. Logics are then identified with the closure of the empty set of axioms under specific rules of inference. When comparing mathematical theories in *classical logic*, it is customary to say that a translation from a language \mathcal{L}_1 to a language \mathcal{L}_2 consists of an ordered pair $\tau = \langle \delta, F \rangle$ where δ is the domain of the translation and F is a recursive mapping associating each n -ary relation symbol $R(y_1, \dots, y_n)$ of \mathcal{L}_1 with an \mathcal{L}_2 -formula $F(R)(y_1, \dots, y_n)$. The translation τ commutes with the connectives and δ relativizes the quantifiers so that, e.g. $(\forall x\varphi)^\tau := \forall x(\delta(x) \rightarrow \varphi^\tau)$. A translation τ from the language \mathcal{L}_1 of a theory T_1 to the language \mathcal{L}_2 of a theory T_2 is then an interpretation of T_1 into T_2 if for every set of \mathcal{L}_1 -sentences Γ and \mathcal{L}_1 -sentence φ , if $\Gamma \vdash_{T_1} \varphi$, then $\Gamma^\tau \vdash_{T_2} \varphi^\tau$ (where, as usual, $\Gamma \vdash_T \varphi$ is a shorthand for $\Gamma \cup T \vdash \varphi$, and Γ^τ is $\{\varphi^\tau \mid \varphi \in \Gamma\}$).

Finally, T_1 and T_2 are mutually interpretable if T_1 is interpretable in T_2 and vice versa. Given these definitions, we could then characterize logical strength for theories in classical logic by saying that a theory T_1 has greater or equal logical strength than a theory T_2 just in case there is an interpretation of T_1 in T_2 , and that they have the same logical strength just in case they are mutually interpretable. However, since we aim to deal with mathematical theories formulated in a given non-classical logic as well, we generalize the notion of interpretation above, and call a translation from \mathcal{L}_1 to \mathcal{L}_2 any recursive mapping that associates formulas of \mathcal{L}_2 with primitive concepts of \mathcal{L}_1 and that is recursively extended to more complex formulas by suitably commuting with the logical constants. An interpretation is then a translation that preserves provability in a such given logic.

Let us consider a few examples that will be relevant also for our later discussion. The proposed characterization entails that the following pairs of theories have the same logical strength: $\text{PA} + \neg\text{Con}(\text{PA})$ and PA as mentioned, ZFC (Zermelo-Fraenkel set theory with the Axiom of Choice) and ZF , PA and ZF_{Fin} (ZF with the Axiom of Infinity replaced by its negation).⁴ To mention examples

⁴The interpretation of finite set theory in arithmetic is due to Ackermann (1937). For the interpretation of the axiom of choice, the classical references are Gödel (1948) and Cohen (1963).

of non-classical theories, the theory $\mathbf{PA}_{\mathbf{k3}}(P)$ – that is Peano arithmetic formulated in the three-valued Strong-Kleene logic $\mathbf{K3}$ and in the language with an additional predicate P , whose interpretation may not be classical – is mutually interpretable (relative to the logic $\mathbf{K3}$) with $\mathbf{PA}_{\mathbf{k3}}(P) + \neg\text{Con}(\mathbf{PA}(P))$.⁵ Hence, the proposed characterization entails that the two theories have the same logical strength.

The characterization of logical strength in terms of mutual interpretability provides a precise formal counterpart to the idea that logical strength resides in a theory’s capability of mimicking inferential structures, possibly via translations that reinterpret primitive concepts. However, the characterization is not sufficient to deal with all cases of comparison of logical strength. For instance, we want to be able to compare pure logics, and in that case we want to reinterpret the logical vocabulary itself, whereas the standard notion of interpretation is designed so as to leave the logical vocabulary alone.

Whilst we cannot hope to preserve the meanings of the connectives when translating between logics, it seems that a translation between logics, besides the basic requirement of being recursive, ought at least to (i) be uniform so that, e.g., it is not the case that $p \wedge q$ is translated as $p \vee q$ but $r \wedge s$ is translated as $r \rightarrow s$ and (ii) allow going beyond translating each operator with a single operator, e.g. we want to be able to translate, say, $p \wedge q$ as $\neg(\neg p \vee \neg q)$. A suitable notion of translation is the notion of a *schematic translation* (Prawitz and Malmnäs 1968; Wojcicki 1988; Pellettier and Urquhart 2003). The general idea is that a translation is schematic if the translation of a complex formula is a fixed schema of the translation of its parts. As a result, formulae instantiating the same schema are translated in the same way. So, for instance, if $p \wedge q$ is translated as $p \vee q$, then $r \wedge s$ must be translated as $r \vee s$. But it is possible to translate $p \vee q$ as $\neg(\neg p \wedge \neg q)$.

To define the notion of a schematic translation, we first define the notion of a *schema*. A schema is a map from formulae (and possibly variables) to the formulae instantiating a *schema-string*, i.e. an expression featuring metalinguistic variables such as $\varphi \vee \psi$ or $\forall \alpha \varphi$ (Dewar 2018). We say that a translation from the language \mathcal{L}_1 of a logic L_1 to the language \mathcal{L}_2 of a logic L_2 is *schematic* if it is a recursive mapping τ such that (i) each atom p of \mathcal{L}_1 is assigned a \mathcal{L}_2 formula, and (ii) for each piece \spadesuit of logical vocabulary in \mathcal{L}_1 there is an \mathcal{L}_2 -schema \mathcal{T} such that for all sequences $\varphi_1, \dots, \varphi_\gamma$ of \mathcal{L}_1 -formulae $(\spadesuit\varphi_1, \dots, \varphi_\gamma)^\tau :=$

⁵In particular, in $\mathbf{PA}_{\mathbf{k3}}(P)$ P can appear in induction, and the induction principle of \mathbf{PA} needs to be formulated as a rule to preserve soundness (Halbach and Nicolai 2018).

$\mathcal{T}(\varphi_1^\tau, \dots, \varphi_n^\tau)$. A schematic translation τ from \mathcal{L}_1 to \mathcal{L}_2 is *sound* if for every Γ and φ in the language of L_1 , we have that if $\Gamma \vdash_{L_1} \varphi$ then $\Gamma^\tau \vdash_{L_2} \varphi^\tau$. A schematic translation τ from \mathcal{L}_1 to \mathcal{L}_2 is *exact* if for every Γ and φ in the language of L_1 , we have that $\Gamma \vdash_{L_1} \varphi$ if and only if $\Gamma^\tau \vdash_{L_2} \varphi^\tau$.

Schematic translations played a prominent role in the history of logic. Gödel, via the so-called *negative translation*, showed that there is a (exact) schematic translation of classical logic into intuitionistic logic. In doing so, he established the consistency of classical logic and classical arithmetic (Peano Arithmetic) relative to their intuitionistic counterparts. He also provided the basis of provability logic, justification logic, and Kripke semantics for intuitionistic logic by providing a schematic translation of the latter logic into the modal logic **S4**.

We take sound schematic translatability to be a core component of our account of logical strength. In fact, if we were dealing just with logics, we could simply characterize logical strength by saying that a logic L_1 has greater or equal logical strength than a logic L_2 just in case there is a sound schematic translation of L_2 in L_1 , and that they have the same logical strength if this holds mutually. One would obtain a different notion of logical strength with exact translations instead of sound ones. Although we believe this to be an alternative worth exploring, we here focus on sound schematic translations in order to preserve the intuitive idea that S being a sublogic of T implies that T is at least as (logically) strong as S .

So far we have only afforded the means of comparing either different logics or mathematical theories cast in the same background logic. However, we also want to be able to compare mathematical theories cast in different logics. For instance, we want to compare the logical strength of **ZF** and Heyting Arithmetic (**HA**), the theory whose axioms are those of **PA** but whose logic is intuitionistic logic rather than classical logic.

This kind of case leads us to our full characterization of the notion of logical strength, which is obtained via a two-stage process and subsumes the characterizations of logical strength that would be suitable in the case of logics or in the case of theories cast in the same logic. Given a theory T_1 with logic L_1 and a theory T_2 with logic L_2 , the idea is that to determine whether T_1 is at least as strong as T_2 one first schematically interprets L_2 into L_1 and then interprets T_2 (under the logic L_1) into T_1 .

LOGICAL STRENGTH T_1 is *at least as logically strong as* T_2 iff there is a sound schematic translation τ of the logic L_2 of T_2 in the logic L_1

of T_1 , and there is an interpretation (relative to the logic L_1) of T_2^τ in T_1 .

We say that T_1 is *logically stronger* than T_2 if T_1 is at least as logically strong as T_2 but not vice versa. Our definitions entail that, for mathematical theories formulated in classical logic, logical strength coincides with the familiar notion of interpretability strength. More generally, for mathematical theories in a given logic, logical strength coincides with the notion of interpretability strength relative to that logic. Similarly, when comparing purely logical systems, our characterization of logical strength reduces to the existence of a schematic translation, since we are taking logics to be theories with the empty set of non-logical principles.

We now discuss some applications of our characterization. We begin by considering cases of comparison between logics. Since schematic interpretability preserves undecidability, it is clear that classical predicate logic is logically stronger than classical propositional logic.⁶ The Gödel-Gentzen translation (Troelstra and Schwichtenberg 2003: §2.3) is an exact schematic translation of classical logic into intuitionistic logic. Therefore, intuitionist logic can mimic the structure of classical derivations – modulo reinterpreting some logical vocabulary. As a consequence, intuitionistic logic is as strong as classical logic. Moreover, intuitionistic logic is a sublogic of classical logic, and hence it can be trivially (schematically) translated in a sound way into classical logic. Hence, intuitionistic logic and classical logic have equal logical strength. The Gödel-McKinsey-Tarski translation is an exact schematic translation of intuitionistic logic into **S4**. Hence, **S4** is at least as logically strong as intuitionistic logic. Similarly to the previous case, we can also reproduce the structure of **S4**-derivations into intuitionistic logic.⁷ Thus, **S4** and intuitionistic logic have the same logical strength. In the context of comparison between modal logics, by translating $\Box A$ with $\Box A \wedge A$, one can show that the modal logics **K** and **T** have the same logical strength.

We now turn to applications of our notions to non-logical axioms. The full power of our characterization of logical strength comes into play when we consider theories formulated in different logics. For instance, our notion enables

⁶If one allows non-effective translations, one obtains the unsound result that classic first-order logic and classical propositional logic have the same logical strength (Kocurec 2017).

⁷A sound translation of **S4** into intuitionistic logic can be defined as follows. One can employ the ‘erasure’ translation schema to translate **S4** in classical logic, and then employ the Gödel-Gentzen translation. Transitivity of sound translations then gives us the claim.

us compare **ZF** to **HA**. Clearly, there is a sound translation of intuitionistic logic into classical logic such that:

$$\text{HA} \vdash_{\text{IL}} \varphi \Rightarrow \text{HA} \vdash_{\text{CL}} \varphi.$$

Then one simply interprets **HA** in **ZF** by means of the interpretation that relativizes quantification over natural numbers as quantification over finite ordinals. Since there is no interpretation of **PA** in **ZF**, this establishes that **ZF** is logically stronger than **HA**. Similar reasoning establishes that intuitionistic **ZF** (**IZF** for short) is logically stronger than **PA**.⁸ One first employs the Gödel-Gentzen translation **gg** to obtain:

$$\text{PA} \vdash_{\text{CL}} \varphi \Leftrightarrow \text{PA}^{\text{gg}} \vdash_{\text{IL}} \varphi^{\text{gg}}$$

Then one would need to show that PA^{gg} , qua subtheory of **HA**, is interpretable in **IZF**.⁹ Since **IZF** has (much) higher consistency strength than **HA**, there is no interpretation of the former in the latter theory.

The examples just discussed lend support to the adequacy of our characterization of logical strength based on preservation of inferential structure. One of the main advantages of our characterization is its generality. We are able to compare both logics and theories, and various combination thereof. We believe this generality is essential to the abductive comparison of logics and theories. Without the possibility of comparing theories with different logical and non-logical primitive vocabulary, there is little hope for logical abductivism to succeed.

Yet another advantage of our characterization is that it leads naturally to a precise characterization of scientific strength. It is to this issue that we now turn.

3 Scientific strength

In this section we first discuss Williamson's and Russell's accounts of scientific strength. We then propose our own account.

⁸IZF is obtained by taking the background logic to be intuitionistic and replacing **ZF**'s Axioms of Foundation and Replacement with the Axioms of \in -induction and Collection.

⁹This point requires extra care in defining a suitable notion of interpretation for intuitionistic theories. Our claim is true for reasonable notions of interpretation for intuitionistic theories (Visser 1999).

3.1 Williamson and Russell on scientific strength

Williamson (2017) holds that logical strength entails a ‘looser’ notion of *scientific strength*. For instance, since classical logic proves all instances of $\varphi \vee \neg\varphi$ and intuitionistic logic doesn’t, the former is logically stronger, but also scientifically stronger than the latter: according to Williamson, a general claim – all instances of excluded middle are valid – is scientifically more informative than its negation. Similarly, ‘the time between 3:14 and 3:16’ is more informative than ‘the time between 4:00 and 12:00’. So, although Williamson does not provide a detailed account of scientific strength, both logical form and a certain degree of accuracy are relevant for his view.

Russell (2018) rejects Williamson’s claim that logical strength implies scientific strength. She does so by distinguishing between two senses of scientific strength. According to the first, a logic L is scientifically strong if it is able to decide, for each argument form in a given language, whether the argument is L -valid or not. In this first sense, each logic is as strong as another, no matter how different they are in logical strength: each logic partitions the set of all argument forms into valid and invalid.

Russell describes her second sense of scientific strength as follows:

If our question is ‘which instances of LL can we use?’ (where LL is some disputed logical law) then the logically stronger logic tells us ‘all of them’ whereas the weaker logic says ‘not all of them’ – and this tells us nothing further about which particular instances are untarnished (Russell 2018: p. 12).

In this second sense the trivial logic **Triv** is the strongest logic, because to the question ‘How many instances of the argument form (Γ, φ) can we use?’ it answers ‘All of them’. Classical logic would then seem to be scientifically weaker than **Triv**, but stronger than, say, its logically weaker sublogics **K3**, the Logic of Paradox **LP**, and First Degree Entailment **FDE**. There are in fact some argument forms (Γ, φ) of which, unlike **Triv**, classical logic can accept only some instances. Similarly, there are familiar argument forms, such as $(\Gamma, \varphi \vee \neg\varphi)$ or $(\{\varphi, \neg\varphi\}, \psi)$, whose instances are uniformly licensed by classical logic but fail to be so in **K3**, **LP**, or **FDE**. Therefore, it would seem that there is a sense of scientific strength that is entailed by logical strength. However, Russell claims that this conclusion would be hasty: any sublogic of **Triv** can be extended to a logic that decides which instances of an argument form are acceptable, and which

aren't. In other words, each logic can be extended in such a way that, to the question 'How many instances of the argument form (Γ, φ) can we use?', it no longer provides the uninformative answer 'Not all of them'. Instead, the question is answered by providing a list of acceptable and unacceptable instances. Russell calls this process 'Triv recapture'. Now any logic that is subject to the procedure of 'Triv recapture' ends up being as informative as another. Since this equally applies to logic with substantially different logical strength, Russell concludes that there is no sense of scientific strength that is implied by logical strength.

We believe that both accounts of scientific strength offered by Russell are unsatisfactory. We start with Russell's first account: on this view, all logics are on a par with respect to scientific strength because either $\Gamma \vDash_L \varphi$ or $\Gamma \not\vDash_L \varphi$: according to Russell, a well-defined consequence relation is cast in a set-theoretic (classical) metatheory (Russell 2018: p. 557). However, it's clear that under this characterization the specific properties of consequence relations are not relevant at all for their scientific strength. In fact, it is simply a feature of Russell's classical metatheory that excluded middle holds for logical consequence claims. It follows that, as long as a notion of consequence is well-defined, any logic is as strong as it could be. But if the notion of scientific strength is to play any role in abductive methodology, then it should be capable of discriminating at least between some logics.

To avoid such an essential dependence on classical metatheory, one might try to generalize Russell's first definition of scientific strength by requiring that each logic L is as strong as another one *by its own light*. On this reading, however, Russell's claims cannot be true in general. There is nothing that guarantees that the notion of logical consequence we are employing satisfies bivalence. For instance, if our metatheory is formulated in a paracomplete setting governed by the logic **K3**, it won't in general be the case that ' φ follows from Γ or it's not the case that φ follows from Γ ', because the very notion of consequence may be partial (Nicolai and Rossi 2018). Moreover, in such scenario, it would seem that logical strength does indeed in many cases entail scientific strength. For instance, classical logic *is* able, for each Γ, φ , to determine whether $\Gamma \vDash \varphi$ or $\Gamma \not\vDash \varphi$, whereas **FDE** and **K3** cannot.

Russell's second sense of scientific strength is based on the notion of **Triv** recapture: any logic L can be consistently extended to a logic that decides which instances of a given argument form are valid or not. This understanding of sci-

entific strength faces serious difficulties too. First, it is worth noticing that Russell’s **Triv** recapture is substantially different from standard recapture strategies found in the literature on semantic paradoxes. Let us consider the case-study discussed by Russell. If one’s language amounts to a formal syntax plus a truth predicate Tr , one can provide models of transparent truth – $\text{Tr} \ulcorner A \urcorner$ is intersubstitutable with A in every context – that satisfy classical logic for all sentences without Tr . In other words, if $\mathcal{L}_{\text{Tr}} := \mathcal{L} \cup \{\text{Tr}\}$ is the language under consideration, one can consistently formulate a logic that satisfies all classical principles for \mathcal{L} and the nonclassical principles for \mathcal{L}_{Tr} . This is what is often called ‘classical recapture’ (Field 2008; Beall 2013).

However, this form of recapture is not sufficient for Russell’s purposes. She requires something much stronger – what she calls Small Square Completeness: for *any* argument form in a given language, one has to be able to decide which instances are licensed and which aren’t. For instance, each specific instance of the form $\text{Tr} \ulcorner \varphi \urcorner \vee \neg \text{Tr} \ulcorner \varphi \urcorner$ must be decided one way or another. This is a hugely complex task. If $\text{Tr} \ulcorner \varphi \urcorner$ is interpreted via fixed-point semantics in the style of Kripke (1975), the problem at hand reduces to a decision procedure for the set of paradoxical, or ungrounded sentences. Unlike the simple syntactic decision problem underlying recapture strategies, already in the simplest Kripkean setting (the minimal fixed point) this problem is highly non-effective (Burgess 1986). And these problems become much more complex for more sophisticated constructions such as other Kripkean fixed points, the revision extensions in Gupta and Belnap (1993), the theory of Field (2008), just to mention a few. Moreover, the complexity of the procedure envisaged by Russell is only going to increase if we move from the specific language \mathcal{L}_{Tr} to less rarefied languages closer to English. Therefore, the procedure of **Triv** recapture is simply unmanageable; it is not the case that any logic can be consistently extended to a Small-Square Complete *logic*, unless by logic we mean extensions of highly non-effective infinitary logics whose set of validities is much more complex than the provable sentences of any recursively axiomatised theory.

3.2 Characterizing Scientific Strength

We now come to our approach to scientific strength. Our proposal shares with Williamson’s the idea that scientific strength is more closely related to the informativeness of a theory than logical strength is. Our proposal goes further in that scientific strength is obtained by placing extra conditions on the relation of

being logically stronger. Thus, scientific strength entails logical strength.

Intuitively, logical strength is a coarser grained relation than scientific strength in that it mainly deals with preserving the deductive structure of theories, and hence allows for radical re-interpretation of logical and non-logical vocabulary in derivations. Scientific strength is then obtained by supplementing logical strength with stricter conditions so as to preserve some structural information contained in the theories' primitives. In particular, we no longer allow radical re-interpretations of primitives, but we impose conditions on the preservation, in derivations, of some structural aspects of logical and non-logical constants of theories. For instance, we have seen that PA and $\text{PA} + \neg\text{Con}(\text{PA})$ have equal logical strength, because the arithmetical primitives used to define provability in $\neg\text{Con}(\text{PA})$ can be re-interpreted by PA in a way that does not entail its inconsistency. However, they will not have the same scientific strength, because our extra conditions on interpretations will require the role in derivations of $\neg\text{Con}(\text{PA})$ to be preserved in a much more accurate way.

We formally render these ideas by means of the notion of intertranslatability. Intertranslatability is also known as definitional equivalence (Glymour 1970) and synonymy (De Bouvère 1965; Pellettier and Urquhart 2003). Earlier we distinguished between interpretations, which relate theories with non-logical axioms in the same logic, and schematic translations, which relate logics. Analogously, we now define intertranslatability as applied to both cases. Logics L_1 and L_2 are *intertranslatable* if and only if there are sound schematic translations σ from the language \mathcal{L}_1 of L_1 to the language \mathcal{L}_2 of L_2 and τ from \mathcal{L}_2 to \mathcal{L}_1 such that¹⁰

$$\begin{aligned} \varphi \dashv\vdash_{L_1} (\varphi^\sigma)^\tau & \text{ for any formula } \varphi \text{ of } \mathcal{L}_1; \\ (\varphi^\tau)^\sigma \dashv\vdash_{L_2} \varphi & \text{ for any formula } \varphi \text{ of } \mathcal{L}_2. \end{aligned}$$

Similarly, one says that theories S and T in a given logic are intertranslatable if there are interpretations σ from S to T , and τ from T to S (with both σ and τ relative to the given logic) such that

$$\begin{aligned} \varphi \dashv\vdash_S (\varphi^\sigma)^\tau & \text{ for any formula } \varphi \text{ of } \mathcal{L}_S; \\ (\varphi^\tau)^\sigma \dashv\vdash_T \varphi & \text{ for any formula } \varphi \text{ of } \mathcal{L}_T. \end{aligned}$$

¹⁰For an excellent overview of various notions of translations between logics extending sound and schematic translations, including original contributions, we refer to French (2010).

Since we are dealing both with pure logics and theories featuring non-logical axioms, we again need to characterize scientific strength in terms of a two-step process.

Intuitively, the idea behind our characterization is that a theory T (where, recall, logics are limiting cases of theories) is scientifically stronger than another theory S if there is some subtheory of T that can faithfully reproduce the logical and non-logical information contained in the inferential structure of S . The idea of ‘faithfully reproducing’ is captured in the strict requirement imposed to the translation by the notion of intertranslatibility. In particular, intertranslatibility requires that both theories recognize (via provability) that the translations that relate them are ‘companion’ to each other in the way they process the original information: when the two translations are suitably combined, they return the original information.

SCIENTIFIC STRENGTH A theory T_1 is scientifically as strong as T_2 if (i) T_1 is at least as logically strong as T_2 , (ii) the logic L_2 of T_2 is intertranslatable with a sublogic of L_1 – say, with $\tau : L_2 \rightarrow L_1$ –, and (iii) there is a subtheory T_0 of T_1 such that T_2^τ is intertranslatable with T_0 (with respect to the logic L_1).

Condition (i) in the characterization of scientific strength may be dropped in certain, well behaved cases, for instance when we deal with mathematical theories cast in classical logic. However, we chose to keep it in the general case because we aim to provide a template to deal with a large class of logics, for which the notion of interpretation may be underspecified. This makes it difficult to prove that condition (i) is redundant in full generality.

We now show that the definition delivers intuitively acceptable verdicts on the comparative scientific strength of theories. We start with examples of theories formulated in the same logic.

Since scientific strength entails logical strength, it obviously follows that any theories that do not have the same logical strength do not have the same scientific strength either. For instance, **ZFC** plus the assertion that there exists an inaccessible cardinal is scientifically stronger than **ZFC** which, in turn, is scientifically stronger than **PA**. For T a reasonable classical theory containing a modicum of arithmetic, $T + \text{Con}(T)$ is logically stronger than T , and properly so, since $T + \text{Con}(T)$ is not interpretable in T (Lindström 2003: Ch. 7). It is worth noticing that $\text{Con}(T)$ is a Π_1^0 -sentence of the language of arithmetic, i.e. a

purely universal claim. In general, the addition of an independent Π_1^0 -sentence results in a scientifically stronger theory. This last example obviously extends to theories in different languages that interpret a sufficiently strong arithmetical theory. So our characterization of scientific strength vindicates Williamson’s claim that a universally quantified sentence adds informativeness to a theory. More generally, our characterization entails that a theory is always scientifically as strong as any of its subtheories.

However, our notion of scientific strength is also flexible enough to accommodate cases of theories that *prima facie* deal with different mathematical domains. One example concerns set theory with and without urelements. By a result of Löwe (2006), **ZF** and **ZF** plus urelements are intertranslatable. Therefore, they have equal scientific strength. A similar phenomenon concerns **ZFC** and **ZFA** (**ZFC** without Foundation plus Aczel’s (1988) Anti-Foundation Axiom). **ZFC**-sets can be interpreted in **ZFA** as well-founded sets. **ZFA**-sets can be interpreted in **ZFC** as equivalence classes of graphs with lowest rank. Such interpretations yield the intertranslatability of the two theories (Visser and Friedman 2014). This example shows that sameness of scientific strength does not amount to sameness of meaning of the theories’ primitives, but only to equivalence with respect to salient aspects of a theory’s primitives that are relevant in the structure of derivations. Our characterization also applies to theories formulated in different signatures. Consider, for instance, the theory **ZF_{Fin}**. Although this theory is not intertranslatable with **PA** (Enayat et al. 2011: Thm. 5.1), it becomes so once one adds to it the claim that every set has a transitive closure (Kaye and Wong 2007).

Crucially, our analysis of scientific strength yields natural counterexamples to Williamson’s implication from logical to scientific strength. A striking example concerns set theory with and without the axiom of choice. **ZF** and **ZFC** have the same logical strength but not the same scientific strength. In particular, **ZFC** is not intertranslatable with **ZF**. Therefore, **ZFC** is scientifically stronger than **ZF** (Enayat 2016).¹¹ This nicely fits with the intuition that the addition of the axiom of choice to **ZF**, although innocent from the point of view of mere consistency strength, results in an increase of informativeness of the axioms. Similarly, although adding the Continuum Hypothesis or its negation to **ZFC**

¹¹Enayat shows that, for extensions of **ZF** in the language \mathcal{L}_ϵ of set theory, the relation of bi-interpretability – a slight weakening of the notion of intertranslatability – reduces to the sub-theory relation. This yields that the two theories cannot be bi-interpretable, and therefore not intertranslatable.

does not increase its logical strength, it does increase its scientific strength. As anticipated, canonical consistency statements display a similar behaviour: although $\text{PA} + \neg\text{Con}(\text{PA})$ has the same logical strength as PA , it is scientifically stronger than PA . There is in fact a subtheory of $\text{PA} + \neg\text{Con}(\text{PA})$ that is intertranslatable with PA , but the converse does not hold (Visser 2006: Cor. 9.4). A similar phenomenon holds for $\text{ZF}(\text{C})$ and $\text{ZF}(\text{C}) + \neg\text{Con}(\text{ZF}(\text{C}))$, as well as full second-order arithmetic Z_2 and $\text{Z}_2 + \neg\text{Con}(\text{Z}_2)$.

We now turn to the comparison of logics. We said in §2.2 that classical predicate logic is logically stronger than classical propositional logic. Since classical propositional logic is a subtheory of classical predicate logic, it follows that classical predicate logic is also scientifically stronger than classical propositional logic. We can also show that classical propositional logic is scientifically stronger than the many-valued propositional logics $\mathbf{K3}$, \mathbf{LP} and \mathbf{FDE} . That classical propositional logic is as scientifically strong as $\mathbf{K3}$, \mathbf{LP} and \mathbf{FDE} obtains because of the sublogic relation. For the other direction, we can show that none of $\mathbf{K3}$, \mathbf{LP} and \mathbf{FDE} can define the classical connectives. Since translational equivalence for logics entails that the connectives of one logic can be defined in the other without reinterpreting propositional letters (Pelletier and Urquhart 2003: Thm. 2.8), this establishes the failure of intertranslatability.

Here is our proof for $\mathbf{K3}$. If $\mathbf{K3}$ were intertranslatable with classical propositional logic, then it would feature formulas $N(\cdot)$ and $O(\cdot, \cdot)$ defining in $\mathbf{K3}$ classical negation and disjunction. However, in $\mathbf{K3}$, one can prove by induction on its complexity that for any formula φ containing only one propositional letter p , φ and $N(p)$ are $\mathbf{K3}$ -logically equivalent, where $N(p)$ can be one of:

$$p, \neg p, p \vee \neg p, \neg(p \vee \neg p).$$

By employing the explosion law for p and $\neg p$, and excluded middle for $p \vee \neg p$ and $\neg(p \vee \neg p)$, one can see that none of these alternatives are possible.¹²

¹²In more detail: since $p, \neg p$ classically entails q , we would have

$$p, N(p) \vDash_{\mathbf{K3}} q$$

However, this cannot be the case if $N(p) \equiv p$, if $N(p) \equiv \neg(p \wedge \neg p)$, if $N(p) \equiv p \vee \neg p$. If $N(p) \equiv \neg p$, we can use $O(p, q)$. In $\mathbf{K3}$, there are only the following forms $O(p, q)$ can take:

$$p \vee q, p \vee \neg q, \neg p \vee q, \neg p \vee \neg q, \neg(p \vee q), \neg(\neg p \vee q) \vee \neg(\neg q \vee p)$$

But $\mathbf{K3}$ does not entail:

$$p \vee \neg p, p \vee \neg\neg p, \neg p \vee \neg p, \neg p \vee p, \neg(p \vee \neg p), \neg(\neg p \vee \neg p) \vee \neg(p \vee p),$$

There are also logics that despite having the same logical strength have different scientific strength. One notable example is given by the modal logics **K** and **T**. We have seen that they have equal logical strength. However, a result of Pellettier and Urquhart (2003: Th. 4.5) entails that **T** is scientifically stronger than **K** (and vice versa) because they are not intertranslatable. The reason for this is that, since both logics have the finite model property, translational equivalence requires isomorphism of classes of finite models. However, since **K** is a sublogic of **T**, there are models of **K** of size n that are not models of **T**. The same result entails that the logics **B**, **S4**, **S5**, **T**, **B** all differ in scientific strength. There are nonetheless logics that have equal scientific strength. For instance, by a result of Lenzen (1979), the modal logics **S4.4** and **KD45** are intertranslatable.

Finally, what has been said so far also enables us to compare theories in different logics by means of scientific strength. In general, if S is a subtheory of T and formulated in a sublogic of the logic of T then S will be scientifically weaker than any extension of T which is not scientifically as strong as T . For instance, **HA** is a subtheory of **PA**, and therefore, by Visser's result on extensions of **PA**, **HA** is scientifically weaker than any proper extensions of **PA**. Similarly, **HA** is scientifically weaker than any theory that is properly logically stronger than ZF_{Fin} plus the assertion that every set has a transitive closure.

As we have seen in the case of logical strength, the examples just presented show that our notion of scientific strength gives intuitive verdicts for an interesting class of combinations of logics and mathematical theories. The core insight behind our notion of scientific strength is to preserve some structural information of the logical and non-logical primitives of a theory in derivations. This is realized by requiring that the re-interpretations of such primitives need to be, so to say, inverse to each other.

4 Abductivism and its strengths

We have presented a framework to analyze the notions of logical and scientific strength. By employing translations between theories, the framework allows one to compare the logical and scientific strength of theories in a formally pre-

since

$$\neg(\neg p \vee \neg p) \vee \neg(p \vee p) \not\equiv_{\mathbf{K3}} \neg\neg p \vee \neg p \not\equiv_{\mathbf{K3}} p \vee \neg p.$$

cise way.

The framework is directly applicable to the debate on logical and mathematical abductivism. Williamson (2017) and Russell (2018) analyzed logical strength essentially in terms of the subtheory relation. This fails to capture many interesting cases of theory comparison. Our framework allows theory comparison between theories that are not cast in the same language. Nonetheless, it also clarifies how the subtheory relation fits into a more general account of logical strength. In particular, being a proper subtheory of another theory implies being logically weaker than it.

One important question for the abductivist concerns the relation between logical and scientific strength. According to Williamson, logical strength entails scientific strength, essentially because more deductive power yields more information. If this is perhaps a plausible picture when comparing theories cast in the same language, it becomes harder to defend when one must translate between theories. For, if not suitably regimented, translations may compromise the information contained in theorems, and this is not compatible with theories having the same scientific strength. For instance, facts such as the interpretation of $\text{PA} + \text{'PA is inconsistent'}$ in PA rely essentially on distorting the information contained in $\text{'PA is inconsistent'}$. It then follows that logical strength cannot entail scientific strength.

By ensuring that the consequences of a theory are translated in accordance to suitable information-preserving constraints, our proposal maintains the generality given by understanding logical strength in terms of translations, while providing a notion of scientific strength as a refinement of the logical one. As a result, scientific strength implies logical strength but not vice versa: not all translations involved in the relation of logical strength are adequate for scientific strength. For instance, for PA to be scientifically as strong as $\text{PA} + \text{'PA is inconsistent'}$, the structural role played by $\text{'PA is inconsistent'}$ in derivations should be preserved, and PA has to be inconsistent after all. Hence, our notion of scientific strength gives its due to the intuitive idea that scientific strength has to do with the information contained in a theory.

Our framework combines notions of reducibility and equivalence that are usually employed in different domains. Interpretability strength is the standard tool to compare mathematical theories, schematic translations are generally employed to compare pure logics, and intertranslatability is a standard measure of theoretical equivalence for scientific theories. Therefore, our framework paves

the way to a unified approach to the comparison of formal theories. The specific combination of notions of reducibility employed in our characterization of logical and scientific strength delivers especially intuitive verdicts when applied to canonical examples. However, several alternatives are possible. For instance, faithful interpretability – in which not only provability, but also unprovability is preserved via the translation – may replace the looser notion of interpretability. Analogously, instead of focusing on sound translations in the comparison of pure logics, one can consider the stricter notion of exact translation. Finally, instead of intertranslatability, which is occasionally considered to be too strict for theoretical equivalence (Weatherall 2019), can be replaced by looser notions such as bi-interpretability (a.k.a. weak intertranslatability, homotopy equivalence) or categorical equivalence (Halvorson 2019). These alternatives will be considered in future work.¹³

References

- Ackermann, W. (1937). Die Widerspruchsfreiheit der allgemeinen Mengenlehre. *Math. Ann.*, 114(1):305–315.
- Aczel, P. (1988). *Non-well-founded sets*, volume 14 of *CSLI Lecture Notes*. Stanford University, Center for the Study of Language and Information, Stanford, CA. With a foreword by Jon Barwise [K. Jon Barwise].
- Beall, J. (2013). A simple approach towards recapturing consistent theories in paraconsistent settings. *Review of Symbolic Logic*, 6(4):755–764.
- Burgess, J. P. (1986). The truth is never simple. *Journal of Symbolic Logic*, 51(3):663–681.
- Cohen, P. (1963). The independence of the continuum hypothesis. *Proc. Nat. Acad. Sci. U.S.A.*, 50:1143–1148.
- De Bouvère, K. (1965). Synonymous theories. In Addison, J., Henkin, L., and Tarski, A., editors, *Symposium on the Theory of Models*, pages 402–406. North-Holland.
- Dewar, N. (2018). On translating between logics. *Analysis*, 78:622–630.
- Enayat, A. (2016). Variations on a Visserian theme. In van Eijck, J., Iemhoff, R., and Joosten, J., editors, *Liber Amicorum Alberti: a tribute to Albert Visser*, pages 99–110. College Publications, London.

¹³This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 758540) within Luca Incurvati’s project *From the Expression of Disagreement to New Foundations for Expressivist Semantics* and from a NWO Fellowship held by Carlo Nicolai.

- Enayat, A., Schmerl, J. H., and Visser, A. (2011). ω -models of finite set theory. In *Set theory, arithmetic, and foundations of mathematics: theorems, philosophies*, volume 36 of *Lect. Notes Log.*, pages 43–65. Assoc. Symbol. Logic, La Jolla, CA.
- Field, H. (2008). *Saving truth from paradox*. Oxford University Press, Oxford.
- French, R. (2010). *Translational Embeddings in Modal Logic*. PhD thesis, Department of Philosophy and Bioethics, Monash University.
- Glymour, C. (1970). Theoretical realism and theoretical equivalence. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1970:275–288.
- Gödel, K. (1947). What is Cantor’s continuum problem? *Amer. Math. Monthly*, 54:515–525.
- Gödel, K. (1948). The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory. *Uspehi Matem. Nauk (N.S.)*, 3(1(23)):96–149.
- Goodman, N. (1955). *Fact, Fiction, and Forecast*. Harvard University Press.
- Gupta, A. and Belnap, N. (1993). *The Revision Theory of Truth*. MIT Press.
- Halbach, V. and Nicolai, C. (2018). On the costs of nonclassical logic. *Journal of Philosophical Logic*, 47(2):227–257.
- Halvorson, H. (2019). *The Logic in Philosophy of Science*. Cambridge University Press.
- Hjortland, O. T. (2017). Anti-exceptionalism about logic. *Philosophical Studies*, 174(3):631–658.
- Kaye, R. and Wong, T. L. (2007). On interpretations of arithmetic and set theory. *Notre Dame Journal of Formal Logic*, 48(4):497–510.
- Keas, M. N. (2017). Systematizing the theoretical virtues. *Synthese*, pages 1–33.
- Kocurec, A. (2017). On the concept of a notational variant. In Baltag, A., Seligman, J., and Yamada, T., editors, *Logic, Rationality and Interaction: 6th International Workshop*, pages 284–298. Springer, Dordrecht.
- Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, 72:690–712.
- Lenzen, W. (1979). Epistemologische betrachtungen zu [s4,s5]. *Erkenntnis*, 14:33–56.
- Lindström, P. (2003). *Aspects of incompleteness*, volume 10 of *Lecture Notes in Logic*. Association for Symbolic Logic, Urbana, IL; A K Peters, Ltd., Natick, MA, second edition.
- Lipton, P. (2004). *Inference to the Best Explanation*. Routledge/Taylor and Francis Group.
- Löwe, B. (2006). Set theory with and without urelements and categories of interpretations. *Notre Dame Journal of Formal Logic*, 47(1):83–91.
- Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press.

- Nicolai, C. and Rossi, L. (2018). Principles for object-linguistic consequence: From logical to irreflexive. *Journal of Philosophical Logic*, 47(3):549–577.
- Pellettier, F. J. and Urquhart, A. (2003). Synonymous logics. *The Journal of Philosophical Logic*, 32:259–285.
- Prawitz, D. and Malmnäs, P. E. (1968). A survey of some connections between classical, intuitionistic and minimal logic. In Schmidt, H. A., Schütte, K., and Thiele, H. J., editors, *Contributions to Mathematical Logic, Proceedings of the Logic Colloquium, Hannover 1966*, pages 215–229. North-Holland Publishing Company, Amsterdam.
- Priest, G. (2005). *Doubt Truth to Be a Liar*. Oxford University Press.
- Priest, G. (2006). *In Contradiction: A Study of the Transconsistent*. Oxford University Press.
- Priest, G. (2016). Logical disputes and the a priori. *Logique et Analyse*, 236:347–366.
- Putnam, H. (1968). Is logic empirical? *Boston Studies in the Philosophy of Science*, 5.
- Quine, W. V. (1990). *Pursuit of Truth*. Harvard University Press.
- Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review*, 60(1):20–43.
- Read, S. (2018). Anti-exceptionalism about logic. *Australasian Journal of Logic*.
- Russell, B. (1973). The Regressive Method of Discovering the Premises of Mathematics. In Lackey, D., editor, *Essays in Analysis*, pages 272–283. George Allen & Unwin, London. Originally read before the Cambridge Mathematics Club, 09 March 1907.
- Russell, G. (2018). Deviance and vice: Strength as a theoretical virtue in the epistemology of logic. *Philosophy and Phenomenological Research*.
- Troelstra, A. and Schwichtenberg, H. (2003). *Basic proof theory*, volume 43 of *Cambridge Tracts in theoretical computer science*. Cambridge University Press.
- van Fraassen, B. (1980). *The Scientific Image*. Oxford University Press.
- Visser, A. (1999). Interpretations over Heyting’s arithmetic. In *Logic at work*, volume 24 of *Stud. Fuzziness Soft Comput.*, pages 255–284. Physica, Heidelberg.
- Visser, A. (2006). Categories of theories and interpretations. In *Logic in Teheran*, Lecture Notes in Logic 26. Cambridge University Press.
- Visser, A. and Friedman, H. (2014). When bi-interpretability implies synonymy. *Logic Group preprint series*, 320.
- Weatherall, J. O. (2019). Part 1: Theoretical equivalence in physics. *Philosophy Compass*, 14(5):e12592. e12592 10.1111/phc3.12592.
- Williamson, T. (2013). Logic, metalogic and neutrality. *Erkenntnis*, 79(S2):1–21.
- Williamson, T. (2017). Semantic paradoxes and abductive methodology. In Armour-Garb, B., editor, *Reflections on the liar*. Oxford University Press, Oxford.
- Wojcicki, R. (1988). *Theory of Logical Calculi. Basic Theory of Consequence Operations*. Kluwer, Dordrecht.