

# The Problem with 'the Background Logic Problem'

Jonas R. Becker Arenhart  
Federal University of Santa Catarina, BR

BECKER ARENHART, J. R. : The Problem with 'the Background Logic Problem'

*Philosophica Critica*, vol. 6, 2020, no. 2, ISSN 1339-8970, pp. 2-29.

Anti-exceptionalism is the view that logic is, among other things, responsive to a posteriori evidence. This brings three questions to the fore: i) what is evidence in logic? ii) how to determine when evidence favours a particular logic, and iii) what kind of logic we use in such disputes? We argue that problem iii), when added to the thesis that natural language does have a specific logic, makes the whole project of logical theory choice untenable. The reason, we argue, is that the logic we are supposed to have vitiates any characterization of evidence, making a change of logic unlikely. We do that by considering two case studies involving adoption of paraconsistent systems, where the very idea that we have a logic operating in natural language prevent any possibility of change. We then go on and suggest, in broad lines, a version of anti-exceptionalism without the natural language logic hypothesis.

*Keywords:* Logical abductivism – Logical theory choice – Evidence – Paraconsistency – Logic of natural language

## 1 Introduction

The multiplicity of systems of logic leads to a perplexing philosophical problem: given that one of the uses (the so-called *canonical application*) of logic concerns the study of reasoning, or valid inferences, in natural language, is there any system of logic which may be called 'the correct one' for such an application? In case of an affirmative answer, how are we supposed to know which one is the correct system? Could it be that we are actually using a 'wrong' system all along? In order to address such questions, the recent epistemology of logic, in a large measure, has taken a rather

revolutionary turn, distancing itself from the aprioristic conception of logic, and took logic as responsive to *a posteriori* evidence, just as much as other empirical sciences. In this new setting, there is a ready answer to the question of which logic to choose: "logical theories are justified by their abilities to best accommodate relevant data. In other words, logical theories are justified by abductive means" (Martin 2020, 2).

Now, claiming that logic must accommodate data and that it uses abductive methodology for theory choice still leaves many gaps requiring further articulation. As a bare claim that logical theories need to accommodate the data, nothing prevents that the data come from *a priori* rational intuitions, for instance. This is not how most abductivists see the issue, though, and the abductivist account gets some more substance when wedded to a position known as *anti-exceptionalism* in logic (anti-exceptionalism and abductivism need not be identified<sup>1</sup>). Anti-exceptionalism is characterized, in particular by Hjortland (Hjortland 2019, 251), as involving the following claims (Hjortland 2017):

- Gradualism: theories of logic are continuous with theories in science
- Revisionism: theories of logic are revisable in terms of evidence
- Non-apriorism: the evidence for such theories is *a posteriori*

When put together, anti-exceptionalism and abductivism lead us to a general idea concerning logical theory choice which may be understood as follows: there is nothing special about logical theories, they are part of the enterprise of providing objective knowledge of the world, and the theory (or theories) we currently adopt may require changes in light of recalcitrant evidence, even of evidence coming from empirical sciences. The precursor of such a position is certainly Quine's web of belief, as presented in the final section of *Two Dogmas of Empiricism*, where it was suggested that logic is not immune to empirical evidence (Quine 1963), although one may adopt the same strategy without adhering to some of the specificities of the Quinean view, such as forms of holism and rejection of the analytic/synthetic distinction. In more recent times, then, people have been trying to articulate the view in more details, not only because it is interesting in itself, but also because of what is currently perceived as a failure of alternative epistemologies for logic<sup>2</sup>. The main challenges to the abductivist/anti-exceptionalist view of logical theory choice could be put in providing an answer to the following questions<sup>3</sup>:

---

<sup>1</sup> See Martin (2020)

<sup>2</sup> Martin (2020) provides for the details.

<sup>3</sup> See Hjortland (2019) and Martin (2020) for relevant discussion.

- a) The *evidence problem*: what counts as evidence for a system of logic?
- b) The *support problem*: how does evidence work to support a particular system of logic?
- c) The *background logic problem*: evaluating evidence in favour or against a system requires, in particular, that one provides for *prima facie* valid inferences leading from the evidences available to the correct or more appropriate system (or to a maintenance of the current system, if the evidence does not lead to a change in logical system). How is it possible, given that it is valid inference itself which is at stake?

Now, it could well be that one may provide for separate answers to each question, forming a coherent version of anti-exceptionalism. As we have mentioned, anti-exceptionalism still requires further articulation, and people have been developing it on a step by step basis, leaving the background problem without an answer (Hjortland 2019) (Martin 2020). However, as we shall argue here, this piecemeal treatment of the problems only postpones the major difficulties that the background logic problem generates, and ignores the fact that the background problem somehow impacts on the other two problems as well. Without a clear treatment of the background logic problem, answers to questions a) and b) may not work as expected. This happens because once one assumes that logic must be used to identify and evaluate evidence for logical theory choice, *the logic one uses may get into the way of a proper evaluation and identification of evidence*.

Fortunately, it seems to us, most of the difficulties raised by the background logic problem come rather from one further assumption which is generally conflated with the background logic problem: *that natural language has a logic of its own, and that it is this logic we use to evaluate evidence, and that it is this logic which stands to face the evidence, being always ready to be changed in the face of recalcitrant evidence*. It is this addition, we claim, which generates the troubles typically associated with the background logic problem. Furthermore, such troubles seem to be good reasons for us to reject the addition, rather than the need of a background logic.

To make things clearer, let us distinguish the background logic problem from the following additional claim:

- d) The *natural language logic*: reasoning in natural language proceeds according to a specific logic.

When we speak about application of logic to natural language, we follow the literature in the search for determining which logic is the one

operating in natural language, the one describing *legitimate validity*, and also, we want to choose a logic that does the best job in doing so. Clearly, the logic we currently employ may not be the best suited for the task, just as it is generally thought that Euclidean geometry is not the best geometry to describe physical space. We may have used Euclidean geometry as the correct geometry for a long time, but after General Relativity we changed to a non-Euclidean theory. The same could happen to logic, it is suggested.

Now, c) and d), when brought together, amount to the claim that in order to identify and evaluate evidence for or against a system of logic, *we have no choice but to use the logic we already have as the background logic*. The aim of this paper is to bring to light a tension between d) and the very project of changing logic in light of evidences. The original plan of applying abductive method to logical choice has as one of its goals that, with some combination of data and reasoning, one could end up in a situation in which it is recognized that there is good evidence to change our logic, which means that we end up discovering that the most appropriate logic is not the one we were using to begin with. However, we shall argue, given the conflation of c) and d), once a logic is assumed to be operating in the natural language, and that this is the background logic over which the process of characterization and evaluation of evidence is conducted, we just will not find evidence for changing it. The very idea of changing logic in light of evidence gets stopped before it starts.

Recognizing this difficulty motivates rejection of d), and may help us advance on the formulation of a more thoroughgoing version of anti-exceptionalism where c) is not seen as problematic. Our claim is that *there is no real tension in having to use a logic to evaluate the relation of evidence with a system of logic, provided that this evaluation is performed in an appropriate manner. The problem lies in the claim that there is a logic we use in natural language and that this logic is playing a role in the description and evaluation of evidence for logical theory choice*.

While anti-exceptionalism will clearly have difficulties addressing the nature of evidence in the scenario where current proposals are formulated, it is not so clear whether and how this kind of difficulty may affect exceptionalist epistemologies of logic, that is, epistemological approaches out of the abductivist/anti-exceptionalist spectrum. Typically, such epistemologies do not need to account neither for logical revision in terms of evidence, and nor for the relation of logical theories with *a posteriori* evidence; at least, they are not typically framed that way, so that the very notion of evidence needs not be characterized and related to logical theories. However, such exceptionalist epistemologies have difficulties of their

own, which themselves have led to the proposal of anti-exceptionalism as an attempted solution (see Martin 2020). In this sense, then, our discussion here is focused exclusively on this abductivist/anti-exceptionalist approach to the epistemology of logic; whether similar difficulties obtain or not for other views on the subject is an issue that we shall not touch on here.

The structure of this paper is as follows. Section 2 provides for the background on logical abductivism and anti-exceptionalism required to make the paper self-contained. It also advances the thesis of the natural language logic, which is a common addition to abductivism and anti-exceptionalism. Section 3 examines two case studies to illustrate the tension created by the fact that we have a background logic in natural language, and that this logic may be revised. Section 4 draws some lessons from the previous discussion, basically identifying the natural language logic hypothesis as the one to blame for the difficulties. It is indicated how an anti-exceptionalist story may be told about the procedure of logical theory choice that solves the background problem in a more reasonable way. We conclude by Section 5.

## 2 The Idea of Logical Choice

The idea that a logical theory is not imposed on us, but rather is a matter of rational choice, opens the door for abductivism as a methodology for theory choice (and whenever we speak about ‘logical choice’, we are speaking about choice of a *logical theory*). As it is discussed by the pioneer work of Routley (1980), and later on by Priest (2006a) chap. 8 and chap. 10, Priest (2016), Hjortland (2017) among others, choice of a system of logic is much on a par with *belief change*. That means in particular that choice of a system of logic involves much the same kind of procedure as any process of rational theory change involves: there is a wide range of theoretical virtues to be taken into account when considering rival alternatives, such as simplicity, non-*ad-hocness*, heuristic fruitfulness, and, more importantly, adequacy to the data (Priest 2016) (Hjortland 2017, 2020). As a result, logic is far from being a privileged science on this aspect (this is a mantra for anti-exceptionalism). As Routley has clearly put it some time ago:

“Choice of a logical theory is a special case of the choice of a theory or a system, and choice of these does not differ in principle from choice of such diverse items as a new house, a winner (e.g. of a gymnastics or equestrian contest), or of a recording of a symphony” (Routley 1980, 81).

In a nutshell, the proposed model for logical selection works as follows: given a list

$$c_1, c_2, c_3, \dots c_n$$

of criteria we believe a correct system of logic ought to have (adequacy to the data, simplicity, fruitfulness, explanatory power, non-*ad-hocness*, and so on), we proceed to select a range of values (say, from 1 to 10) to evaluate a system  $T$  according to each of those criteria by using a measure function  $m$ . So, suppose system  $T$  is simple, but not that simple; perhaps we believe it deserves only 6 on that count, so that  $m(s, T) = 6$  (here, “ $s$ ” stands for the criterion of simplicity). The same must be done with the evaluation of other criteria. The method also allows that we attribute different relevance for each of the criteria, so that each criterion has a respective weight  $w$  (adequacy to the data, for instance, is more important than simplicity). As a result, every criterion receives a weight and a value according to the measure  $m$ , and once these are settled, we may calculate what Priest (2016) calls a *rationality index*  $p$  for theories  $T$ , *i.e.* a weighted sum of each of the criteria. The result is:

$$p(T) = m(c_1, T)w_1 + m(c_2, T)w_2 + \dots + m(c_n, T)w_n$$

The theory that scores higher in this process is the most rational choice for the specific selection of theoretical virtues and weighting. Of course, the model is only ideal; it is not supposed to be applied in real cases of logical dispute as a practical calculation, it just somehow rationally reconstructs the typical dialectics that usually appears in such actual disputes. What is really important for the method, as Priest puts it, is that it codifies aspects of the practice of logicians, it brings to daylight the fact that “[w]hen people argue for a particular logical theory, what they are doing, in effect, is trying to show that their preferred candidate fares better on one or more of the criteria than a rival” (Priest 2016, 41).

Notice that there are a lot of trade-offs involved in selecting the weight to be attributed to the criteria to be applied in theory selection. One may consider that simplicity, although highly praised, may be sacrificed for the sake of explanatory power, for instance, so that a simpler system is not necessarily the best, when a rival system has more explanatory power (and scores similarly on other criteria). Also, one may value generality as an important virtue for systems of logic, praising more highly a system that accounts for more cases of situations than others, even if for that criteria such as consistency and simplicity have to be sacrificed<sup>4</sup>. All in all, the very choice of criteria and the respective attribution of weight is a further

---

<sup>4</sup> See the discussions in Routley (1980), for instance.

matter to be discussed, and one about which there is no complete agreement, neither in the sciences in general, nor in logic in particular; it depends very much on one's goals in developing a system of logic, and on how one conceives the role of logical theories in relation to their targets. Different priorities may be attributed by people having different goals, and no wonder that this happens frequently.

Although there is dispute concerning these criteria for theory selection, it may come as no surprise that there is consensus about what could be taken as the single most important criteria, *adequacy to the data*; this is the virtue most praised of all. It also generates some of the most difficult problems for abductivists and anti-exceptionalists.

“The main concern is to identify what counts as evidential confirmation of logical theories. The motivation is that ‘fit with the data’ is arguably the most important of the theoretical virtues associated with abduction. (In other words, it is assigned the greatest weight.) After all, in inference to the best explanation, something has to be explained. That something is the evidence. A theory can explain the evidence better than another, either by more fully accounting for the evidence, or by offering, say, a simpler or more unified explanation. The issue of evidential confirmation is at the heart of abductivism” (Hjortland 2019, 257).

If we are to evaluate theories according to their adequacy to the data, then, we must know what the data are and how they favour one theory over others (so that some theories may score better than others on the issue of adequacy to the data). These are problems a) and b) from section 1.

In a nutshell, evidence has been treated as coming from at least three sources (here, we follow Hjortland's (2019) useful catalogue of evidence as it is employed so far in logical discussions, and which may be used by an anti-exceptionalist). One may have evidence coming

- i) from *intuitions* about the validity and invalidity of some inferences. For instance, from some intuitively appealing counter-examples, one could hold that the law of excluded middle is not valid.
- ii) from the *global coherence of a system of logic along with other independently well-confirmed hypotheses* from science. When used as the mechanism of inference underlying well-confirmed empirical theories, a system of logic should cohere with the already known facts, and not lead from independently well-established premises to dubious conclusions, and also, not reject inferences that lead one from independently well-established premises to established conclusions.

iii) from a kind of *indispensability argument*: a system *S* may be the one that best accounts for the data because it is indispensable to empirical science.

It is not our goal in this paper to discuss the appropriateness of the characterization of evidence just described, or whether there are any other sources of evidence in logic. It is our claim that there will be a problem in properly getting to identify and evaluate the evidence once one assumes that there is a background logic operating in natural language. Basically, recall, we want to know what is the evidence available, because we are interested in determining which logic scores best at dealing with the evidence. Unfortunately, for those associating a logic with our natural language (which is basically item d) of the Introduction, usually conflated with c), accepting that it is this logic which is used to describe and evaluate the evidence, there will be a difficulty.

That natural language has an associated logic is a common claim among logicians in philosophy departments, not only those that are anti-exceptionalists. Hjortland, for instance, argues that disagreements about logic are disagreement about validity 'simpliciter' (Hjortland 2019, 252-253), a supposed notion of validity in the wild that one is attempting to describe through a system of logic. Which logic should we use in order to reason about evidence, and to ground the arithmetical reasoning that is required in order to make the computations involved in the application of the aforementioned method? We must use *the logic we have* in order to evaluate the evidence. Priest's answer to this specific question is framed precisely in these words:

"... we use [t]he logic (and arithmetic) we have. If we were trying to establish logical knowledge from first principles, then any use of logic would generate a vicious regress. But we are not: our epistemic situation is intrinsically situated. We are not *tabulae rasae*. In a choice situation, we already have a logic/arithmetic, and we use it to determine the best theory – even when the theory under choice is logic (or arithmetic) itself" (Priest 2016, 51).

This is clear: one must use the logic one has. That is an affirmation that something like d) holds. Which is the logic we have is not clear, though. There is clearly a major problem here of *logical identification*, because one would certainly be willing to know which exactly is the logic we use. That is an empirical question, which only makes sense if we ever do use a specific logic in natural language. We shall not deal with the identification problem here, but our argument shall point to the fact that there seems to be good reasons for an anti-exceptionalist to hold that there is no logic in

natural language (or, that this is the assumption that generates the most difficult problems for the anti-exceptionalist project).

One could claim that the system used to identify and evaluate the evidences is the same we are proposing as the correct one. After all, using another system could lead us to reason in terms of inferential patterns we deem incorrect. Perhaps, this could be defended along the following lines:

“The choices of system and metasystem – more generally, system and extrasystematic adjuncts – are by no means entirely independent. It is not satisfactory for example, to reject classical logic systemically, e.g. as involving mistakes or illegitimate assumptions (such as the law of excluded middle), and to use it metasystemically without further ado or qualification; for to do so would be to proceed by what are confessedly mistaken paths. Such choices of system and metasystem are of course valuable for *limited* specific purposes, such as convincing a classicist of a certain result, getting a lead on how a good proof might work by studying defective classic proofs, persuading a classicist that there is nothing unintelligible about a certain notion since it has a semantics that conforms to his standards, etc. But such choices are not generally satisfactory: they fail to cohere. Choice of a logical theory involves not merely delineation of a system or type of system, but also of requisite parts of metatheory – in a way that conforms to criteria for selection of the systemic theory” (Routley 1980, 94).

By characterizing a system in terms of a metalanguage that has a logic we do not accept, we are proceeding in terms of “confessedly mistaken paths”, which “fail to cohere”. Then, the claim appears to be that the same logic must be used to characterize the system in semantic terms (that is, as the background logic of the metasystem), and also in the object system. Perhaps the same kind of argumentation could be used to also advance which one is the background logic: in order to characterize evidence and evaluate it, we should use the logic we deem as the correct one.

The main point of having uniformity between system and metasystem concerns the fact that advancing a system of logic requires determining many distinct notions, such as the conception of truth that is involved, the meaning of the connectives, and a notion of validity. The failure of coherence concerns all the distinct aspects involved in the characterization of a system. In particular, the meanings of the connectives may be mistakenly characterized if the wrong metasystem is used:

“Any intuitionist or dialetheist takes himself to be giving an account of the correct behaviour of certain logical particles. Is it to be supposed that their account of this behaviour is to be given in

a way that they take to be incorrect? Clearly not. The same logic must be used in both 'object theory' and 'metatheory'" (Priest 2006a, 98).

That seems fair enough. Using classical metatheory, for instance, to characterize intuitionistic logic, one has the resources of classical *reductio ad absurdum* and the unrestricted validity of the law of excluded middle. This clearly is not acceptable for intuitionists. One is then open for charges of incoherence and, also, proceeding from confessedly mistaken paths. Following these advices would require that the background logic should be the very same system proposed as the best suited to the data.

But one could take another route too. The fact that system and meta-system must be the same in proposing an alternative logic still does not settle the issue of the logic we have. That is, one may propose an alternative logic with coherent system and metasystem, but this new system needs not be the system we have. Priest hinted at this point precisely:

"The choice of a logic is, as I have pointed out, a fairly major project, and many theoretical notions are part of the theory under choice. These are likely to include those relevant to the (metatheoretic) semantics of the logic. And, presumably, the (meta)logic of that semantics should be the logic itself — not the received logic. Thus, a theory that endorses intuitionistic or a paraconsistent logic should use that very logic in framing its own semantics. (Or if not, it is liable to face some charge of incoherence.) In other words, we, the theorists, use the received logic in performing our evaluation; but the theories to be evaluated are allowed to use their own logics 'internally'" (Priest 2016, 51-2).

Here we have again the claim that in order to characterize a system, the metalanguage must be precisely the same logic as the object theory, not *the received logic* (*viz.*, classical logic). Otherwise, one runs the risk of "charges of incoherence". Each logic is allowed, then, to be *presented* not only as an object system available for choice, but, also, to be *used* in order to act as a background system to determine the semantics of the object system (the truth conditions of logical operators in object language, for instance, must be framed in terms of the same connectives in the metalanguage). However, the last two lines also make one further thing clear: "we use the received logic in performing our evaluation". The background logic, which must be the logic we have, is the received logic (which Priest identifies with *classical logic*). However, Priest is an exception of this aspect; most philosophers, as we have already commented, simply skip discussion of the background logic, and with it, about the natural language logic (which

comes typically conflated with the background logic problem, recall). This will be of utmost importance in the next section.

Before moving on, let us take stock and see what we have got. The plan is that once distinct systems are available, there is a dispute on which one is more appropriate or correct. In order to do so, we must check which one scores best on virtues we choose as important for a system of logic to have when applied to inferences in natural language. To characterize each system, we use the very same system in the metalanguage, so that each competing system is characterized in its own terms. Once we have the systems characterized, we evaluate how each one of them fares on accounting for the evidence. The difficulty, as we shall argue now, is that while it is clear that the semantics of a system must be framed in terms of the system one is advancing, it is not clear that the evidence, when involving logical terminology (as it typically does), should also do so. One may have the option of using the received logic (as Priest suggests), or of using the rival logic. Any path chosen will bring enormous difficulties for the abductive proposal.

### 3 The Tension

Now, it is time to check how evidence gets identified and evaluated according to this picture. Once we do that, we will have answered to questions a) and b). As we shall argue, when one keeps in mind c) and d), which are also ingredients of the framework, one shall have trouble for the abductivist methodology. Recall that Priest conceded that evaluation is developed *inside the received theory*: “we, the theorists, use the received logic in performing our evaluation; but the theories to be evaluated are allowed to use their own logics “internally”” (Priest 2016, p. 52). The problem is that this still does not touch on the problem of evidence characterization when such evidence involves logical terminology. But that does not mean that evidence is not determined. Basically, claim d), the claim that we do have a logic in natural language, seems to saddle us with the received logic when it comes to spell out the behaviour of logical components involved in the evidence. That shall bring great troubles for any attempt to change logic in the face of evidence. But that is not bad enough: if one uses an alternative candidate logic to characterize the behaviour of logical particles, evidence also does not seem to be able to prompt logical revision. However one does the characterization, logical revision seems to get hindered by the framework we have (and that spots it as the problematical assumption). We shall argue for this claim by considering two case studies. The cases are representative of important disputes about logical theory, and point to a widespread phenomenon. The first case is more general, and concerns

paraconsistency without further specification. Most paraconsistent logicians argue that, given that we do reason in the presence of contradictions, but still, without triviality, our underlying logic must be paraconsistent. Evidence speaks in favour of adopting paraconsistent logic when one considers our deductive behavior. The second case is also about paraconsistency, but involves some rather specific ingredients of a logical theory: it addresses Graham Priest's claims that evidence suggest that some contradictions are true, leading to the view that truth and falsity overlap for some truth-bearers. This goes one step further than the previous case, because it is an argument for dialetheism and paraconsistency, while paraconsistency may not need dialetheism.

Our point in choosing precisely these two case studies is that they both concern presentation of evidence that consists typically of a derivation in natural language which already involves use of logical terminology. Given the cautious remarks of Priest and Routley we have quoted before, that every logical term must be completely defined and understood inside its own framework, in order to avoid that we proceed by confessedly mistaken paths and to avoid charges of incoherence, the question that naturally arises is: what about the logical terminology that appears when one is evaluating and presenting evidence for a system of logic? Should it be understood in the terms of the received logic, or in the terms of the concurrent system that is being advanced as a more able candidate? We have seen that Priest seems to suggest the former, but some quotes by Routley also suggest the latter. As we shall argue, both options prevent that we properly evaluate evidence if this evaluation is supposed to eventually lead to a change in logic.

### 3.1 Case 1: Evidence for Paraconsistent Logics

A system of logic is *paraconsistent* if it invalidates the law of explosion: in paraconsistent logics, from premises  $A$  and  $\neg A$ , one cannot infer every formula  $B$  of the language of the system (see da Costa, Krause, and Bueno 2007). If conjunction behaves as expected, that also means that from  $A \ \& \ \neg A$  one cannot infer any  $B$  whatever. Now, in informal terms, it means that, from a contradiction, not everything follows. One of the main claims by paraconsistent logicians is that this is clearly more adequate to the evidence available, given that in many contexts people are faced with contradictions, but do not infer everything (a contradiction does not lead to triviality). If classical logic were the correct logic, we would have no choice but to infer that everything is the case. Does that mean that paraconsistent logics should be adopted?

According to Routley (1980, 96), it is a *fact* that there are inconsistent and non-trivial theories and situations (Priest 1984, 128), makes similar claims, but given that they are related with the paradoxes to be evaluated in the next case, we postpone discussion of it for the next section). Here, a fact is taken to be a kind of hard fact, so that if a system of logic does not account for it, it is simply out of the game of theory choice. Clearly, that puts the received logic as straightforwardly inadequate, and different versions of paraconsistent logics must compete among themselves to check for which one is the correct logic (every sensible logic competing should be paraconsistent, then). Explosive logics are non-starters. In particular, Routley explains, the point is that for some theories  $T$ , and some deductions from  $T$ , we have both  $A$  and  $\neg A$  following, but we do not have every  $B$  holding. These theories have an underlying paraconsistent logic, and these deductions are based on a paraconsistent logic. The deductions are the evidence that classical logic is not adequate to the facts, which means it fails to account for correct reasoning.

One moment of pause for us to get the situation clear. Let us concentrate on an example. Suppose we are dealing with Cantor's naive set theory (one could change the example to any case of alleged inconsistency in science, say, the early calculus, or the Bohr model of the atom). What is being claimed is that a derivation of a contradiction appears in such a theory. Indeed, it is well-known that Russell's paradox, for instance, is derivable from apparently uncontroversial assumptions (we need not discuss them now). As it is known, the presentation of Russell's paradox involves defining Russell's set from the principle of naive abstraction, that is, one easily obtains  $R = \{x \mid x \text{ is not a member of itself}\}$ . By considering separately the two possible cases, of whether  $R$  is a member of itself or not, one derives that  $R$  both belongs to itself and does **not** belong to itself. This is clearly a contradiction, reached by a short derivation from a theory, and it does involve logical terminology, in particular, a negation.

Now, let us concentrate on this boldfaced negation, and let us keep in mind that we are claimed to have a logic in natural language. Is the derivation of Russell's paradox evidence for a change of logic, as claimed, or is it not? Suppose that our logic is the received logic, *i.e.*, classical logic. Then, we presume, the **not** in the derivation of the contradiction is to be understood in terms of classical logic, and, to be fair to classical logic, its characterization is made in terms of the classical metalanguage. That leads us to a classical contradiction in the case of the derivation of Russell's paradox. What does it mean in terms of evidence to be accounted for? Well, it means bad news for Routley's (and the paraconsistentists

in general) claims. Indeed, from that particular point of view, the theory in question is to be seen as *classically inconsistent and trivial*. How, then, can one claim that this is a case of inconsistent and non-trivial theory? In other words: how can such a derivation be seen as *a fact* about inconsistent and non-trivial theories? Framed in terms of the received logic, that is simply an impossibility! Every inconsistent theory will be automatically trivial. There is simply no space for a fact to appear that will be evidence against explosion as an inference rule! That is, one cannot expect evidence for a contradiction to hold (they never hold, in this framework), without everything else holding.

Notice that this does not make any claim that classical logic is right. All that is being claimed is that, once classical logic is assumed as being our current logic, and the one that describes the behaviour of negation (in particular), then, a derivation of a contradiction using this negation is explosive, and the derivation itself just cannot be seen as evidence against classical logic. The reading of the evidence in terms of our current framework simply prohibits that it be evidence against it.

But one could reason differently. Just as the characterization of every system of logic must be made in a metalanguage that is the same as the object logic, one could claim, the evidence, and in particular the negation of our example, must be characterized in terms of the logic one is claiming better suited to account for the evidence. In other words, one must understand the **not** used in the derivation of Russell's paradox in terms of a paraconsistent negation. Let us evaluate such a possibility.

For a negation to be paraconsistent, in semantic terms, what we need is quite simple:

"It is clear that in order to give a counterexample to the principle of explosion we need a weaker negation and a semantics in which there is a model  $M$  such that  $A$  and  $\neg A$  hold in  $M$  ( $\neg$  is a paraconsistent negation) but for some  $B$ ,  $B$  does not hold in  $M$ " (Carnielli – Rodrigues 2015, 62).

A paraconsistent negation is characterized, in semantic terms, as a negation which allows for models in which a formula and its negation may receive designated values, and which requires that at least one formula, in that same model, does not receive designated values. Notice what that implies: the characterization of a paraconsistent negation assumes, beforehand, that a contradiction may have a model, and that it does not lead to triviality. This fact has important consequences for the claim that the derivation of Russell's paradox is evidence for a paraconsistent negation: the claim that we have derived a contradiction and do not derive everything

from that contradiction was already assumed in the very definition of paraconsistent negation. So, seen from the perspective of a paraconsistent logic, when this is taken as the logic in which the evidence is framed, it results obvious (true by the very definition of paraconsistent negation) that we have a contradiction without triviality. But that obviousness does not arise from the derivation itself, which was being presented as *data* or *fact* for the adoption of such a negation, but rather from the background logic used to characterize the negation, which introduces such a feature right from the start. In this sense, when seen from this perspective, the derivation cannot be counted as evidence for paraconsistent logic, because, given that paraconsistent logic is assumed as the background logic, it just puts there what one was expecting to extract. One gets a paraconsistent rabbit from a paraconsistent hat, but that is clearly not what one wanted.

As a result, the claim that we use a logic to evaluate and frame the evidence vitiates the debate. On the one hand, using classical logic simply blocks any claim that we could ever find evidence for paraconsistent logics; the facts just tell us another story. On the other hand, using paraconsistent logics to frame and evaluate the evidence simply puts by hand, as it were, what was expected to be learned from the evidence. Both readings of the situation put one logic in charge of the behaviour of logical particles we use in the inferences we make in natural language, and attribute what we could call '*a theoretical profile*' to the negation employed in the derivation that was being advanced as evidence. This very attribution of a theoretical profile makes change of logic in terms of evidence difficult, to say the least; basically, there can be no recalcitrant evidence.

### **3.2 Second Case Study: the Liar Paradox**

Now, let us consider a case where we also have a derivation of a contradiction, but which implies a revision not only of our inferential apparatus, but also of some of the basic concepts of the background logic, that is, the concepts of truth and falsity. We shall consider the derivation of the Liar paradox, and Graham Priest's claim that such a derivation requires that one changes to a paraconsistent logic with a specific view of the behaviour of truth (Priest 2006) (Priest 2006a). Here, a 'theory of truth' is, in general outline, a description of the behaviour of truth and falsehood. The dialetheist approach is general enough to be consistent with distinct approaches to truth, such as truth as correspondence and truth as coherence, so that we need not enter into such level of detail (see Priest 2006a, chap. 1).

The case of the Liar paradox is very similar to the case of Russell's paradox. The Liar sentence  $L$  is a sentence saying of itself that it is not true:

$L$  iff  $\langle L \rangle$  is not true.

Here,  $\langle \rangle$  is a name-forming device, so that  $\langle L \rangle$  is the name of  $L$ . Again, there is a derivation by a reasoning by cases, where the cases involve assuming that  $L$  is true or  $L$  is not true. If it is assumed that  $L$  is true, one easily derives that  $L$  is also not true, and if it is assumed that  $L$  is not true, one easily derives that  $L$  is true. In the end, (given that  $L$  is true or  $L$  is not true is assumed as derivable from the empty set of premises),  $L$  is both true and **not** true. Priest argues that none of the steps in the derivation is to be spotted as unsound, and the conclusion is a contradiction. Also, from that contradiction it is not the case that every formula follows, so the contradiction does not lead to explosion. As a result, the logic is paraconsistent, and more, given that one is dealing specifically with the concept of truth, it must be accepted that some propositions are true and not true, which, on a conception of a transparent notion of truth, leads us to the falsity of the proposition. That is, some propositions are true and false.<sup>5</sup>

So, here we are again: we have a derivation leading to a contradiction, which is also said not to explode. This time, however, the derivation is on the vernacular, using resources of natural language itself, and Priest claims, in many distinct places, that it is the logic of natural language that leads us to a contradiction. Priest, Berto, and Weber (2018) sec. 3.2, point clearly to that fact:

“Overall, such paradoxes as the Liar provide some evidence for the dialetheist's claim that some contradictions are *provably* true, in the sense that they are entailed by plain facts concerning natural language and our thought processes. Extended Liar paradoxes like ‘This sentence is not true’ are spelt in ordinary English. Their paradoxical characteristics, dialetheists argue, are due exactly to the intuitive features of ordinary language: unavoidable self-reference; the failure of metalinguistic hierarchies, which only produce languages that are expressively weaker than English; and the obvious presence of a truth predicate for English, ‘is true’, which is characterized, at least extensionally, by either the Tarskian T-schema or rules amounting to the transparency of truth” (Priest – Berto – Weber 2018).

---

<sup>5</sup> See Priest (2006) and Priest (2006a) for details.

The same point was raised by Priest before:

“We have seen that there are excellent reasons driving us to the conclusion that English is semantically closed. This means that there are true contradictions in English, sentences such that both they and their negations are true. However, obviously not all English sentences are true. (In fact only a minute fraction of sentences of English are paradoxical.) So it follows that the classical rule of inference *ex falso quodlibet* ( $A \wedge \neg A/B$ ) is invalid. There are cases where the premiss is true and the conclusion is not. In short, the underlying logic of English is paraconsistent. Moreover any adequate account of the semantics of English will have to face semantic closure and the existence of contradictory truths. This is true of Davidson’s, Montague’s or any other account of the semantics of English. There are no problems here. The semantics of paraconsistent logics show exactly how true contradictions can be handled” (Priest 1984, 128).

Once that is given, the question, naturally, is: is that derivation evidence for a change from the received logic to a paraconsistent logic? Well, it all depends, again, on how the evidence is identified and profiled. The bold-faced **not** in the derivation is certainly represented by a negation in systems of logic, but one must ask again, which theory of negation should we use to characterize this evidence? One cannot evade the question, given the supposition that natural language does have a working logic. So, if evidence is characterized and evaluated in terms of the received logic, then, clearly, the derivation must be considered a derivation of a contradiction, and an explosive contradiction that does lead to triviality. Seen through these lights, it is simply wrong to say that the contradiction derived is not explosive, because it is. In this sense, then, the derivation cannot be seen as evidence for a shift to a paraconsistent logic. In other words: the datum presented, when seen in the terms of the received logic, simply speaks against any claim that the derivation does not lead to triviality (and, of course, the classical logician provides for an explanation of how one can live with a contradiction without going trivial)<sup>6</sup>.

Still from the point of view of the received logic, which incorporates its own account of the formal behaviour of truth, there is never evidence for a case that a proposition is true and still have a true negation (or, a proposition being true and false). This happens because of the behaviour of negation in relation to the truth values (at least on what concerns the theory of truth advanced as an intended notion of truth for classical logic): whenever one proposition receives one of the truth values, its negation receives the other, and no proposition receives both. So, talk of true contradictions

---

6 See for instance Michael (2016) and Steinberger (2016).

as being pointed to as the result of the derivation is pointless if that is the background logic according to which the evidence is analysed and evaluated. The derivation is a derivation of a false contradiction (every contradiction is false in classical logic with its intended semantics), and we are recommended to reject at least some step leading to it. Let us say it again: whenever one allows that the background logic in terms of which evidence is presented is classical logic, the derivation is not evidence *against* classical logic. One simply cannot see the Liar as a derivation of a true contradiction if the logic is assumed to be classical logic.

But what if we change the character of the negation employed to identify and characterize the evidence? Once again, one could claim that, to be fair to the evidence presented in favour of the theory of negation and truth advanced by Priest (and dialetheists in general), the derivation should be examined in terms of his favourite theory of negation (that is what Routley seems to suggest, when claiming that a uniformity in language and metalanguage should be adopted). Notice that this would require that the background logic should allow for true contradictions, and in particular, if it is to be completely fair, it must allow that some truth-bearers receive true and false as their truth-values. Now, if that is going to be allowed, of course, truth and falsity, as a truth-value *glut*, must be available in order to be attributed to some truth-bearers. The theory of truth must allow that truth and falsity sometimes overlap, and that the negation of a sentence that is true and false is also true and false. That background apparatus would allow that some sentences are true and false at the same time, so that the derivation of the Liar would be indicating that sometimes derivations end up in gluts.

The problem is that, once the logical apparatus is characterized in those terms, the evidence presented is accounted for by a theory that should be precisely getting support from the evidence. However, that is not what happens, then: what we are doing is introducing the theory in the background logic in order to characterize the evidence, and then, claiming: 'look, the evidence points to that theory as the correct one'. Certainly, that cannot be helpful if the evidence should guide us in logical theory revision. Using the dialetheist framework to characterize the evidence amounts to what in other contexts is characterized as a self-fulfilling prophecy: if we accept that the logical apparatus behaves in a certain way, then, some selected derivations may be seen as corroborating the selection of the apparatus. Once again, we only take from the evidence the theoretical content we put there to begin with. This is hardly a case in favour of a change from classical to paraconsistent logic, and also hardly a case in favour of a change of the theory of truth from classical to dialetheist.

Martin (2020) has analysed Priest's arguments using the Liar in favour of a change of logic, with the accompanying change in the theory of truth. Priest argues not only in terms of the derivation of the Liar and related paradoxes of self-reference, but also in terms of what kind of desiderata a theory accounting for the paradoxes should satisfy. Among them, we could list that the theory should preserve what is taken to be intuitively valid inferences, and allow for semantic closure (because English, as a natural language, does so). Could these further considerations be brought to the overall balance and favour the dialetheist against the classical logician? As we have been arguing, when we stick to the most important part of the methodology of logical choice, the one concerning adequacy to the data, then, the very idea that the data must be framed in specific logical terms, and that this framing prevents theory change, because it either does not allow one to see it as evidence against the adopted logic, or because it is difficult to see it as favouring a framework that is put by hand there to lead to the expected result, then, discussion of extralogical factors seem to lose part of the interest. The difficulty of accounting for the data seems to be mining the rest of the discussion. These further considerations, it seems, would also not be based on evidence, but on other non-evidential factors<sup>7</sup>.

But, even if one were to enter into such debates with Priest on the desiderata a satisfactory logical theory should satisfy, it seems that the discussion will not favour his approach, given that these desiderata involve conditions that are, some of them, already logically-oriented, and similar problems could arise. Consider the claim that a logical theory should preserve some obvious claims of natural language, such as the fact that in natural language we make attributions of truth values and have a semantically closed language. What is a classical logician to make of it? Certainly, there is a disagreement here on the aims of a logical theory. The classical logician simply does not buy into that goal as a most important one, and takes derivations such as the Liar as evidence that natural language is inconsistent (in the classical sense), and that some further solution is needed.<sup>8</sup> Why should a classical logician preserve that? Obviously, here there is infiltration of the theoretical apparatus of the classical logician in judging the further stated desiderata of a logical theory, just as Priest had his own desiderata in light of his own apparatus. That is, in the measure that such desiderata involve logical concepts, they are also judged from

---

<sup>7</sup> Which ones, is a matter for discussion elsewhere, but see Hjortland (2019) for the difficulties of using non-evidential factors in logical theory selection.

<sup>8</sup> See, for instance Scharp – Shapiro (2017) for a discussion on conceptual revision involving truth in the light of paradoxes.

the theoretical apparatus of the received logic (for the classical logician), and they do not gain much purchase from that point of view. Something similar, of course, could be said from Priest's point of view. Once a dialetheist background logic is assumed, these desiderata are fulfilled, and gain prominence. But this only points to the fact that, when seen from the privileged position of the theory to be adopted, the evidence speaks in favour of it.

#### 4. Diagnosis and a Way Out

Perhaps it is already time for a diagnosis of what has been discussed, and to locate what we believe to be the source of the difficulties that we have highlighted. Recall that the problems we found in the evaluation of evidence seems to suggest the following: in cases where the alleged evidence for logical change involves logical vocabulary, the assumption that we do operate with a definite logic in natural language (even if with the wrong logic, for the time being) seems to require that the logical vocabulary present in the evidence should be understood as operating according to the prescriptions of the logic we already are supposed to have and use. In such cases, the examples we have studied indicate that hardly the evidence presented can count as evidence in favour of an alternative logic that is being proposed as a better account of the evidence. Basically, the evidence must be understood in terms of the already adopted logic, and it is precisely this reading of the evidence that prevents that the evidence should count against the logic we are said to have: how could it, given that the evidence behaves precisely as prescribed by our current logic? That leads to no possibility of recommendation of change of logic in terms of the evidence available, given that the evidence suits perfectly well the logic we do have. When one changes the perspective and adopts the challenging logic as the preferred framework according to which such a vocabulary should be understood, then, again, the data do favour the new candidate logic, but only because they behave as described by a logic put there by hand, and then, cannot be counted as evidence in favour of the new logic.

The source of the problem should be clear by now: the trouble is being caused by *the claim that we do have a logic operating on natural language, which is also the source of the profile attributed to the evidence, and which describes the behaviour of the logical apparatus present in the evidence.* This supposition, Priest's 'the logic we have', (which appeared under the heading d) in the Introduction) vitiates the information coming from natural language by conferring a well-defined theoretical character to the data beforehand, and by doing so, it prevents that evidence could count

*against* the adopted logic, and also prevents that the evidence could count in *favour* of a rival logic. Our proposal, then, is that it is this claim that should be given up, leading us to embrace a form of *logical nihilism* (see also Cotnoir 2018): natural language does not have a well-specified logic.

Logical nihilism does not mean abandoning the very idea that a system of logic can be chosen for given purposes, and that one of them may be better suited to deal with the evidences than others. That also does not mean that there is no background logic. That is, an anti-exceptionalist may be a logical nihilist. From the point of view of a logical nihilist, issues a) – c) may be settled, and with advantages, when the idea of a logic in natural language is disposed off. The remaining of this section will be a proposal, but it tells, in broad strokes, how a conception of logic and logical choice may go on, what avoids the difficulties that were presented earlier, by adopting logical nihilism in an anti-exceptionalist context (certainly, logical nihilism requires further motivation of its own, but we trust that the kind of problem presented here can be avoided by the nihilist, and this is no small deal).

Let us begin by making clear what we are giving up by abandoning the claim that natural language has a logic of its own, whose characterization is thought to be the goal of logicians. What is being left behind is the very idea of logic as an absolute *canon of reasoning*; we also abandon the idea that the aim of the activity of logicians is attempting to find out something that is already there ‘in the wild’, the idea that there is a notion of ‘validity simpliciter’. Woods has attempted to characterize this view of logic:

“I focus on disputes about which logic to adopt as our most basic canon of logical implication. There is little to no dispute about whether we can adopt distinct logics for instrumental purposes. Nearly all cases of situation-specific reasoning, such as drawing out local commitments from potentially inconsistent data sets, can be treated as instrumental applications of formal methods; the most interesting questions about logical revision focus on our most general canons of implication, such as which logic should we take as the background logic in which to evaluate logical relations between propositions. My discussion should be understood accordingly” (Woods 2019, 1204).

Rejecting thesis d) leaves us only with what Woods calls ‘instrumental applications’, or, perhaps, it demystifies the canonical application, by bringing it to the exact level of just another application, to where it belongs. To see logic as mathematical theories applied for a purpose is a key notion here, and one that substitutes the image of the search for the true logic capturing something out there, once the nihilist picture enters the stage.

This is a first step in the anti-exceptionalist sketch we wish to advance; it is very much in touch with the view of logic as a model, where the features of the model depend very much on our purposes for using the model (see Shapiro 2014).

The second ingredient is a naturalistic approach, more oriented by the practice of logicians. Martin suggests this strategy, but with restricted purposes (Martin 2020). From a naturalistic perspective, logic, as a science, can only be judged from the point of view of science itself, and not by a first philosophy. This means that the practice of logicians must be seen as a reference for our philosophizing about logic, and not our philosophizing about logic as a guide for our practice of logic (logicians don't typically await for a philosophical authorization on which systems are appropriate to be developed). In this sense, the relation between evidence and logical systems, as seen from the practice of logicians, may be seen as follows: what logicians do is to select some argumentative patterns or *salient features of argumentative contexts* (such as the quantity and nature of truth-values involved, for instance, or peculiar behaviour of logical operators, or specific inferential patterns to be held as valid) they are interested in codifying, and then, to propose systems that model such features. From a nihilist perspective, there is no question of truth of a system, but only of success of achieving the goals, which is how we judge such attempts. Also, that seems to be an accurate description of what logicians do when working.

Achieving goals as a standard of success of a logical system may be backed as follows. Logic, as part of science, is a rational activity, and as such, pretty much goal-oriented. As summarized by Bueno and da Costa:

*"The goal-oriented component.* One of the main features of rationality, especially when we consider the behavior of certain agents, is the fact that rationality is a goal-oriented activity. In this sense, a rational behavior clearly depends on the goals one may have. In the context of science, the goal involves the search for some sort of truth or regularity to make sense of experience" (Bueno – da Costa 2007, 386).

The goals one has in developing a system of logic will play a major role in theory selection and in selecting those features that count as evidence in favour or against choice of a given system. We restrict ourselves here to specific goals that are related with studying and codifying inferences that appear in natural language. Given natural language inferences, in the most varied contexts, such as classical mathematical inferences, inferences in contexts of vagueness, in contexts involving contradictions, contexts

involving lack of perfect information, future contingents, and so on, what logicians do is to offer *mathematical models* of how inferences may proceed in such cases, keeping some salient features of the practices that are most prominent as a major target of such formalization, depending on their goals (and this is just another way of describing Woods' claim of *instrumental application of formal methods*). We choose some basic features of the context that will behave in expected ways, and by doing so, narrow the field of application, idealizing and abstracting from other practices or features that are not related with our more specific aim.

Under this practice-oriented picture, we have indications of how to solve the problems a) – c), as we have been mentioning. To begin with, abandoning the natural language logic hypothesis, there are no vitiated data about logical consequence that one should attempt to capture in a system. *Goal fulfilment*, then, is a substitute for talk about the truth of a logical system, given that we no longer have anything else the system could be literally true of (also, we believe, claiming that a system of logic is *true* is a category mistake, given that truth is also a notion that must be developed inside a system of logic, but that is another issue). As put by Smith (2011), the concept of valid inferences is not to be found in natural language, it is typically heavily loaded with theory, and which theory we select depends very much on our goals and aims. If you think that there is notion of validity out there, he suggests, think again:

“If you think that there is, start asking yourself questions like this. Is the intuitive notion of consequence constrained by considerations of relevance? — do *ex falso quodlibet* inferences commit a fallacy of relevance? When can you suppress necessarily true premises and still have an inference which is intuitively valid? What about the inference ‘The cup contains some water; so it contains some H<sub>2</sub>O molecules’? That necessarily preserves truth (on Kripkean assumptions): but is it valid in the intuitive sense? — if not, just why not?” (Smith 2011, 29).

That is, the notion of validity is heavily theoretical, and depends on a set of previous choices about how to properly understand logical consequence; and these choices, we claim, depend on our aims, on the specific of the application (the ‘salient features of the application field’ we mentioned). That is, one always starts by selecting the aspects of practice that are held fixed and that are to be modelled, and doing that involves choosing how to articulate a most appropriate framework to deal with the context, in particular, describing the notion of truth involved, how many truth values are applied, the behaviour of logical connectives, and so on (as Priest has put it, logic is a major enterprise, requiring articulation of diverse concepts;

Smith's quote before just adds to the cart). The fixing of the free parameters depends on specific goals one may have in mind. These specifics are not to be found fully characterized in the practices themselves, but are selected as aspects deserving study, and once selected, they inform the conceptual apparatuses that we will use to address such practices. One judges then *the expediency of using a system to deal with a given aim*. Matters of dealing with the evidence, then, are not to be understood in terms of truth and falsity, but of choice of the most appropriate framework to deal with some stated goals.

This is perhaps perfectly illustrated by classical logic and its relation with its original stated goal: capture of inferences valid in classical mathematics. Classical logic was born as applied mathematics, as the mathematical study of mathematical reasoning, in the hands of Frege, Russell and Whitehead (even though they would certainly not agree with this way of putting the issue). Given the context of application it was supposed to deal with (mathematical reasoning), one has some clear salient features and parameters that must be part of the model, and others that cannot: one can, for instance, abstract from tenses of verbs, given that mathematical objects do not change with time, and one may also dispense with many features of language that are not relevant for the aim in mind, such as cases involving incomplete information, vagueness of expressions, and so on. The result is in large measure successful for that goal, with most people believing that classical logic, involving material conditional, does capture reasonably well the kind of inference that classical mathematicians use. Notice that this does not involve, so far, any judgment on the adequacy of the resulting theory to account for inferences outside of that field of application. Some may even believe that a single model must account for all such distinct contexts, but this is an issue that needs not concern us here.

Something similar may be said about paraconsistent logics. When a logician decides, for some reason or purpose, that some contradictions may be tolerated, and that a theory where contradictions appear is worth studying with the presence of such a contradiction, the salient features chosen as facts and data do comprise the permanence of such contradictions. Adjusting the details leads to distinct paraconsistent systems, but the fact is that selection of what is relevant, what one has as a goal as inferences that seem worth preserving for theoretical studies, guide the choice of a system, and even the reading of the data. After the development of paraconsistent logics, one can choose to consider Russell's paradox as a good fact to be accounted for, as something that should not be eliminated by logical reasons. However, when it first appeared, there was only one way to regard the paradox, and it required a reformulation of the theory to

preserve consistency. Now, the naturalist about logic can allow that study of both options (to keep the debate between paraconsistent and classical logicians only) is fruitful, and that the derivation, in itself, does not privilege any of them. It does not point to any options being more correct, because there is nothing there to a system of logic be correct about. It is a matter of presenting models that are adequate to our aims, and our aims sometimes involve preserving Russell's set, other times, not. If there were a fact of the matter as to whether Russell's set is allowable or not, one of the options would be wrong, and affirming this, at least in the current state of development of logic and foundations, means to judge the practice of logic from a privileged point of view that no one has.

This, then, is a first stab on the relation between theory and evidence which seems to cover pretty well what logicians do on their practices when *selecting* data to be modelled by a formal system. But what about *the background logic problem*? Once we reject the claim that there is a logic in natural language, the claim that a logic must be in the background is not problematic. From the naturalist perspective we are advancing here, the background logic problem gets solved in the following way: given that systems of logic are seen as models of determined inferential practices, as applied mathematics, we develop these mathematical systems, as every logician, inside some set theory. As situated agents, we use the mathematics we have: an informal version of ZFC, or something to that effect, most of the times. There may be reasons for changing the general framework, but a naturalist will be happy with following the practice, and developing the theory of formal languages, formal semantics, and formal proofs, inside such a comfortable framework (with methods of proof such as proof by induction, *reductio ad absurdum*, and many others, available, as well as a well-known theory of cardinal numbers, among other commodities). Does this involve assumptions that some logicians would not agree with? Certainly, but then, it is a matter of them claiming that our practice is wrong, leading us to *a revisionary approach* to mathematics and logic. It is up to them, anyway, to suggest changes in the practice and convince the scientific community that a more fruitful approach is available, something that is a hard task these days, given that it is in tune with classical mathematics, which is the basis of our contemporary science.

Notice what this answer is not: it is not the claim that whenever one is trying to model some informal inference appearing in natural language by the use of a formal system, that informal inference is already loaded with classical logic. By abandoning the natural language logic hypothesis, we are free to select something we believe is an important behaviour of a logical particle in natural language and fix it as a pattern to be codified (such

as failures of excluded middle when dealing with contexts involving imperfect information, let us say), leaving other aspects aside (in many cases, for instance, tense of the verbs is not important, and may be left aside). This is a natural part of the modelling strategy in science in general, and in logic in particular, and one just develops the model inside classical set theory, as any logician would, and attempts to highlight the adequacy of the model, as we have discussed earlier in this section.

Certainly, this approach suits also the developments of systems of logic that are traditionally seen as requiring a revision in classical logic in the background logic, such as intuitionists and dialetheists. By using a common set theory as our metalanguage we are able to discuss the relevant evidence and the use of determinate systems as appropriate models. Even intuitionistic logic and dialetheistic logic may be developed, and typically are, inside classical set theory. Those claiming that this is wrong, and that the correct development requires a non-classical language right from the start are not being naturalists, but rather requiring the correction of our current practices. They typically see natural language as having a right logic, which is then used to judge on the evidence, falling on the problems of logical choice we have already discussed.

Of course, that does not mean that the current situation will not change. Alternative set theories may be required for some purposes deemed worthwhile, and, one may envision future revisions of our logical background. The very idea of using a set theory in order to develop a system of logic is a recent research program, appearing only in the middle of the twentieth century. Scientific practice is not frozen in time, not immune to the idiosyncrasies of history, and the same happens to logic. However, in so far as logic is integrated to mathematics (as it now is), it may be seen as a branch of mathematics, and it is part of the enterprise of current developments under a set theory. One may opt for a non-naturalistic approach to logic, but if one is a naturalist, this kind of practice will have to play an important role.

## 5. Conclusion

In this paper, our efforts were focused on dislodging a frequently occult premise in discussion of logical choice: that we need a canon of inference, a logic of natural language, and that it is this logic that guides our evaluation and identification of evidence when it comes to logical theory choice. As we have argued, this natural language logic ends up infecting the way that the evidence is framed, by impregnating the data with the theory we are said to adopt. The result, we have claimed, is that logical theory change

is made impossible in the light of evidence, if this hypothesis is assumed. Basically, we have done that by discussing two case studies involving paraconsistency. Paraconsistent logicians typically use alleged cases of inferences involving contradictions without triviality as evidence that our canon of inferences should be paraconsistent, and not classical. If the assumption that there is a background logic holds, and evidence is framed in terms of the background logic, then, two cases appear: if the background logic is assumed to be classical logic, then, there is no evidence available for a paraconsistentist: every derivation of a contradiction will necessarily lead to triviality, and the data is not evidence for paraconsistent logic. On the other hand, if evidence is framed in terms of the paraconsistent logic, then, there is nothing else to judge, the data cannot count as evidence in favour of the paraconsistent logic, given that the logic has already been assumed as paraconsistent.

We have also advanced an alternative picture of logical theory choice in the absence of the problematic assumption. The picture is a logical nihilist version of anti-exceptionalism, wedded to naturalism in logic. Once the goal of providing for a description of the logic of natural language, using one logic of natural language as the background logic is abandoned, a coherent picture emerges. Logic is seen as providing models for inferential practices. Given that such practices do not codify by themselves any logical theory, one selects some relevant aspects of it and provides for some mathematical models of those selected aspects. One judges not the truth of a model, but its expediency in order to achieve some goal. This is done all the time, and seems to be a platitude when we check what logicians really do. This picture was coupled with a naturalist view of the background logic: we just use the mathematics we have (and that means, broadly, ZFC), as the metalanguage, in which formal systems are developed and studied. This is clearly in accordance with the practice of logicians, and this speaks in favour of the story we have told.

## References

- BUENO, O. – da COSTA, N. C. A. (2007): Quasi-truth, Paraconsistency, and the Foundations of Science. In: *Synthese* 154, 383 – 399.
- CARNIELLI, W. – RODRIGUES, A. (2015): On the Philosophy and Mathematics of the Logics of Formal Inconsistency. In: Béziau, J.-Y. – Chakraborty, M. – S. Dutta (eds.): *New Directions in Paraconsistent Logic*. New Delhi: Springer, 57-88.
- COTNOIR, A. J. (2018): Logical Nihilism. In: Kellen, N. – Pedersen, N. J. L. -Wyatt J. (eds): *Pluralisms in Truth and Logic*. Palgrave Macmillan, 301-329.
- da COSTA, N. C. A. – KRAUSE, D. – BUENO, O. (2007): Paraconsistent Logic and Paraconsistency. In: *Handbook of the Philosophy of Science. Volume 5: Philosophy of Logic*. Amsterdam: Elsevier, 655-775.

- HJORTLAND, O. T. (2017): Anti-exceptionalism about Logic. In: *Philosophical studies*, 174, 631-658.
- HJORTLAND, O. T. (2019): What Counts as Evidence for a Logical Theory? In: *Australasian Journal of Logic*, 16 (7), 250-282.
- MARTIN, B. (2020): Identifying Logical Evidence. Forthcoming in *Synthese*.
- MICHAEL, M. (2016): On a "Most Telling" Argument for Paraconsistent Logic. In: *Synthese* 193, 3347-3362.
- PRIEST, G. (1984): Semantic Closure. In: *Studia Logica* 43, 117-29.
- PRIEST, G. (2006): *In Contradiction: a Study of the Transconsistent*. 2. edition. Oxford: Oxford University Press.
- PRIEST, G. (2006a): *Doubt Truth to be a Liar*. Oxford: Oxford University Press.
- PRIEST, G. (2016): Logical Disputes and the a priori. In: *Princípios* 23 (40), 29-57.
- PRIEST, G. – BERTO, F. – WEBER, Z. (2018): Dialetheism. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy (Fall 2018 Edition)*. Web. 13. 12. 2020 <<https://plato.stanford.edu/archives/fall2018/entries/dialetheism/>>.
- QUINE, W. v. O. (1963): Two Dogmas of Empiricism. In: *From a Logical Point of View*. In: *9 Logico-Philosophical Essays*, 20-46. 2. edition. New York and Evanston: Harper and Row.
- ROUTLEY, R. 1980. The Choice of Logical Foundations: Non-classical Choices and the Ultralogical Choice. In: *Studia Logica* 39 (1), 77-98.
- SCHARP, K. – SHAPIRO, S. (2017): Revising Inconsistent Concepts. In: Armour-Garb, B. (ed.) *Reflections of the Liar*, 257-280. Oxford: Oxford University Press.
- SHAPIRO, S. (2014): *Varieties of Logic*. Oxford: Oxford University Press.
- SMITH, P. (2011): Squeezing Arguments. In: *Analysis* 71, 22-30.
- STEINBERGER, F. (2016): Explosion and the Normativity of Logic. In: *Mind* 125, 385-419.
- WOODS, J. (2019): Logical Partisanship. In: *Philosophical Studies* 176, 1203-1224.

*The production and revision of this paper benefited from a Research Fellowship at the Institute Vienna Circle, at the University of Vienna, Austria, and also from a Capes-Humboldt Experienced Researcher Fellowship, held at the Ruhr University – Bochum, Germany.*

**Jonas R. Becker Arenhart, PhD.**

Federal University of Santa Catarina

Department of Philosophy

Campus Universitário Reitor João David Ferreira Lima, s/n

Trindade – Florianópolis

Santa Catarina – Brazil

88040-900

e-mail: [jonas.arenhart@ufsc.br](mailto:jonas.arenhart@ufsc.br)

Orcid: <https://orcid.org/0000-0001-8570-7336>