

Logical Abductivism: Challenges and Prospects

Elke Brendel

<https://orcid.org/0000-0001-9701-9768>

Filippo Ferrari

<https://orcid.org/0000-0001-6770-8124>

Filippo Mancini

<https://orcid.org/0000-0002-1061-7982>

Abstract

Logical abductivism is the view that rational theory choice in logic occurs through an abductive method: our justification for accepting a logical theory relies on whether that theory outperforms alternative theories on some relevant criteria. This contribution examines the core principles of logical abductivism, focusing on some core abductive criteria and their role in theory choice (sections 1 and 2). Section 3 explores the limitations of logical abductivism through the contrasting applications of it by two philosophers, Graham Priest and Timothy Williamson. Using broadly the same abductive method, they arrive at divergent conclusions: Priest endorses a paraconsistent logic, while Williamson endorses classical logic. It is argued that this divergence stems primarily from differing views on the nature of logical data and their underlying metaphysical commitments. This highlights a key limitation of abductivism: its dependence on extralogical assumptions, making it non-neutral in guiding logical theory choice. Section 4 deals with two epistemological challenges to logical abductivism: the Adoption Problem, which questions whether certain basic logical principles can be adopted, and the Background Logic Problem, which highlights the circularity in revising fundamental logical principles while relying on them for such revisions. Section 5 concludes the contribution by discussing prospects for logical abductivism.

1. Introduction

Some of the core questions in the *epistemology of logic* are: how do we acquire justification for the correctness of a logical theory? What guides us in selecting in an epistemically responsible way the logical theory to adopt? In recent years, *logical abductivism* has been advocated as a novel view in the epistemology of logic which provides answers to these questions by emphasizing a certain methodological similarity between the justification and selection of theories in logic and the sciences. According to logical abductivists, the rational process of justifying and choosing theories, both in logic and in science, should be based on *abductive reasoning*, understood here as a form of *inference to the best explanation*. Like scientific theory choice, they claim, one should select the logical theory that best accounts for the evidence and data and, furthermore, performs best with respect to scientific virtues such as *unifying and explanatory power*, *non-adhocness*, *fruitfulness*, *strength*, *simplicity*, etc. As Williamson writes:

“Scientific theories are compared with respect to how well they fit the evidence, of course, but also with respect to virtues such as strength, simplicity, elegance, and unifying power. We may speak loosely of inference to the best explanation, although in the case of logical theorems we do not mean specifically causal explanation, but rather a wider process of bringing our miscellaneous information under generalizations that unify it in illuminating ways.” (Williamson 2017, 334)

Accordingly, whether, say, the theory of classical logic, or that of some non-classical logic is the correct one to choose should be determined based on explanatory considerations, and in particular on the goodness of the explanations these logical theories offer for the relevant facts.¹ Although several substantial issues remain open, such as which facts (or data, or evidence)² a logical theory actually allows us to explain,³ many philosophers of logic, e.g. JC Beall (2017), Ole Hjortland (2017), Graham Priest (2014, 2016), Gillian Russell (2015), and Timothy Williamson (2017), have embraced logical abductivism placing it among the main epistemic views in contemporary debates in philosophy of logic.

Logical abductivism often aligns with logical anti-exceptionalism (e.g., Hjortland 2017, 2019), which views logic as methodologically continuous with science,⁴ and rejects traditional accounts which see logic as an a priori discipline dealing in analytic truths. However, while many abductivists are anti-exceptionalists, abductivism is independent of several core theses of logical anti-exceptionalism (see Hlobil 2020; Russell 2019).

2. Abductive Criteria of Logical Theory Choice

Key criteria for employing a logical theory in an abductive methodology include non-triviality, adequacy to data, a balance of strength and simplicity, informativeness, explanatory and unifying power for its intended aims, and intertheoretic coherence—i.e., alignment with other theories relevant to its potential applications.⁵

The criterion of non-triviality is a necessary condition for the rational acceptability of a logical theory, which is usually not explicitly mentioned by logical abductivists (since it is normally taken for granted). A logical theory *T* is trivial if every arbitrary proposition is provable or valid in *T*.⁶ But for a logical theory *T* that is even to be considered as a possible candidate within the process of an abductive theory choice, there must be propositions that are *not provable* or *invalid* in *T*.⁷

Adequacy to data is probably the most significant abductive criterion (see, for example, Hjortland 2019, 257; Priest 2014, 217; Priest 2016, 32; Williamson 2017, 334f.). A logical theory *T* only scores well in an abductive theory choice if *T* adequately fit the data associated with the phenomena that are to be modeled and explained by *T*.⁸ While a trivial theory can never qualify as the abductive “best” theory, adequacy to data, unlike non-triviality, is neither absolute nor fully objective—it varies in how well a theory accounts for the data. Moreover, explanatory value depends on an overall evaluation of all criteria. A theory that imperfectly fits the data may still succeed if it offers greater explanatory advantages than its competitors.

What qualifies as data in justifying and selecting a logical theory? As in the sciences, the data in logical abductivism are diverse. Priest argues that the relevant data are intuitions—or more precisely, considered judgments—about the validity of inferences in ordinary natural language reasoning.:

“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 2016, 41)

Priest notes that when intuitions are unified into a logical theory as general laws or rules of inference, they often lose their strong initial plausibility: “Very often, a form of inference strikes us as correct only because of an impoverished diet of examples” (Priest 2016, 43). Initially intuitively plausible inferences can also turn out to be *invalid*, as counterexamples may arise that did not become apparent outside of logical theorizing. Within logical frameworks, we sometimes discover that principles seeming intuitive in ordinary contexts are *invalid*: for instance, there might be reasons to reject modus ponens when paradoxical sentences are involved. Furthermore, some structural principles are only plausible under idealized assumptions (e.g. monotonicity or transitivity). As noted, in an abductivist methodology, even seemingly plausible logical principles must be evaluated against the theoretical virtues in abductive theory choice.

The question of the precise nature of logical data crucially depends on what one takes a logical theory to be about and what is considered to be its intended domain of application. As Jonas R.B. Arenhart points out: “Evidence for the adequacy of a logical theory [...] comes from either linguistic facts or from facts about the world” (Arenhart 2024, 2401). Ole Hjortland calls these two options the ‘metalinguistic approach’ and the ‘non-metalinguistic approach,’ respectively (Hjortland 2019). In general, metalinguistic approaches take validity, understood as properties of arguments, to be the primary object of study of logic, while non-metalinguistic approaches hold that logic is primarily about some of the most general features of the world.⁹ Priest clearly advocates a metalinguistic view while Williamson adopts a non-metalinguistic approach (Williamson 2013 and Williamson forthcoming). According to Williamson, logic is a discipline that must not be identified with its metatheory, which is particularly concerned with validity and logical consequence relations. The data logic is concerned with are abstract structures. For example, a logical truth, such as $\forall x(x = x)$, does not contain metalinguistic notions. It rather concerns something like the structure of identity and the scope of the universal quantifier. Furthermore, according to Williamson, the object languages of mathematical logics are also not metalinguistic in character, as, for example, set theory which studies, in particular, the structure of the hierarchy of sets.

As noted, both Priest and Williamson advocate logical abductivism. Moreover, they agree that an abductive approach in logic does not amount to a rejection of a priori (or armchair) epistemology. While the boundary between a priori and a posteriori justification is debated, intuitions about validity and semantic considerations of logical terms are generally seen as key a priori pieces of justification. For both, these forms of a priori (or armchair) justification are data to be addressed. We have already seen that Priest identifies intuitions about the validity of arguments as relevant evidence in assessing a logical theory. Priest also emphasizes that when we theorize about the validity of arguments, ‘a theory of the meaning

of logical words’ plays an important role (see Priest 2016, 46). Although Williamson adopts a non-metalinguistic approach, he nevertheless claims that, in using an abductive methodology, logic “can still remain a primarily ‘armchair’ discipline” (Williamson 2016, 263). According to Williamson, “there is no restriction to knowledge gained in some special ‘conceptual’ or ‘a priori’ or ‘intuitional’ or ‘armchair’ way” in an abductive methodology (Williamson 2016, 268). However, an abductive assessment of a logical theory in terms of a priori evidence, is “not the last word” (Williamson 2016, 274). Such a logical theory can be revised in light of further evidence (which may be a posteriori in nature)¹⁰ and the application of further theoretical virtues (e.g. strength, simplicity, intertheoretic coherence, and non-triviality). Thus, as Arenhart points out, logical abductivism, is not a ‘new’ epistemology of logic that replaces the ‘old’ a priori epistemology. It is rather a complement to traditional epistemology of logic (see Arenhart 2024, 2399).

3. Limits of Logical Abductivism

As Gillian Russell suggests, an abductivist methodology in logical theory choice “permits a hope that we might eventually be able to determine which is the correct logic to everyone’s satisfaction” (Russell 2019, 549). The thought is that if we can agree on a method for comparing competing logical theories based on the various criteria discussed above and aggregating the results (using an agreed-upon method of aggregation), we might reasonably expect to converge on a single logic (or a specific set of logics) as the outcome of abductive reasoning. However, as we will discuss in this section through an analysis of two prominent applications of the abductive methodology in logical theory choice, this expectation remains unfulfilled.

Unexplained phenomena often serve as a catalyst for abductive comparisons between rival theories. These phenomena, which include anomalies, paradoxes, or cases where existing theoretical resources fall short, provide data that a theory must address. Logical paradoxes—such as the *Liar paradox*—offer a particularly interesting area for abductive comparison between competing logical theories. As it is well known, a classical logic that is capable of encoding its own syntax and contains a fully transparent truth predicate ‘*Tr*’ for which an unrestricted disquotational schema (T-schema) holds, leads to a provable contradiction, as, in particular, the famous *Liar paradox* shows, where the T-schema can be rendered as: $Tr(\sigma) \leftrightarrow \sigma$ (for all sentences σ of the logical theory). So, on pain of triviality, one cannot keep both, all of classical logic and a fully transparent truth predicate. In light of the *Liar paradox*, Williamson admits that “[a]t first sight, the semantic paradoxes constitute unusually promising ground for an abductive critique of classical logic” (Williamson 2017, 339). However, for Williamson, an abductive cost-benefit analysis clearly points in favor of keeping classical logic while restricting the T-schema. As we have already seen, Williamson advocates a non-metalinguistic approach in the philosophy of logic. A logical theory unifies and systematizes general structures that underlie scientific and mathematical thought. The truth predicate and the T-schema are metalinguistic, and, according to Williamson, are as such not central components of scientific theories, which are not metalinguistic. Thus, Williamson argues that, based on criteria like simplicity and explanatory power, classical logic coupled with a restricted T-schema is abductively preferable to non-classical alternatives that maintain an unrestricted T-schema:

“For any complex scientific theory, especially one that involves some mathematics, will make heavy use of negation, conjunction, disjunction, the quantifiers, and identity. Thus, restricting classical logic will tend to impose widespread restrictions on its explanatory power, by blocking the derivation of its classical consequences in particular applications. By contrast, most scientific theories make no use whatsoever of a truth predicate and quotation marks, because they are not metalinguistic theories. Hence restricting disquotation makes no difference to their explanatory power. Once we take such extralogical applications into account, restricting classical logic must involve a vastly greater abductive loss than does restricting disquotation.” (Williamson 2017, 340)

Unlike Williamson, Priest advocates for a non-classical logic based on abductive reasoning—specifically, a paraconsistent system known as the “Logic of Paradox” (LP, see Priest 1979; Priest 2006a, Ch. 1). In LP, rather than restricting the T-schema, certain logical principles that hold unrestrictedly in classical logic—most notably *ex contradictione quodlibet* and *modus ponens*—are subject to limitations. As noted earlier, Priest supports a metalinguistic approach to the philosophy of logic. For him, intuitions about validity in natural language serve as the primary data a logical theory must address. It is an important feature of natural language that it allows assertions being made about its own linguistic entities, in particular, ascribing the properties of truth or falsity to sentences. Truth is thus a central semantic concept of natural language, and the T-schema formulates, in terms of disquotation, an immediately plausible adequacy condition for the application of the concept of truth. For Priest, restricting the T-schema would therefore result in a significant loss of expressive strength and explanatory power of a logical theory. Furthermore, Priest’s view on the metaphysical status of contradictions differs radically from Williamson’s. Some contradictions, according to Priest, are true (*dialētheiai*) which express facts to be accounted for by a logical theory. The self-referential ‘Liar sentence,’ stating its own falsity and giving rise to the *Liar paradox*, is precisely an example of *dialētheiai* for Priest. He argues that *dialētheiai* exist not only in logic and semantics but also as inconsistent facts in the empirical world, such as contradictory situations arising at the instant of a change in motion. (see Priest 2006a, Ch. 11). So, *dialētheiai* can even be empirical data that should be considered by a logical theory and not (as contradictions in classical logic) be avoided by all means. In particular, the classical *ex contradictione quodlibet* principle should therefore not be considered valid—and paraconsistent logics accommodate this requirement.

As illustrated, Williamson and Priest both use an abductive methodology to justify their preferred logical theories, applying similar criteria: non-triviality, data adequacy, a balance between strength and simplicity, and expressive power. Despite this, they reach contrasting conclusions: Williamson supports classical logic, while Priest advocates for paraconsistent logic, each based on abductive reasoning. This example shows that even when logicians agree on abductive criteria, assign them similar weight, and use comparable methods, applying abductive reasoning in choosing a logical theory can still lead to different results. Such divergence often stems from fundamental disagreements about the nature and scope of a logical theory or which data are relevant for explanation before abductive reasoning is applied. Furthermore, even if proponents of logical abductivism agree on the phenomena a logical theory must account for, differing metaphysical interpretations of these phenomena

can still lead to the justification of distinct logical theories. For instance, considering adequacy to data, if one views the *Liar paradox* as an avoidable metalinguistic phenomenon, irrelevant to the core applications of a logical theory, it does not provide strong abductive grounds for rejecting classical logic. However, if one sees the *Liar paradox* as an instance of true contradictions reflecting facts about the actual world, abductive reasoning is far more likely to support a dialetheist, non-classical logic. All this suggests, as Ulf Hlobil has pointed out (Hlobil 2020), that abductivism cannot serve as a neutral arbiter, at least not for foundational disputes, such as disputes about paraconsistency versus classicality. The underlying thought is that unless we agree on what the relevant data for assessing the adequacy of a logical theory are, the abductive machinery looks like an idle wheel in the process of adjudicating disputes like those between Priest and Williamson. This highlights an interesting limitation in the applicability of abductive methodology within contemporary debates in the philosophy of logic. While this limitation does not constitute an insurmountable objection to logical abductivism, it does reveal a structural concern regarding its use in logical theory choice. This concern stems from the *vexata quaestio* of what qualifies as data in the context of logical theories—an issue on which there’s ample disagreement in the current debate.

Disagreement in logical theory choice between advocates of an abductive methodology in the epistemology of logic might seem to discourage adherence to logical abductivism, but upon closer inspection, this also has a constructive potential. Reflecting on possible opposing views regarding the nature of logic and its relevant data can help uncover the deeper reasons behind the disagreement between logical abductivists like Priest and Williamson. Furthermore, a logical disagreement can be assessed in critically scrutinizing the respective abductive reasons that lead to a particular logical theory choice.¹¹ Although an abductive weighing of the benefits and costs of alternative logical theories does not always lead to a conclusive result, it can, nevertheless, provide important rational insights into a logical disagreement and the advantages and disadvantages of the respective logical theories. This approach may help in clarifying the core of many current logical disagreements, avoiding the concern, of broadly Quinean origin, that such disputes dissolve into mere verbal differences. This provides a first important positive aspect of logical abductivism and highlights one of the key prospects of adopting this methodology.

4. Two epistemological challenges: the Adoption Problem and the Background Logic Problem

In this section, we discuss two epistemological challenges to logical abductivism: the so-called *Adoption Problem*, which questions whether it is genuinely possible to adopt certain fundamental logical principles, and the so-called *Background Logic Problem*, which exposes a circularity in attempting to revise basic logical principles while simultaneously relying on them to carry out such revisions. In the subsequent and closing section we will discuss some prospects for logical abductivism in relation to the two challenges presented here.

The Adoption Problem (AP) originated from a seminar that Saul Kripke held at Princeton University in 1974 and from a lecture he gave at Pittsburgh the same year. The material was subsequently elaborated in some of his later contributions. His main thoughts on this matter

are discussed in Kripke (2023). Recently, several philosophers¹² have contributed to the discussion around AP. The reason behind the interest in this problem is that it would have “deep and wide implications for the epistemology of logic” (Birman, 2023, 37). Thus, one of the aims of this section is to explore what these implications are, especially for logical abductivism.

Given the large number of logical theories available today, the expression “to adopt a logic” is commonly used within the community of logicians and philosophers, although in a rather generic way, since it applies to quite different situations.¹³ AP is about whether it is possible to adopt certain logical principles. This may have consequences on the possibility of adopting a whole logical theory, since to adopt a logical theory is to adopt, among other things, its constituent logical principles. But note that in this case, “to adopt” has a very specific technical meaning: adopting a logic is understood as not merely accepting it as correct, but, crucially, also applying its principles in one’s reasoning (among other things). As Birman writes:

“[Adoption is] a two-phase process consisting in (1) the acceptance of a basic logical principle, and (2) the development, *in virtue of Phase 1*, of a practice of inferring in accordance with that principle.” (Birman 2023, 37)

The crucial point, then, is that the development of an inferential practice in accordance with a certain logical principle that is not based on the acceptance but on brute training or psychological conditioning does not qualify as adoption. Therefore, in this perspective, adoption is the acquisition of procedural knowledge (phase 2) by means of propositional knowledge (acceptance, phase 1).¹⁴ Thus, having established this meaning of “adoption,” AP is the claim that:

“certain basic logical principles cannot be *adopted* because, if a subject already infers in accordance with them, no *adoption* is needed, and if the subject does *not* infer in accordance with them, no *adoption* is possible.” (Birman 2023, 37)

AP has been justified as follows. Let us consider the logical principle of *universal instantiation* (UI): universal generalizations imply each of their instances. Suppose we want to have a cognitive agent S adopt UI, who is completely unaware of it: that is, S has never made an inference that conforms to the UI pattern and has no notion of the UI principle itself. It is important to stress that, since we want it to be a case of adoption, S must develop an inferential practice in accordance with UI based on the acceptance of UI. The point is that the mere acceptance of UI does not seem to be able to do this job. To see this, suppose S accepts that all logicians are funny and that Kripke is a logician, but since she is entirely unfamiliar with UI by hypothesis, she does not draw the conclusion that Kripke is funny.

(P ₁) All logicians are funny	
(P ₂) Kripke is a logician	
	(C) ?

Then, we tell S that universal generalizations imply each of their instances and, given our authority on the matter which is fully recognized by S, she accepts it. Thus, we have the following situation, where it is yet to be determined what conclusion S is able to draw:

- (P₁) All logicians are funny
 (P₂) Kripke is a logician
 (UI) All universal generalizations imply each of their instances

 (C) ?

To test whether S has really adopted UI (i.e., whether she has accomplished phase 2 as well) we ask her if Kripke is funny. But, to our surprise, she replies she does not know. The subject is stuck. To be guided by UI in performing what we may call the “basic-inference” above, she would have to appreciate that “All logicians are funny” is a universal generalization, and then to make an intermediate step in reasoning via what we may call the “principle inference,” from “All universal generalizations imply each of their instances” to “*This* universal generalization (i.e. “All logicians are funny”) implies each of its instances.” However, by hypothesis she is unable to do this since this very inference is itself an instance of UI.

To fully grasp the significance of AP it is crucial to appreciate that two sets of inferences are involved: the *basic-inference* and the *principle-inference*. And, according to Kripke, it is precisely because the ability of performing the basic-inference depends on the ability to perform the principle-inference that adopting UI is not possible. In other words, the general point here is that there is no reasoning pattern that (a) starts from only the premises “All logicians are funny”, “Kripke is a logician” and “All universal generalizations imply each of their instances,” (b) uses only logical principles that the subject is able to use by hypothesis, and (c) reaches the conclusion “Kripke is funny.” As suggested in Padró (2015), it can be shown that the same dynamic arises with other basic logical principles, such as *modus ponens*, *adjunction*, and possibly others.¹⁵

Let us now briefly introduce and discuss the Background Logic Problem (BLP). One notable formulation comes from Crispin Wright, first presented in Wright (1986) and later revisited in Wright (2021). The issue is presented as a challenge to a family of epistemologies of logic of Quinean lineage—epistemologies that are, in one form or another, endorsed by many proponents of logical abductivism.

One of the core tenets of Quine’s epistemology of logic is that no logical principle is immune to the possibility of empirical revision. When a situation of recalcitrance happens such that a certain observation O, which is taken by a theory T to be a consequence of a set of initial conditions I, is found to be incorrect, because as a matter of fact our evidence E is incompatible with O, everything that is involved in reaching the acknowledgement of such a situation of recalcitrance is open to revision. This includes the theory T, the claim that E agrees with I but disagrees with O, the epistemic good standing of E, the logic L that allows us to derive within T the testing conditional, as Wright calls it, that I implies O ($I \rightarrow O$), and, crucially, the claim that such a conditional is indeed an L-consequence of T—i.e., a statement that is derivable from T by some rule R of L. Wright’s objection, to put it concisely, is that it

is incoherent to maintain that situations of recalcitrance as those just illustrated allow us to hold accountable not just the empirical premises of the scientific theory but also any aspect of the logico-inferential apparatus involved, in particular the good standing of the rule R which allowed us to derive the testing conditional from T. To even appreciate that there is a problem (such as a situation of recalcitrance) in the first place, we need to presuppose the epistemic good standing of R. As Wright puts the point:

“There is incoherence in the idea that the case for revising a rule of inference, R, might rest on a derivation of a Problem—a situation of “recalcitrance”—in circumstances where the derivation relies essentially on R itself. The key thought is that the belief that you really have a Problem, which rationality requires you to remedy, must rely on the belief that your derivation is sound, so on a belief that R is good. If you then query that, you undermine your reason for thinking that you have a Problem in the first place.” (Wright 2021, 341)

Thus, according to Wright, even conceding that logic is by and large unexceptional, we must take some of its most basic principles (e.g., *modus ponens* and UI) as being by default in epistemic good standing and thus removed from the dynamics of any process of empirical theory testing.

BLP has direct consequences in the unrestricted application of an abductive methodology for logical theory choice. As we have seen, an important element that set the stage for devising alternative to, say, classical logic is the presence of anomalies (like paradoxes) or recalcitrant data (e.g. considered judgements of competent speakers of English concerning the meaning of ‘if...then’). To appreciate that there is a problem (such as an anomaly or a situation of recalcitrance) in the first place, we need to presuppose the epistemic good standing of some of the basic principles (of classical logic, in this example) that are needed to derive the anomaly or to appreciate that a certain prediction that is entailed by the theory and which is taken to be inconsistent with our ‘observation’ (i.e. considered judgements about the meaning of ‘if...then’ in natural language) is indeed a consequence of the theory.

Let’s discuss an example. This is intended to be a fictional example for illustrative purposes. Our classical theory (T_c) which defines ‘if...then’ in terms of material conditional, together with a condition of full competence as a user of English by a subject S (initial condition I) entails that S is compelled to accept any instance of *modus ponens* (observation O). Then we have our piece of recalcitrant evidence E which is given by S’s considered judgements about one instance of *modus ponens*, namely one of the (counter-)examples discussed in McGee (1985),¹⁶ which is deemed invalid by S. Now E is taken to be compatible with I but incompatible with O. To appreciate that this situation is indeed a challenge for the theory, we need to assume a background logic L and a rule R (which is, in our case, *modus ponens*) such that it allows us to appreciate that $I \rightarrow O$ is indeed a consequence of T_c . In this situation, E cannot be taken to be simultaneously in conflict with both O and R since R is involved in the derivation of the problem. Doing so would mean to undermine our reason for thinking that there is a problem to tackle in the first place. As mentioned above, this example is fictional, and we claim no accuracy with respect to the actual case made by McGee and the underlying reasoning that led him to take the examples discussed in his 1985 paper to be counterexamples to the validity of *modus ponens*.

BLP appears to impose structural constraints on the methodology of logical revision in logical abductivism—at least in cases where the theorist finds themselves in a predicament like the one described in our fictional example. In such scenarios, which are entirely possible and coherent, not every component of a logical theory can be open to revision. In particular, the very rule of inference used to identify the problem calling for a revision cannot itself be contested without undermining the problem’s initial recognition.

5. Some Prospects for Logical Abductivism

In this last section, we discuss some prospects of logical abductivism especially in relation to the two broad challenges discussed in the previous section.

Why is AP a challenge for logical abductivism? In theory, logical abductivism allows for the possibility of selecting a logical theory as the winner of an abductive comparison among several theories, even if this theory requires adopting one or more fundamental principles that are precisely those deemed inadmissible by the reasoning behind AP. So, in theory, AP present a potential challenge to logical abductivism, raising deep concerns about this methodology for choosing logical theories. There are, however, a couple of points to make in relation to this potential challenge. Even admitting that AP undermines the adoption of certain logical theories, this is not incompatible with logical abductivism *per se*. AP imposes constraints on which logical theories can be the subject of logical theory choice, but it is not in conflict with the use of the abductive methodology in that selection process. Restrictions are not necessarily a flaw, and it should be also considered whether they apply in a similar way to other competing epistemic views about logic. Moreover, the qualifier *in theory* is crucial here. In practice, most—if not all—logical theories considered as alternatives to classical logic in abductive comparisons are based on sub-classical logics. These theories preserve as much of classical logic as possible while imposing restrictions on the validity of one or more traditionally accepted classical principles. As a result, no adoption of previously unavailable logical principles is required. This demonstrates that, in the context of many contemporary debates in the philosophy of logic—particularly the dispute between Timothy Williamson and Graham Priest, which serves as our case study—AP loses much of its significance. A completely different scenario arises when, through an abductive comparison, a reasoner *S* is compelled to shift from endorsing a logical theory grounded in classical logic to one based on a contra-classical logic—for instance, logical theories based on some connexive logics or Abelian logics, which are neither extensions nor sub-logics of classical logic. Such a shift would require *S* to abandon certain classically valid principles and adopt entirely new logical principles that are fundamentally incompatible with classical logic. These principles are not mere restrictions or weakening of existing classical principles. Rather, if added to classical logic, would make the logic trivial. In this scenario—a genuine and open possibility for proponents of logical abductivism—AP would manifest its full force.

BLP and the fact that we always already presuppose and apply (at least certain instances of basic principles of) a logical theory in the background when we engage in a logical theory choice certainly have an impact on an abductive methodology in the epistemology of logic. It seems to put pressure on the idea of logical abductivism as a certain neutral and objective

method that, in particular, is capable of resolving disagreement between proponents of rival logical theories in an unequivocal manner. So, does BLP show that the use of an abductive methodology in logical theory choice is severely limited and insufficiently epistemically justified?

In section 3, we have seen, from the example of the disagreement between Williamson and Priest, that metaphysical or even empirical assumptions about the scope and the data to be explained by a logical theory can significantly influence the result of an abductive logical theory choice. BLP shows that the data to be explained could already be informed by the used background logic. In this vein, Arenhart and Molick write that:

“[...] the metalanguage and metalogic we have (or think we have) infiltrates in the process of theory choice not only by the inferences accepted, but also by interfering on how we judge simple issues such as *the choice of relevant factors* for theory evaluation. Philosophical agendas infiltrate, consciously or unconsciously, in these discussions. Consider, for instance, the logical facts which a system must accommodate in order to be appropriate. What is taken to be a logical fact, or the data, is already logic-laden, as it were, and the facts that must be taken into account already reveal the preferences of those in the dispute.” (Arenhart and Molick 2020, 275)

So, for example, it seems to be impossible, without already presupposing a paraconsistent background logic and a paraconsistent negation as part of a metalanguage, to state that there are inconsistent non-trivial theories. A classical logician, however, cannot on pain of self-refutation make such a statement (see Arenhart and Molick 2020, 277).

To be sure, we cannot perform a logical theory choice without already reasoning logically—and thereby using instances of certain logical principles. If these are instances of, say, paraconsistent logical principles and if we use these instances to characterize alleged relevant evidence or data, such as non-trivial contradictions, then it seems that this evidence can be used to support a paraconsistent logical theory. If, however, we use instances of classical logical principles in the background logic (such as *ex contradictione quodlibet*), then even provable contradictions, as the *Liar paradox*, are not regarded as non-trivial—and therefore need not be treated as evidence against classical logic. But does this mean that the background logic *always* functions as a kind of “self-fulfilling prophecy” (see Arenhart 2020, 19) such that the answer to the questions *What are the relevant data?* and *Which is the best logic to account for them?* is already decided by the used background logic? Does the interdependence between the data to be explained and the background logic necessarily predetermine the outcome of an abductive logical theory choice, reinforcing biases for or against a particular theory?

It seems that this is *not* necessarily the case. The choice of a background logic should be guided by abductive considerations as well. The logic that operates in the background of an abductive logical theory choice should use instances of logical principles that are as simple and generally accepted as possible. Furthermore, logical reasoning that is used in the background when choosing a logical theory, typically consists of instances of logical principles that do not have to be considered as necessarily universally valid and as part of a well-specified logic (see Arenhart’s approach to solving BLP in terms of logical nihilism and

logical naturalism in Arenhart 2020, 22). Additionally, it is possible to question one's own preferred logic and the evidence characterized by it. For example, provable contradictions within a classical logical theory could be seen as an indication that the preferred theory is not correct. A classical logician, even if she is using instances of classical principles in the background logic, could consider provable contradictions as a reason to abandon classical logic and, possibly, convert to a paraconsistent logic. Alternatively, she could try to deal with contradictions within the framework of her classical logical theory. She could, for example, endorse some kind of classical solution to the *Liar paradox*.

A logician is not much different from a natural scientist when being confronted with anomalies. Anomalies can often be explained (away) with additional assumptions within the framework of the preferred theory. For example, even if a classical logician does have the intuition that there are non-explosive contradictions, she could, nevertheless, dismiss this intuition as flawed (just as we can have deceptive intuitions regarding exponentially growing sequences or infinities, which are corrected by mathematical theories). Alternatively, a classical logician could limit the scope of her endorsed logic and the data to be explained in such a way that potentially inconsistent non-trivial contradictions are excluded (as it is the case in most areas of mathematics). But if she accepts inconsistent non-trivial contradictions, she must admit that there is an incoherence between her endorsed classical logic and the consequences that follow from it—and, thus, restrict, modify or reject the classical logic she employs. The fact that she herself uses classical principles when restricting, modifying or rejecting classical logic does not appear to be self-refuting. Similar considerations apply to proponents of non-classical logical theories. For example, a paraconsistent logician can critically reflect on whether there are non-trivial contradictions. Doing so does not depend on whether she thereby employs instances of classical or paraconsistent logical theories. Moreover, should it turn out that the paraconsistent logic she prefers is trivial, then a paraconsistent logician will modify or even reject her preferred paraconsistent logic, since it violates the most important abductive criterion in logical theory choice, namely the non-triviality of a logical theory.¹⁷

The above considerations suggest that an abductive methodology is not a procedure by which one can clearly and objectively select the one correct logic. The result of an abductive comparison of logical theories depends, *inter alia*, on assumptions about the nature and scope of a logical theory, as well as on what is considered as the data that such theories must account for. Both AP and BPL are usually regarded as challenges to logical abductivism. If AP poses a challenge, it does so insofar as it imposes restrictions on the logical theories under consideration in the selection process, which, however, in practice, don't concern current instances of disagreement between classical and non-classical logicians. In the case of BLP, while the influence of background logic on logical theory choice imposes some limitations, these are not severe enough to make an abductive methodology epistemically unjustified or question-begging. Instead, BLP challenges the idealization of logical abductivism as a purely neutral and objective method.

Acknowledgments

We are grateful to the editors of this volume and Maria Sekatskaya for their invaluable support, effort, and patience. Special thanks go to the members of the European Network for the Philosophy of Logic (EuPhilo) and participants of the Bologna-Bonn-Padova (BoBoPa)

seminar series, whose vibrant discussions have contributed to this work. We also thank the anonymous referee for their insightful comments on earlier drafts. The research for this work was supported by the DFG (Deutsche Forschungsgemeinschaft), research unit FOR 2495. In addition, Filippo Ferrari would like to acknowledge the support of PRIN 2022 PNRR “Understanding Scientific Disagreement and its Impact on Society” CUP J53D23017180001 funded by the European Commission (Next Generation EU) of which he is PI. Filippo Mancini’s work was funded by PRIN 2022 PNRR (Project ID: P20225A73K; CUP: C53D23009020001), funded by the European Commission (Next Generation EU).

References

- Arenhart, B. J. R. 2024. “Abductivism as a New Epistemology for Logic?” *Erkenntnis* 89(6), 2397–2415. <https://doi.org/10.1007/s10670-022-00634-7>
- Arenhart, J. R. B. and Molick, S. 2020. “On the Very Idea of Choosing a Logic: The Role of the Background Logic.” In *Abstract Consequence and Logics - Essays in Honor of Edalcio G. de Souza*, edited by A. Costa-Leite, 267–286. London: College Publications.
- Beall, JC. 2017. “There Is No Logical Negation: True, False, Both, and Neither.” *Australasian Journal of Logic* 14(1), 1–29.
- Birman, R. 2023. “The Adoption Problem and the Epistemology of Logic.” *Mind* 133(529), 37–60. <https://doi.org/10.1093/mind/fzad009>
- Boghossian, P. and Wright, C. 2024. “Kripke, Quine, the ‘Adoption Problem’ and the Empirical Conception of Logic.” *Mind* 133(529), 86–116. <https://doi.org/10.1093/mind/fzad011>
- Brendel, E. 2025. “Logical Abductivism and the Liar Paradox.” In *Demarcating Logic and Science: Exploring New Frontiers*, edited by S. Molick. New York: Springer.
- Carnap, R. 1937. *The Logical Syntax of Language*. New York: Routledge & Kegan Paul Ltd.
- Da Costa, N. and Arenhart, J. 2018. “Full-Blooded Anti-Exceptionalism about Logic.” *Australasian Journal of Logic* 15(2), 362–380. <https://doi.org/10.26686/ajl.v15i2.4865>
- Devitt, M. and Roberts, J. R. 2024. “Changing Our Logic: A Quinean Perspective.” *Mind* 133(529), 61–85. <https://doi.org/10.1093/mind/fzad010>
- Dummett, M. 1991. *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
- Erickson, E. 2024. “More Limits of Abductivism About Logic.” *Studia Logica*, <https://doi.org/10.1007/s11225-024-10130-4>.
- Ferrari, F. and Carrara, M. 2025. *Logic and Science*. Cambridge: Cambridge University Press.

- Hjortland, O. 2017. “Anti-Exceptionalism about Logic.” *Philosophical Studies* 174(3), 63–658. <https://doi.org/10.1007/s11098-016-0701-8>
- Hjortland, O. 2019. “What Counts as Evidence for a Logical Theory?” *Australasian Journal of Logic* 16(7), 250–282. <https://doi.org/10.26686/ajl.v16i7.5912>
- Hlobil, U. 2020. “Limits of Abductivism About Logic.” *Philosophy and Phenomenological Research*, 103(2), 320–340. <https://doi.org/10.1111/phpr.12707>
- Kripke, S. A. 2023. “The Question of Logic.” *Mind* 133(529), 1–36. <https://doi.org/10.1093/mind/fzad008>
- Lipton, P. 2004. *Inference to the Best Explanation*. London and New York: Routledge (2nd ed.).
- Maddy, P. 2002. “A Naturalistic Look at Logic.” *Proceedings and Addresses of the American Philosophical Association*, 76(2), 61–90. <https://doi.org/10.2307/3218629>
- Maddy, P. 2014. “A Second Philosophy of Logic.” In *The Metaphysics of Logic*, edited by P. Rush, 93–108. Cambridge: Cambridge University Press.
- Martin, B. 2020. “Identifying Logical Evidence.” *Synthese* 198(10), 9069–9095. <https://doi.org/10.1007/s11229-020-02618-y>
- Martin, B. and Hjortland, O. 2024. “Evidence in Logic.” In *The Routledge Handbook of the Philosophy of Evidence*, edited by M. Lasonen-Aarnio and C. Littlejohn, 467–482. New York: Routledge.
- McGee, V. 1985. “A Counterexample to Modus Ponens.” *The Journal of Philosophy*, 82(9), 462–471.
- Mohammadian, M. 2021. “Abduction – the Context of Discovery + Underdetermination = Inference to the Best Explanation.” *Synthese* 198(5), 4205–4228. <https://doi.org/10.1007/s11229-019-02337-z>
- Padró, R. 2015. *What the Tortoise Said to Kripke: The Adoption Problem and the Epistemology of Logic*. Ph.D. dissertation. City University of New York.
- Popper, K. 1963. *Conjectures and Refutations. The Growth of Scientific Knowledge*. London & New York: Routledge.
- Priest, G. 1979. “Logic of Paradox.” *Journal of Philosophical Logic* 8(1), 219–241.
- Priest, G. 2006a. *In Contradiction. A Study of the Transconsistent*. Oxford: Clarendon Press (2nd ed.).
- Priest, G. 2006b. *Doubt Truth to be a Liar*. Oxford: Clarendon Press.

- Priest, G. 2014. "Revising Logic". In *The Metaphysics of Logic*, edited by P. Rush, 211–223. Cambridge: Cambridge University Press.
- Priest, G. 2016. "Logical Disputes and the a priori." *Principios: Revista de filosofía* 23(40), 29–57.
- Putnam, H. 1969. "Is Logic Empirical?" In *Boston Studies in the Philosophy of Science: Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*, edited by R. S. Cohen and M. W. Wartofsky, 216–241. Springer, Dordrecht.
- Quine, W. V. 1951. "Two Dogmas of Empiricism." *Philosophical Review* 60(1), 20–43.
- Russell, G. 2015. "The Justification of the Basic Laws of Logic." *Journal of Philosophical Logic* 44(6), 793–803. <https://doi.org/10.1007/s10992-015-9360-z>
- Russell, G. 2019. "Deviance and Vice: Strength as a Theoretical Virtue in the Epistemology of Logic." *Philosophy and Phenomenological Research* 99(3), 548–563. <https://doi.org/10.1111/phpr.12498>
- Schurz, G. 2008. "Patterns of Abduction." *Synthese* 164(2), 201–234. <https://doi.org/10.1007/s11229-007-9223-4>
- Schurz, G. 2020. "Abduction as a Method of Inductive Metaphysics." *Grazer Philosophische Studien* 98(1), 50–74.
- Shapiro, S. 2014. *Varieties of Logic*. Oxford University Press.
- Stanley, J. and Williamson, T. 2001. "Knowing How." *The Journal of Philosophy* 98(8), 411–444.
- Williamson, T. 2013. *Modal Logic as Metaphysics*. Oxford: Oxford University Press.
- Williamson, T. 2016. "Abductive Philosophy." *The Philosophical Forum* 47(3-4), 263–280.
- Williamson, T. 2017. "Semantic Paradoxes and Abductive Methodology." In *Reflections on the Liar*, edited by B. Armour-Garb, 325–346. Oxford: Oxford University Press.
- Williamson, T. (forthcoming). "Is Logic about Validity?" In *The Oxford Handbook of the Philosophy of Logic*, edited by F. Ferrari, E. Brendel, O. Hjortland, G. Sagi and G. Sher. Oxford: Oxford University Press.
- Wright, C. 1986. "Inventing Logical Necessity." In *Language, Mind and Logic*, edited by J. Butterfield, 187–209. Cambridge: Cambridge University Press.
- Wright, C. 2021. "Making Exceptions." *Philosophical Topics* 49(2), 333–345.

¹ Given the scope and space constraints of this paper, this general characterization of abduction as a method for justifying theories should suffice. For a detailed classification of abduction patterns, such as selective versus creative abduction, and its role in inductive metaphysics, see Schurz (2008, 2020). For an analysis of the relationship between abduction and inference to the best explanation in science, refer to Mohammadian (2021).

² Here, we consider these terms as synonyms, though this is not a necessary choice.

³ See, e.g., Hlobil (2020, Sec. 3), Martin (2020) and Martin and Hjortland (2024).

⁴ For a systematic discussion of the relationships between logic and science see Ferrari and Carrara (2025).

⁵ These criteria, to be understood as neither exhaustive nor exclusive, are also proposed and discussed in Brendel (2025), mainly about the question of how they influence an abductive methodology in the choice of a logical theory that provides an adequate solution to the *Liar paradox*.

⁶ In classical logic, where the principle of explosion (*ex contradictione quodlibet*) holds, provable contradictions within a theory (T) lead to its triviality. Thus, inconsistency becomes a necessary criterion for logical theory choice. However, in paraconsistent logic, where the explosion principle does not apply (see Section 3), we propose the broader criterion of non-triviality as the primary and most important abductive criterion.

⁷ Similarly, Karl Popper's criterion of falsifiability for empirically significant theories can be taken in this sense when he writes: "Every 'good' scientific theory is a prohibition: it forbids certain things to happen. [...] Irrefutability is not a virtue of a theory (as people often think) but a vice" (Popper 1963, 48).

⁸ Logical abductivists often group 'adequacy to evidence or data' with theoretical virtues such as unifying power, non-ad hocness, strength, and simplicity, rather than treating it as a distinct abductive criterion. Critics argue that this approach diverges from traditional accounts of inference to the best explanation for empirical theories (e.g., Lipton 2000), where theories are first evaluated for empirical adequacy, and only then are theoretical virtues applied to identify the 'best' theory (see Erickson 2024).

⁹ Metalinguistic approaches include Carnap (1937), Dummett (1991), Shapiro (2014) and Priest (2016), while some examples of non-metalinguistic approaches are Maddy (2002, 2014) and Williamson (2017).

¹⁰ A prominent example of a logical theory that was justified by a posteriori evidence in the form of empirical data, is quantum logic. As Hilary Putnam has argued, empirical facts about quantum phenomena provide justifying grounds for revising classical logic (see Putnam 1969).

¹¹ Even if one adopts a non-metalinguistic view of logic, Williamson's abductive arguments for classical logic may not be entirely convincing. For instance, it is disputed whether classical solutions to semantic paradoxes—such as hierarchical, axiomatic, substructural, or contextualist approaches—clearly outperform non-classical alternatives in terms of expressive strength, explanatory power, and simplicity (see Brendel 2025). Additionally, it remains debatable whether scientific reasoning consistently and unambiguously adheres to the laws of classical logic (e.g., da Costa and Arenhart 2018). On the dialetheist side, the metaphysical status of certain contradictions as facts in the world also warrants critical scrutiny.

¹² Birman (2023), Devitt and Roberts (2023), Boghossian and Wright (2023), among others.

¹³ Among the most famous pioneers who believed it possible to adopt one logic over another, we find Quine (1951), Putnam (1968) and Carnap (1937).

¹⁴ We refer here to Gilbert Ryle's distinction between knowing-that (declarative or propositional knowledge, characterized by its representational nature) and knowing-how (procedural knowledge, often unrepresented or subconsciously represented). This distinction has been challenged by philosophers such as Jason Stanley and Timothy Williamson (see Stanley and Williamson 2001).

¹⁵ As for other examples of basic logical principles that are believed to be subject to AP, Kripke (2023) mentions the *law of non-contradiction* and Michael Levin suggested *simplification* in conversation with Birman.

¹⁶ The first example discussed by McGee is the following:

“Opinion polls taken just before the 1980 election showed the Republican Ronald Reagan decisively ahead of the Democrat Jimmy Carter, with the other Republican in the race, John Anderson, a distant third. Those apprised of the poll results believed, with good reason:

If a Republican wins the election, then if it’s not Reagan who wins it will be Anderson.

A Republican will win the election.

Yet they did not have reason to believe.

If it’s not Reagan who wins, it will be Anderson.” (McGee 1985, 462).

¹⁷ If the trivialization result arises solely from applying classical principles in the metatheory and would not occur under consistent use of paraconsistent reasoning in the metatheory, a paraconsistent logician need not reject their paraconsistent object theory. The key question is whether the paraconsistent metatheory is abductively justified—particularly whether it is non-trivial and meets other abductive criteria satisfactorily. For example, the model-theoretic semantics typically used to introduce the paraconsistent logic LP (via strong Kleene semantics) relies on classical principles. Unlike the object-language predicates "true" and "false" in LP, the meta-language notions such as "valid (in LP)" and "invalid (in LP)" or "having the designated value (in the model of LP)" and "not having the designated value (in the model of LP)" remain exclusive, as in classical logic. Consequently, a metatheoretic sentence asserting its own invalidity (in LP) would lead to a contradiction: it would have a designated value if and only if it did not, resulting in triviality. However, this does not imply that classical logic must ultimately determine the abductive justification of logical theories. Nor does it mean a paraconsistent theory must be rejected because a classical metatheory leads to trivialization. Instead, the result suggests that classical metatheories are not always suitable for non-classical object theories, and a paraconsistent metatheory might better avoid such trivialization.