



Intuitions, theory choice and the ameliorative character of logical theories

César Frederico dos Santos^{1,2} 

Received: 27 October 2020 / Accepted: 21 July 2021 / Published online: 2 August 2021
© The Author(s) 2021

Abstract

Anti-exceptionalists about logic claim that logical methodology is not different from scientific methodology when it comes to theory choice. Two anti-exceptionalist accounts of theory choice in logic are abductivism (defended by Priest and Williamson) and predictivism (recently proposed by Martin and Hjortland). These accounts have in common reliance on pre-theoretical logical intuitions for the assessment of candidate logical theories. In this paper, I investigate whether intuitions can provide what abductivism and predictivism want from them and conclude that they do not. As an alternative to these approaches, I propose a Carnapian view on logical theorizing according to which logical theories do not simply account for pre-theoretical intuitions, but rather improve on them. In this account, logical theories are ameliorative, rather than representational.

Keywords Abductivism in logic · Predictivism in logic · Carnapian explication · Ameliorative analysis · Logical intuitions

1 Introduction

In recent years, the view that “logic is in the same epistemic boat as other scientific theories,” as Bueno and Colyvan (2004, p. 156) put it, is becoming increasingly accepted among philosophers of logic. To contrast with the previous orthodoxy, which regarded

This article belongs to the topical collection “Anti-Exceptionalism about Logic”, edited by Ben Martin, Maria Paola Sforza Fogliani, and Filippo Ferrari.

This work was supported by FAPEMA (Foundation for Research and Scientific and Technological Development of Maranhão) under Grant Number BD-08292/17.

✉ César Frederico dos Santos
c.f.dos.santos@vu.nl ; cesar.frederico@ufma.br

¹ Department of Philosophy, Vrije Universiteit, Amsterdam, Netherlands

² Department of Philosophy, Federal University of Maranhão, São Luís, Brazil

logic as *exceptional* in that the epistemic justification of logical theories would require evidence and methods other than those commonly used in other sciences, this view was named “anti-exceptionalism about logic” (Hjortland 2017; Williamson 2007). Once logic is equated in epistemic terms with other sciences, wherein observation and experimentation constitute the main sources of evidence, the question of what counts as evidence for logical theories becomes pressing. In response to this question, a number of philosophers have proposed guidelines for theory choice in logic (e.g., Bueno and Colyvan 2004; Resnik 2004; Priest 2016; Williamson 2017; Martin and Hjortland 2020).

Two of these accounts are abductivism (Priest 2016; Williamson 2017) and predictivism (Martin and Hjortland 2020). In these accounts, logical theories are viewed as intended to provide a faithful description of pre-theoretical logical intuitions. Accordingly, conflicts between verdicts about validity given by a logical theory and verdicts about validity given by intuitions are seen as *prima facie* reasons to reject the former. Despite the importance these accounts attribute to intuitions, there is little clarity about what logical intuitions are and why, or whether, they are reliable.

I start this paper by investigating the prevalence and the reliability of pre-theoretical logical intuitions. After outlining abductivism and predictivism as methods for theory choice in logic in Sect. 2, in Sects. 3, 4 and 5 I draw on results from the psychology of reasoning and on Dutilh Novaes’ (2021) Prover–Skeptic dialogical model of deduction to argue that truly pre-theoretical intuitive judgments about validity are *less reliable* than judgments about validity informed by a logical theory. If this is so, intuitions should not count as a reason to reject logical theories, contrary to what abductivism and predictivism assume.

Given the unreliability of pre-theoretical intuitions, what seems more plausible, I submit, is that pre-theoretical intuitions constitute only a starting point for logical theorizing. Logical theories do not simply capture pre-theoretical intuitions, but rather *improve* on them. In other words, logical theories are *ameliorative*, rather than representational. In the two final Sects. (6 and 7) of this paper, I draw on Carnap’s (1950, 1963) view of scientific and logical theorizing as *explicative* and on Haslanger’s (2012) account of ameliorative philosophical analysis, accounts in which theory choice is *pragmatic*, to argue that what determines theory choice in logic are, first, the investigative aims of the theorist and, second, data about logical theories themselves (e.g., meta-theorems) showing whether a given theory can satisfy the pursued aims.

As in most anti-exceptionalist accounts, here I take logical theories to be theories of validity. However, here logical theories do not *model* a pre-theoretical notion of validity; rather, they lay down *explications* (in Carnap’s sense) of validity that may be more or less satisfactory depending on one’s investigative aims. The result is a yet anti-exceptionalist account of logical theorizing, but one that sees logical theorizing as similar to scientific theorizing not because both are representational undertakings, but rather because both improve our thoughts and practices, often flying in the face of our pre-theoretical intuitions.

2 Abductivism and predictivism

Priest (2016) and Williamson (2017) have advanced similar accounts of theory choice in logic. Although they disagree on the outcome of the application of the criteria they propose, they agree that logical theories should be chosen on the basis of abductive arguments, i.e., inferences to the best explanation. On their accounts, logical theories compete as candidate explanations of a given phenomenon. The best explanation is the one that satisfies a number of theoretical virtues to the highest degree, such as simplicity, strength, unifying power, and adequacy to the data. This latter criterion—adequacy to the data—is particularly important, since it relates the proposed explanation to its target phenomenon. For Priest, the target phenomenon of logic is the notion of validity. Validity, however, is not an observable phenomenon in the same sense that the phenomena studied by physics or chemistry are, which poses a difficulty for Priest from the outset.

In the criterion of adequacy to the data, what counts as data? It is clear enough what provides the data in the case of an empirical science: observation and experiment. What plays this role in logic? The answer, I take it, is our intuitions about the validity or otherwise of vernacular inferences (Priest 2016, p. 41).

The gist of Priest’s answer lies in the words ‘intuition’ and ‘vernacular.’ His claim seems to be that we have pre-theoretical intuitions about validity,¹ on which we rely for making inferences in ordinary situations. These everyday inferences are conducted “in the vernacular,” rather than in a formal language; “maybe the vernacular augmented with a technical vocabulary (such as that of chess, physics or whatever); maybe the vernacular augmented with mathematical apparatus; but the vernacular nonetheless” (Priest 2006, pp. 169–170). The pre-theoretical notion of validity that guides vernacular inferences constitute, according to Priest, the data against which competing logical theories should be tested. Some inferences conducted in the vernacular “strike us” as valid, whereas some others “strike us” as invalid. “Any account that gets things the other way around is not adequate to the data” (Priest 2016, p. 42).

Williamson has a different view about the target phenomenon of logical theories. For him, logical theories are intended to capture the most general aspects of the world. Accordingly, he accepts that empirical evidence can confirm or disconfirm logical theories, although indirectly. Roughly, Williamson claims that competing logical theories A and B can be “empirically tested” in the following way: take a set of well-confirmed scientific sentences Γ ; derive all its empirical consequences according to A , generating the set of sentences Γ_A , and according to B , generating the set of sentences Γ_B ; finally, compare how well Γ_A and Γ_B fit with empirical data (Williamson 2017, p. 334). However, Williamson also admits other sources of evidence, not related with predictions deduced from scientific theories: “[e]vidence here is not confined to observations. We

¹ As will be made clear in Sect. 5, here I take pre-theoretical logical intuitions to be judgments about which propositions follow from other propositions in informal arguments. Such judgments usually come about without explicit reasoning. The capacity to make them comes from the maturation or acquisition of basic cognitive skills, such as language usage, or through experience with a specialized kind of argumentative practice Dutilh Novaes (2021) calls “Prover-Skeptical dialogues.”

may use anything we know as evidence” (Williamson 2017, p. 335). In particular, we can recruit pre-theoretical intuitions.

For example, in the case of propositional modal logic, we may know that the coin could have come up heads, and could have not come up heads, but could not have both come up heads and not done so, and on that basis eliminate this proposed law: $(\diamond p \ \& \ \diamond q) \rightarrow \diamond(p \ \& \ q)$. By contrast, this law identifies a useful pattern in the modal data: $(\diamond p \ \vee \ \diamond q) \rightarrow \diamond(p \ \vee \ q)$. In that sense, we can verify some predictions of the law by using our pretheoretic ability to evaluate particular modal claims (Williamson 2017, p. 336).

He claims that this sort of “pre-theoretical modal knowledge [is] accessible to almost any reasonable, intelligent person” (Williamson 2013, p. 427). What is more, he claims that pre-theoretical modal knowledge is much more relevant than evidence coming from natural science, at least when it comes to the study of modal logic.

Although nothing in theory precludes the application of results from any branch of natural science to the present enquiry [i.e., the study of modal logic], we have seen little evidence that they would be of much help in practice. It would hardly be relevant to carry out special experiments or make special measurements. A combination of logico-mathematical reasoning with *elementary modal knowledge* in particular cases turns out to be far more useful (Williamson 2013, p. 423, emphasis added).

Thus, for Williamson, as for Priest, pre-theoretical intuitions (or pre-theoretical “knowledge,” as Williamson calls it) have a central role to play as evidence for the acceptance or rejection of logical laws. The same view is shared by many other philosophers. For example, speaking of “the classical research tradition in logic,” Bueno and Colyvan say that “[t]he aim of logic is taken to be to provide an account of logical consequence that captures the *intuitive notion* of consequence found in natural language” (Bueno and Colyvan 2004, p. 168, emphasis added). This echoes Tarski, who notes that “in making precise the content of this concept [logical consequence], efforts were made to conform to the everyday ‘pre-existing’ way it is used” (Tarski 2002, p. 176). Resnik elaborates on this point:

When it comes to describing our inferential practice or the reasoning used in some branch of science or mathematics, logicians, *like empirical linguists*, try to achieve the best systematization of their data that they can ... Here we find logicians relying both upon data concerning our inferential practice and their *intuitions*—both normative and metaphysical—concerning the facts of logic (Resnik 2004, pp. 180–181, emphases added).

In sum, the idea seems to be that logical theories are aimed at capturing either pre-theoretical intuitions directly or capturing the features of an underlying phenomenon these intuitions convey information about. Either way, logical theories are thought of in representational terms: they *describe* a given phenomenon.

In an attempt to capture this widespread trait of logical practice—the view that intuitions about the validity of inferences expressed in ordinary language count as data for logical theories—Martin and Hjortland (2020) have recently advanced a model

of anti-exceptionalist logical methodology that they call “logical predictivism.” On their view, a logical theory is evaluated with respect to its ability to make successful predictions. In contrast with Williamson’s account, though, in Martin and Hjortland’s account the relevant predictions do not concern empirical consequences inferred from sets of well-confirmed scientific sentences, but rather predictions about which “vernacular arguments,” to use Priest’s term, should be counted as valid according to the logical theory at stake. If the theory predicts that a certain argument will be judged valid by competent reasoners and they do judge it so, this counts as confirmation for the theory; otherwise, this counts as disconfirmation.

On their account, logical theorizing involves two steps. For example, a logician who is concerned with the identification of the logic underpinning informal mathematical proofs would start by identifying which kinds of inference are judged correct by mathematicians. If the logician identifies the use of, say, Modus Ponens in some informal mathematical proofs mathematicians judge correct, this counts as *prima facie* evidence for taking Modus Ponens as a valid logical rule. The second step is testing the hypothesis that Modus Ponens is valid. Since this hypothesis allows for the prediction that the occurrence of Modus Ponens in other informal mathematical proofs will also be regarded as valid, additional mathematical proofs containing inferential steps based on Modus Ponens that are judged correct by mathematicians count as evidence for the maintenance of Modus Ponens as a logical principle; otherwise, i.e., if mathematicians saw these inferential steps as incorrect, this would count as counter-evidence for the general validity of Modus Ponens in mathematics. In sum, “what’s taken as a reliable indicator of validity, and thus suitable data to test the consequences of the theory, are the judgements of mathematicians regarding acceptable informal proofs” (Martin and Hjortland 2020, p. 13). Naturally, these judgments are not made by means of formal tools and axiomatized logical theories, but rather by recruiting their pre-theoretical intuitions about the validity of mathematical proofs.

Martin and Hjortland claim that the same approach can be applied to capture the pre-theoretical notion of “what follows from what” involved in everyday arguments. In this case, not only the judgement of mathematicians, but also the judgment of laypeople “over the correctness of arguments, or over whether some conclusion ‘follows from’ some premises, are treated as data and taken to be *prima facie* reliable indicators of validity” (Martin and Hjortland 2020, p. 16). They recognize, however, that, differently from the judgments of mathematicians, the judgments of ordinary people about validity may be unreliable.

After all, we are well aware from cognitive psychology of the unreliability of individuals’ logical reasoning under certain conditions (see, for an introduction, Evans [19]). Thus, it seems either we must admit that the proposed data is unreliable, or we need to pre-identify certain agents as reliable judges of which propositions follow from others in particular arguments (Martin and Hjortland 2020, p. 17).

They identify the difficulty but do not offer a solution to it, since they “only aim to show what sense can be made of a logical methodology that treats such judgements as reliable” (Martin and Hjortland 2020, p. 17). This is a difficulty not only for predictivism, but also for abductivism. If there are unreliable deductive reasoners, the

abductivist also has to pre-identify them and discount their judgments as data. As a result, a problem that both accounts have to tackle is that, in order to discriminate between reliable and unreliable judges, the logician has to have a prior notion about what counts as a valid inference. Only with this prior notion in place, can the logician discount some judgments as erroneous. In the next three sections, I discuss the difficulties of pre-identifying reasoners with reliable logical intuitions that could guide theory choice in logic.

3 Unreliable reasoners from the perspective of classical logic

The findings from cognitive psychology Martin and Hjortland mention in the quotation above belong to a tradition in psychology of reasoning that uses syllogistic and classical logic as its normative theories. Oaksford and Chater (2020) call this tradition the *logical paradigm* in psychology of reasoning. One of the most famous experiments of this paradigm is the Wason selection task. This task is designed to test participants' ability to reason about conditionals. In the original formulation of this task, participants are presented with four cards. Each card has a number on one side and a letter on the other. The cards are laid down on a desk so that only one of its sides is visible. For example, participants are shown the cards with *N*, *T*, *6*, and *8* written on the side facing up. Then, participants are asked to verify, by turning over all and only the relevant cards, the following conditional: "If there is an *N* on one side of the card, then there is a *6* on the other side." This conditional has the form 'if *p* then *q*' and, according to the rules of material implication, it is false only if *p* is the case and *q* is not. Therefore, participants reasoning according to the canons of classical logic should turn over the cards with *N* and *8* facing up. However, this is not what most participants do; they turn over only the card *N*, or the cards *N* and *6* (Elqayam 2018). In its original formulation with numbers and letters, less than 20% of participants give the logically 'correct' answer (Ragni et al. 2017). The conclusion is that people without training in logic have poor intuitions about material implication. Commenting on the consequences of this result for his claim that pre-theoretical intuitions are a reliable guide for theory choice in logic, Priest writes:

it needs to be said that the intuitions in question here need to [be] of a robust kind, purged of clear performance errors. As the literature on cognitive psychology shows, people make not only mistakes, but systematic mistakes, such as those involved in the Wason Card test. What makes these clear mistakes is that once the matters have been pointed out to the people concerned, they can see and admit their errors. Neither is this done by teaching them some high powered logical theory: it can be done by showing simply that they get the wrong results. The intuitions invoked in theory-weighting have to be steeled in this way (Priest 2016, p. 16).

Priest's classification of errors in the Wason selection task as "performance errors" seemingly invokes the Chomskyan distinction between *competence* and *performance* in linguistic contexts. According to this distinction, some mistakes people make when speaking—performance errors—may not reflect what they know about the language

they are speaking—their linguistic competence. Analogously, one could argue, as Priest seems to do, that the mistakes people make in the Wason selection task do not show that they do not know how to reason in accordance with material implication; it may be just that they have made performance errors when trying to apply what they know to this specific task.

However, there are reasons to suspect that mistakes in conditional reasoning cannot be fully attributed to performance errors. A way of experimentally distinguishing between performance errors and lack of competence consists in providing subjects with hints that are likely to prevent the competent ones from making performance mistakes. Markovits (1985) adopted this strategy to distinguish between the influence of performance and competence factors in conditional reasoning in adults. He observed that participants who failed in a very easy conditional reasoning task wherein no hint was provided also failed in subsequent tasks wherein hints were provided. Importantly, most of the participants who succeeded when hints were provided had also succeeded in the easier task without hints. He concludes: “[t]hese results are generally inconsistent with any purely ‘accidental’ or performance-base theory of incorrect conditional reasoning in adults” (Markovits 1985, p. 246). Furthermore, in the so called “new paradigm” in the psychology of reasoning (more on this below), “mistakes” in the Wason selection task are neither performance errors nor a sign of lack of competence, but adequate answers based on probabilistic reasoning (Oaksford and Chater 1994). This is not to deny that performance factors may affect conditional reasoning; the point is that competence with conditional reasoning in accordance with the canons of material implication does not seem to be a widespread trait in adults.²

In another experimental task typical of the logical paradigm in the psychology of reasoning, participants are asked to evaluate the validity of simple syllogisms. The experimental setting is simple: participants are presented with a syllogism and asked to judge whether its conclusion follows from its premises. Participants are explicitly instructed to assume that the premises are true, even if they know that the premises are false, and to evaluate the conclusion with respect to the given premises only. Typical of the logical paradigm, experimenters assume that, since syllogisms are presented in natural language, previous training is not required for participants to understand the task. In a study of this kind, Sá et al. (1999) presented participants with the following syllogism:

All living things need water.
Roses need water.
Roses are living things.

Asked to evaluate this syllogism, 68% of the participants judged it valid, although anyone trained in logic easily recognizes it as invalid. This result shows what every logic teacher perceives in the first days of a new class: syllogisms whose conclusion is believable tend to be judged valid (even if the conclusion does not follow logically from the premises), whereas syllogisms whose conclusion is unbelievable tend to be judged invalid (even if the conclusion follows logically from the premises). This

² It is worth noting that this issue is not settled yet, though. There are various models of the mental processes underlying conditional reasoning, and in some of them mistakes in conditional reasoning can be viewed as performance errors (see, e.g., Johnson-Laird et al. 2018).

result is confirmed in many other studies of the same kind. For example, in Evans et al. (1983), 71% of invalid arguments with believable conclusions were judged valid, whereas only 56% of valid arguments with unbelievable conclusions were correctly judged valid (i.e., in 44% of the cases, participants failed to recognize a valid argument because its conclusion was unbelievable).

Experiments like these have repeatedly confirmed that, when evaluating validity, people who are not trained in logic do not take into account only the information contained in the premises—as they should, according to the canons of deductive logic—but bring prior beliefs to bear in the completion of the task. In other words, they reason non-monotonically: prior beliefs, not contained in the premises, can defeat or vindicate an argument. This tendency was dubbed “belief bias,” and it was considered a *bias* under the assumption that deductive reasoning, particularly the rules of deductive reasoning codified in classical logic, should be the standards governing reasoning (Ball and Thompson 2018).

The influence of belief bias is not limited to the evaluation of validity. This tendency permeates many other aspects of our cognitive lives. It is closely related with *confirmation bias*, sometimes called *myside bias*, which is “a tendency [that people have] to find arguments that support their point of view, whether that means supporting a position they agree with or attacking a position they disagree with” (Mercier 2018, p. 404). In other words, just as we tend to judge positively arguments whose conclusion agrees with our beliefs (belief bias), we are also better at finding arguments to support our beliefs than to undermine them (myside bias). These two tendencies are obviously related: myside bias hinders the search for counterexamples that could invalidate an argument with a believable conclusion (i.e., a conclusion that is in line with one’s point of view), and hence invalid arguments with believable conclusions are likely to go unnoticed. By contrast, an argument with an unbelievable conclusion (i.e., a conclusion in tension with one’s point of view) will naturally elicit counterexamples, if there are any, and therefore it is much less likely that an invalid argument with unbelievable conclusion goes unnoticed. (In Evans et al. (1983), only 10% of invalid arguments with unbelievable conclusions were incorrectly judged valid.)

There are various explanations for the causes of belief bias available in the literature on psychology of reasoning (for a review, see Ball and Thompson 2018). Some of them are compatible with the hypothesis that belief bias and other reasoning “mistakes” are caused by performance factors, and that competence with deductive reasoning in accordance with the canons of classical logic is widespread. However, an explanation that is gaining traction in recent years holds that belief “bias,” after all, is not a bias, but a feature of human reasoning. What has been described as the “new paradigm” in psychology of reasoning has it that ordinary reasoning is best modeled by Bayesian probability theory (Elqayam 2018; Oaksford and Chater 2020). “What appear to be erroneous responses, when compared against logic, often turn out to be rationally justified when seen in the richer rational framework of the new paradigm” (Oaksford and Chater 2020, p. 305). According to the new paradigm, everyday inferences are probabilistic and, hence, are, and should be, “knowledge-rich”—i.e., based on background knowledge and therefore non-monotonic—since evaluation of probabilities requires previous knowledge of the domain the inference is about.

A simple instance of Modus Tollens, presented in Oaksford and Chater (2020), illustrates this point. If someone is told that “If John turns the key, the car starts” and “the car didn’t start,” she is more likely to infer that John in fact turned the key but the car had some problem, than that John didn’t turn the key. “This implication is based both on our understanding of relevant aspects of the world (cars do not start spontaneously, they are mostly immobile, etc.) and on the norms of conversation ... We infer the existence of a failed key-turning attempt; otherwise, the car not starting would not be worth mentioning” (Oaksford and Chater 2020, p. 305). Thus, not following what seems to be a Modus Tollens is completely rational in such a everyday situation. Certainly, in cases like this the conclusion will not follow from the premises with full certainty—there will always be the possibility that John really didn’t turn the key—but the point is that in everyday situations we rarely need full certainty.

[D]eductive reasoning is hardly ever instantiated ‘in the wild,’ so to speak, as it is at odds with the strong component of defeasibility in everyday reasoning. In most everyday circumstances, deductive reasoning is overkill: the point is not to infer with absolute certainty what follows necessarily from the available information, but rather what is likely to happen given the available information and a number of background assumptions (Dutilh Novaes 2021, pp. 20–21).

If non-monotonicity is a widespread property of everyday reasoning, monotonic deductive reasoning is not to be found everywhere. This means that, from the perspective of monotonic deductive logics, the logical intuitions of most people will count as unreliable. What can the abductivist and predictivist do in the face of this fact?

4 Obstacles in identifying reliable reasoners

The prevalence of apparently non-monotonic reasoning in laypeople (here, people who were not trained in logic nor mathematics) poses a challenge to both the abductivist and predictivist. In order to pre-identify the reliable reasoners who will provide data for them, they have (at least) four methodological alternatives. All of them face obstacles, as we will see.

A first alternative is to assume that, contrary to the first impression of non-monotonicity, in fact most people reason deductively. It is only that in everyday situations people usually respond to richer arguments than the ones that are explicitly presented to them. The first task, in this case, would be the identification of the implicit deductive arguments subjects are responding to, and then taking their judgments about these arguments as the basis for logical theorizing. The difficulty in this case is that the transformation of putative enthymemes into completely stated arguments requires the assumption, in advance, of a logic: in order to identify which premises have to be added so as to convert an enthymeme into a deductively valid argument, one needs to adopt a notion of validity. Since this notion has to be in place before any analysis of the data can start, its justification must rest on something other than the data.³ That is,

³ Certainly, for many enthymemes a very minimal notion of validity can do the job, but this minimal notion will not get us very far. It is unlikely to help decide, e.g., whether implication has to be relevant, whether contradictory premises trivialize, etc.

abductivists and predictivists would have to choose a logic on some other basis than abductivism and predictivism.

A second alternative is to discount the judgments of those people who reason non-monotonically, retaining as data only judgments that are in line with deductive canons. This would amount to the rejection of the majority of the possible data sample. If such a significant part of the data sample is discarded, what justifies the claim that everyday reasoning is a target phenomenon of logical theorizing? Furthermore, given that this rejection would be based on the view that deductive logics are normative for everyday reasoning, the abductivist and the predictivist would have to justify, in advance, why these logics are preferable. For this, they could not rely on the judgments of the reliable reasoners they have already selected, on pain of circularity. In other words, as above, they would have to explain their preference for deductive logics on some other basis than abductivism and predictivism.

A third alternative is to assume the non-monotonic judgments commonly found in everyday situations as data and try to devise a non-monotonic logic that could account for them. The difficulty in this case is that, by selecting non-monotonicity as a normative principle for correct reasoning in everyday situations, abductivists and predictivists would be eliminating from their data sample exactly that minority of cases where people draw indefeasible conclusions in accordance with reasoning practices that have proved to be fruitful in mathematical and scientific contexts. This could be seen as a kind of pluralism: perhaps everyday reasoning relies on a logic other than the one employed in mathematics and some sciences. But, in this case, why should it be *wrong* to apply a monotonic logic in everyday contexts?

A fourth alternative would be to reject laypeople's judgments altogether when it comes to the choice of deductive logical theories. This would amount to assuming a position like Dutilh Novaes's, according to whom "deduction is a term of art corresponding to practices belonging to niches of specialists [mathematicians, scientists, philosophers], rather than a basic building block of human cognition" (Dutilh Novaes 2021, Ch. 10, section 5, para. 5). If this is so, when it comes to the identification of reliable logical intuitions, only the intuitions of these specialists should matter. Martin and Hjortland consider this alternative:

It may be that logicians do indeed only take into account the judgements of perceived "reliable reasoners", whether this be logicians themselves, philosophers as a whole, or members of professions required to engage in detailed reasoning within their working lives, such as lawyers and scientists. This would certainly explain why logicians do not go in much for empirical studies (Martin and Hjortland 2020, p. 17).

If deductive reasoning is not a basic building block of human cognition, then the ability to reason deductively has to be learned. Therefore, only those people who were trained in certain practices, such as those of mathematics, logic, science, and philosophy, would have the intuitive ability to reliably identify validity.

This hypothesis is in line with Williamson's (2011) response to the critique that experimental philosophers raise against reliance on intuitions in analytic philosophy. Experimental philosophers have argued that intuitions are not a reliable source of evidence for philosophical hypotheses because it has been experimentally observed

that non-philosophers' intuitions about the thought "experiments" philosophers use to elicit intuitions vary across philosophically-irrelevant factors, such as one's cultural background and the order of presentation of questions (Machery 2015). In response to this, Williamson (2011) argues that professional philosophers differ from non-philosophers in that they deploy more reliable intuitions about thought experiments due to their professional training. "[P]hilosophical training substantially reduces the influence of the distorting factors, even short of total eradication" (Williamson 2011, p. 219). Regardless of Williamson being right about the reliability of philosophers in conducting thought experiments in general, certainly his point holds (almost trivially) with respect to the superior ability of logicians and philosophers of logic to recognize valid arguments due to the specific kind of training they receive. The findings from psychology of reasoning mentioned above show that laypeople's deductive-logical intuitions are unreliable, but surely those findings are not a reason to deem philosophers' and logicians' intuitions equally unreliable.

However, particularly with regard to the logical intuitions of logicians and philosophers of logic, there is a further problem. Their intuitions are not *pre-*, but *post-*theoretical, and therefore likely to be deeply influenced by their training and their own philosophical preferences, making their intuitive judgments biased. Machery raises a similar problem against Williamson's claim that philosophical training makes philosophers' intuitions more reliable. "Some evidence suggests that at times expertise even makes experts worse than laypeople because their theoretical commitments bias their judgments" (Machery 2015, p. 196). MacFarlane raises this problem with respect to the logical intuitions of logicians and philosophers of logic themselves:

The dominant methodology for addressing them [questions about validity] involves frequent appeals to our "intuitions" about logical validity. I do not think it should surprise us that this methodology leads different investigators in different directions. For our intuitions about logical validity, such as they are, are largely the products of our logical educations (MacFarlane 2004, p. 2).

MacFarlane (2004, p. 2) refers to these intuitions as "indoctrination biases." It is not difficult to see that the intuitions of a dialetheist about true contradictions will diverge starkly from the intuitions of a classicist, for example. Affected by indoctrination biases, the intuitions of logicians and philosophers of logic should not guide theory choice in logic, on pain of circularity (at least if we assume abductionism or predictivism as the method for theory choice). What is left, then, are the intuitions of mathematicians, philosophers, and scientists who were not trained in logic. But are their intuitions reliable?

5 The pre-theoretical logical intuitions of specialists

So far I have been speaking of intuitions, but I did not address the question of what intuitions are. In philosophy, it is not unusual to see the word 'intuition' associated with some kind of supposedly innate or a priori knowledge, whose origins are obscure. As De Cruz (2015, p. 236) observes, both in psychology and philosophy "intuitions are regarded as assessments that come about without explicit reasoning and that seem

to have some *prima facie* credibility to those who hold them.” This aura of credibility is what makes intuitions philosophically relevant, and the absence of explicit reasoning—intuitions apparently just pop up in the mind—is what makes their origins obscure. Psychological investigation has shed light on this latter aspect. Following McCauley (2011), De Cruz distinguishes between *maturational* and *practiced* intuitions. Maturational intuitions arise early in development, typically in infancy or early childhood, from practices and contents that are mastered without formal instruction, “emerging through mundane interactions between a child and her social and physical environment” (De Cruz 2015, p. 239). Some examples are the linguistic intuitions that allow one to create sentences she has never heard before, or the intuitions that enable one to interpret facial expressions. Practiced intuitions, by contrast, emerge later and only after explicit training or instruction. “The most obvious illustrations are the sorts of good judgments that experts in any field can make in a snap, whether it is an engineer knowing what building material to use in a structure, a chess master knowing what move to make in order to avoid his or her opponent’s trap, or a long-term commuter knowing how the fares work on his or her local transit system” (McCauley 2011, p. 5). Practiced intuitions are a manifestation of expertise; they originate from extensive experience in a specialized domain.

In line with Williamson (2011), De Cruz (2015) suggests that the intuitions philosophers rely on to assess thought experiments are of the practiced kind. By the same token, the post-theoretical intuitions of logicians and philosophers of logic can be viewed as practiced, whereas laypeople’s intuitions about “what follows from what” are likely to be maturational. Just as an engineer is able to intuit, even before “doing the math,” that a certain material is not adequate for a certain kind of structure, a trained logician may be able to intuit that a certain argument is valid (or invalid) even before formalizing and probing it. Since these logical intuitions are not innate nor a priori, but acquired by training and hence post-theoretical, they are likely to be biased by the logician’s training and preferences, as argued above.

Philosophers, scientists, and perhaps some mathematicians who were not trained in logic will not have practiced post-theoretical logical intuitions, but perhaps their experience with argumentation within their own disciplines may somehow “sharpen” their reasoning skills in a way that can produce intuitions about what logicians call deductive validity. This may be so if argumentation in their disciplines instantiates key aspects of what Dutilh Novaes (2021) calls “Prover-Skeptic dialogues.” According to Dutilh Novaes, dialogical practices of this kind gave rise, in historical terms, to the creation of deductive logical theories. Thus, it may be that practitioners of this kind of dialogue could somehow acquire intuitions about deductive reasoning without needing explicit instruction in a particular logical theory.

Prover-Skeptic dialogues are a specialized kind of argumentative dialogue that takes place when interlocutors are jointly seeking for an indefeasible chain of inferences in order to demonstrate whether a certain conclusion follows necessarily from certain premises. Dutilh Novaes’s model of these dialogues involves two characters, Prover and Skeptic. Prover wants to establish a given conclusion Q from some given set of premises Γ . Skeptic, however, is not convinced at the beginning of the dialogue that Γ entails Q . Prover’s objective is to convince Skeptic of this entailment. To this end, Prover starts the dialogue by asking Skeptic to endorse the premises in Γ . If

Skeptic endorses them and does not present a global counterexample to the proposed entailment, Prover proceeds by putting forward a sequence of further statements that she claims follow necessarily from the premises Skeptic has endorsed. These are the intermediate inference steps that, Prover hopes, will eventually lead to Q . At each of these intermediate steps, Skeptic can ask for further clarification (if she thinks the inference is not sufficiently perspicuous), propose a counterexample to the inference, or simply endorse the inferential move. In response to Skeptic's objections, Prover may provide further clarification or modify the inference so as to avoid the counterexample. If both Prover and Skeptic are fully successful in playing their roles, at the end of the dialogue they will have produced a chain of valid inferences showing that Γ entails Q . The contribution of Skeptic in this is fundamental, since it is her disposition to contrive counterexamples and thus identify invalid inference steps that help Prover produce a chain of inferences wherein each link is immune to counterexamples.

Dutilh Novaes's Prover–Skeptic model of deductive dialogues, outlined above, presupposes a strictly controlled interplay of giving and asking for reasons. Dialogues fitting this model are not to be found everywhere, in daily life conversations. Even so, Prover–Skeptic dialogues are still less specialized than formal proofs in a logical theory, since participants to these dialogues are not supposed to use formal tools or explicit logical rules to defend their claims. The conversation about “whats follows from what” in these dialogues may be totally informal and, most important for my purposes here, this conversation does not need to be informed by any logical theory. In this sense, there may be instances of Prover–Skeptic dialogues that are completely “pre-theoretical” with respect to logical theories, i.e., dialogues in which participants do not know logical theories at all and therefore cannot ground their assessments of validity on any logical theory. For example, think of a theoretical physicist trying to show that a certain prediction follows from a given physical theory. Surely, she will use some mathematics, but her inferential steps are unlikely to be explicitly informed by any logical theory. Insofar as other physicists (acting as “Skeptics”) scrutinize her inferences, ask for further clarifications, and perhaps propose some counterexamples, there will be an instance of a Prover–Skeptic dialogue going on where no logical theory plays a role. Another example may be a theologian trying to prove the existence of God from some premises, and having her argument scrutinized by agnostic or atheist Skeptics. I submit that occasions like these are likely to give rise to pre-theoretical deductive logical intuitions among practitioners of these dialogues.

It is worth considering the psychological process through which Prover–Skeptic dialogues can give rise to deductive logical intuitions. According to the results from psychology of reasoning we saw above, before a claim that Q follows from Γ , people are likely to evaluate it under the influence of belief bias. Thus, since Prover is the one claiming that Γ entails Q , under the influence of belief bias—or, more precisely, myside bias—she is *less* likely to find counterexamples to the entailment she herself is proposing (and to intermediate inference steps). By contrast, Skeptic is initially unconvinced that Γ entails Q , and therefore the influence of myside bias predisposes her to contrive counterexamples to ‘ Γ entails Q ’ (and to intermediate inference steps) more easily, if there are any. At the same time that myside bias makes Prover oblivious to problems in her argumentation, it opens Skeptic's eyes to failures in Prover's argumentation. Now it is easy to see that experience with this kind of dialogue is likely to

produce intuitions about which inference steps are indefeasible, i.e., immune to counterexamples (at least in the informal settings where these dialogues take place). Over time, the practitioner of these dialogues *learns* that some inference steps are indefeasible because, in her experience, Skeptics never manage to contrive a counterexample to them. These intuitions will not be maturational but practiced, since their acquisition involves mastering a specialized dialogical argumentative practice.

In this way, scientists or philosophers who do not have formal training in logic can acquire practiced intuitive knowledge about which inferential steps are “good.” Their intuitions are likely to be more neutral than logicians’ and philosophers’ of logic, since they do not come from training in specific logical theories. But can reasoners experienced in Prover–Skeptic dialogues be the reliable judges of the validity of vernacular arguments that (Martin and Hjortland 2020, p. 17, quoted above) are looking for?

This does not seem to be the case. Insofar as logical intuitions do not come from a Platonic heaven, but rather originate from experience with Prover–Skeptic dialogues as I have been arguing, they cannot be more reliable than the dialogical practices that give rise to them. A comparison with other kinds of practiced intuitions illustrates this point. Consider the examples mentioned above from McCauley (2011), of an experienced commuter who is able to intuit the fare she will pay when going to every destination in her city, or an experienced engineer who can intuit the best material for a certain kind of structure. In both cases, if they really need absolutely certain information—because, say, the commuter wants to know the exact fare down to the last cent and the engineer is designing an aircraft—they would do better by *calculating*, rather than intuiting. In the commuter’s case, she would have to consult the price list and the regulations of the transportation company and calculate the fare (or consult the company’s app); in the engineer’s case, she would have to calculate the resistance of the material, following the standard practices and theories of her field. After all, experience with these operations, information, and theories was what gave them their intuitions; but their intuitions are nothing more than educated guesses. When they need more than educated guesses, they need to effectively calculate. The same goes for the post-theoretical logical intuitions of logicians. A logician can intuit at a glance that a certain argument is valid; but in order to make sure that it is really valid, she has to formalize it and use a logical theory to prove its validity.

Experienced practitioners of Prover–Skeptic dialogues are in a similar position. In these dialogues, the main “move” concerning the notion of validity is the presentation of counterexamples to an inference, showing that it is defeasible. The fact that a certain kind of inference was never challenged by a counterexample in one’s experience is what gives experienced practitioners of these dialogues the intuition that that kind of inference is valid. But, naturally, the fact that no one has ever come up with a counterexample does not imply that a counterexample does not exist. Thus, despite intuitions to the contrary, someone who really wants to probe an argument would do better by exposing it to the challenges of Skeptics. Here, again, intuitions are no more than educated guesses and therefore unreliable.

Even the scrutiny by Skeptics, however, will not conclusively show that the argument is valid, since, again, the fact that no Skeptic has come up with a counterexample so far does not imply that a counterexample does not exist. Recall that the dialogues

that give rise to truly *pre*-theoretical logical intuitions proceed without the aid of formal tools, logical theories or proof techniques. Just as logicians have to make use of a logical theory to prove conclusively that a certain argument is valid (in a given logical system), full certainty about the validity of an argument accepted by all Sceptics would demand its formalization and proof in a logical system. After all, this is the point of having a logical theory. Logicians create logical theories to, among other reasons, *prove* that certain kinds of arguments are valid under certain conditions. In the vernacular, we do not have either clarity about what makes an argument valid or fixed and precise definitions of implication, negation, disjunction, etc., on which we could rely to prove that an argument is valid.⁴ A logical theory defines these concepts precisely, and then proofs of validity are made possible. Of course, such proofs do not show that the given argument is *absolutely* valid; demonstrations of validity are always relative to the logical theory where they are made, and depend on the definitions of concepts such as implication, negation, and validity, whether truth-value gaps or truth-value gluts are accepted, and so on. This brings us back to the debate about theory choice in logic. What is clear (or so I hope) is that reliance on pre-theoretical intuitions will not help here, since they are less reliable than the theories in competition (in the sense that intuitions do not allow for proofs of validity).

In the next section I argue that logical theories do not simply capture pre-theoretical intuitions about validity; rather, they improve on these intuitions so as to prove that certain kinds of inference are valid under certain conditions and explicate why they are so.

6 Carnapian explication and ameliorative logical theorizing

The discussion of the previous sections casts doubt on abductivism and predictivism as methods for theory choice in logic. Both accounts presuppose the existence of some data in the form of pre-theoretical logical intuitions against which logical theories could be compared. I have been arguing that there is no pre-theoretical logical intuitions that could serve this purpose. Laypeople’s intuitions about “what follows from what” are non-monotonic. Although they could be seen as favoring non-monotonic logics, it does not seem adequate to rule out deductive logics on the basis that they do not correspond to laypeople’s intuitions. Logicians’ and philosophers’ of logic intuitions are a product of indoctrination and, as such, are biased. Really pre-theoretical logical-deductive intuitions may be found in specialists such as philosophers, scientists, and mathematicians (without extensive training in logic) who are experienced in some sort of Prover-Skeptic dialogues. But their intuitive judgments of validity are unlikely to be more reliable than proofs of validity in a logical system. All things considered, it

⁴ That is why the intuitions of practitioners of Prover-Skeptic dialogues about what follows from what should be deemed unreliable even if one assumes that there is an underlying notion of “validity” in such dialogues. Without clarity about what the logical terms are, what they mean precisely, and lacking proof techniques, even an experienced practitioner may fail in realizing that a given argument should count as invalid (valid) according to the criteria for validity implicit in the practice. Indeed, the fallibility of pre-theoretical logical intuitions is admitted even by abductivists such as Priest, for whom “the data is soft, and can be overturned by a strong theory” (Priest 2016, p. 42).

seems that the invocation of intuitions as a reason to rule out a logical theory seems unwarranted.

To be sure, if anti-exceptionalism about logic is right, logic really could not be different from other sciences in this respect. The history of science is full of episodes where intuitive judgments have been shown to be mistaken. Some examples: the intuition that the Earth is flat; the intuition that continents are static; the intuition that light should behave either as a wave or as particles. Scientific theories are not responsible to pre-theoretical intuitions or common sense; rather, they usually come up with findings and theories that confront our intuitions radically. The same goes for mathematics. Hahn (1980) presents a number of intuitive notions that were proven wrong in geometry. After showing how several intuitive ideas about the behavior of curves turned out to be mistaken, Hahn concludes:

Because intuition turned out to be deceptive in so many instances, and because propositions that had been accounted true by intuition were repeatedly proved false by logic, mathematicians became more and more sceptical of the validity of intuition. They learned that it is unsafe to accept any mathematical proposition, much less to base any mathematical discipline on intuitive convictions (Hahn 1980, p. 93).

Why logic should be different in this regard?⁵ Both the abductivist and predictivist, insofar as they think of intuitions as the data logical theories have to account for, are *exceptionalist* in this sense. But if intuitions are not the data, what is the data? Is there any data that logicians should account for when answering the question *what is validity?*

Haslanger (2012, p. 367) identifies “three common ways to answer ‘What is X?’ questions: *conceptual*, *descriptive*, and *ameliorative*.” Only the first two aim primarily at giving an account of some “data.” A conceptual account of X aims at revealing *our* concept of X “and looks to a priori methods such as introspection for an answer. Taking into account intuitions about cases and principles, one hopes eventually to reach a reflective equilibrium” (Haslanger 2012, p. 367). Intuitions, in this case, constitute the relevant data. A descriptive account of X aims at giving an accurate account of the phenomenon that the concept X is supposed to refer to. The relevant data, in this case, depend on the nature of the phenomenon in question; e.g., if the phenomenon is physical, then empirical data will be relevant. Abductivists and predictivists about logic seem to approach the question ‘what is validity?’ either conceptually or descriptively. If logical intuitions themselves are the phenomenon they intend to account for, their approach is conceptual; if intuitions are seen as conveying information about an underlying phenomenon (be it validity abstractly conceived or the most general aspects of the world, as for Williamson), then their approach is descriptive. Either way, they think of logical theories in *representational* terms.

⁵ Before turning to logical predictivism, Hjortland (2019) criticized reliance on intuitions as data for logical theories on the same grounds. He mentions as examples the scientific concepts of *mass* and *star*, which are very distant from their pre-scientific counterparts, and concludes: “relying on semantic intuitions about the defunct [pre-scientific] concept will lead one astray. Logical theories are no different” (Hjortland 2019, p. 261).

The ameliorative approach to what-is-X questions is not representational. Rather, it starts by asking “[w]hat is the point of having the concept in question ... What concept (if any) would do the work best?” (Haslanger 2012, p. 367). In this approach, the objectives of conceptual analysis are, first, to identify what purposes the concept in question is supposed to serve and, second, to improve on the available concept, or to replace it by a new one, so that it can serve that purposes better. One example is what Haslanger herself does with the concepts of race and gender. Another example, I submit, is what logicians do with the concept of validity.

In the previous section we have seen that pre-theoretical logical intuitions are engendered by experience acquired in some sort of Prover-Skeptic dialogues. These intuitions are best conceived of as educated guesses. Although we may intuit that a certain argument is valid, we cannot be sure that it is really so because the fact that no one has found a counterexample to it may be our fault rather than a demonstration that there are no counterexamples. The development of a logical theory is a way of going deeper in the investigation of validity; the use of formal techniques will ultimately reveal a counterexample or allow us to prove that, under certain assumptions, there is no counterexample. This is the point of having a rigorous account of the concept of validity such as those provided by logical theories.

In this sense, logical theorizing is primarily ameliorative. Logicians do not simply want to faithfully capture pre-theoretical intuitions about validity, but rather to provide a better understanding of what validity is—one which is superior to pre-theoretical notions because it can prove that certain principles and rules of inference are valid under certain assumptions, and also because it reveals what needs to be assumed (which logical constants, how they are to be defined, etc.) if certain principles and rules are to be valid.

The idea that logical theorizing is ameliorative is not new. It can be found already in Carnap (1950), in his conception of logical and scientific theorizing as *explicative*.⁶ On Carnap’s view, the explication of a pre-scientific concept does not aim at faithfully capturing intuitive notions about it; rather, “[t]he task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second” (Carnap 1950, p. 3). Carnap calls the vernacular notion in need of explication the *explicandum*, and the more exact theoretical concept that replaces it the *explicatum*. In logical theorizing, the explicandum is the pre-theoretical intuitive notion of “what follows from what,” and the explicatum is a logical theory.

In the Carnapian approach, a mismatch between the pre-theoretical intuitive notion and the theory is exactly what is expected. If there is ambiguity in the explicandum, then in order to make it exact some of the meanings associated to it have to go. Therefore, “an explicatum that aspires to be exact will necessarily misrepresent the inexact explicandum” (Dutilh Novaes and Reck 2017, p. 202). It is not difficult to see that there is plenty of ambiguity in the intuitive notion of validity. (Smith 2011, p. 27) makes this point:

⁶ Haslanger notes the similarity between her ameliorative approach and Carnap’s conception of explication in a footnote (Haslanger 2012, p. 367, fn. 1); Dutilh Novaes (2020a) offers a systematic comparison of Haslanger’s and Carnap’s approaches.

If you think that there is [an exact pre-theoretical ‘intuitive’ notion of valid consequence], 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 premisses and still have an inference which is intuitively valid? What about the inference ‘The cup contains some water; so it contains some H_2O molecules’? That necessarily preserves truth (on Kripkean assumptions): but is it valid in the intuitive sense?—if not, just why not?

If the pre-theoretical notion of “what follows from what” allows for so many different precisifications, logical theories should not be judged according to how much of pre-theoretical intuition they account for, as if similarity to intuitions constituted evidence for a logical theory. According to Carnap, similarity is required only to the extent that it justifies the claim that the explicatum is an explication of the explicandum. In Dutilh Novaes and Reck’s (2017, p. 203) view, “the issue of similarity in explication (and in formalization more generally) is partly an issue of intentionality, an issue of aboutness.” That is, the final product of logical theorizing may be very far away from the intuitive notions which motivated it, provided that it is still about “what follows from what” in a suitable sense.

Smith’s (2011) general point is that a rigorous theoretical notion can be said to capture faithfully an informal vague notion only if some elucidation, sufficiently precise even if yet informal, of the informal notion has already been provided. This initial refinement of the intuitive notion corresponds to Carnap’s requirement that an explication should start with a preliminary narrowing of the meaning of the explicandum “in order to prevent the discussion of the problem from becoming entirely futile” (Carnap 1950, p. 4). Without this preliminary clarification, theorists risk talking past each other, since the polysemy of the explicandum may allow for quite different approaches.

An example of such preliminary narrowing is Tarski’s condition of material adequacy for logical consequence. (Tarski 2002, p. 176) notices that “the concept of following [in everyday language] is not distinguished from other concepts of everyday language by a clearer content or more precisely delimited denotation, the way it is used is unstable,” and therefore

the task of capturing and reconciling all the murky, sometimes contradictory intuitions connected with that concept has to be acknowledged a priori as unrealizable, and one has to reconcile oneself in advance to the fact that every precise definition of the concept under consideration will to a greater or lesser degree bear the mark of arbitrariness (Tarski 2002, p. 176).

That said, Tarski assumes as “the point of departure” for his analysis of the concept of logical consequence “certain considerations of an intuitive nature” (Tarski 2002, p. 183), namely, his conditions of material adequacy for accounts of logical consequence: necessary truth-preservation and validity-preserving schematic substitution. In doing so, he selects two intuitions about validity among the possibilities contained in the broader intuitive notion of validity. But a different preliminary clarification could select different intuitions as a point of departure. An example is the criterion of containment, according to which in a valid inference, the conclusion must be causally

or epistemically contained in the premises. Precifications of the notion of validity that include the criterion of containment were common in medieval discussions on consequence and are related to the development of relevance logics in contemporary times (Dutilh Novaes 2020b).

The claim that the pre-theoretical notion of validity can be specified in different manners has been used to defend pluralism about logic (e.g., Beall and Restall 2006; da Costa and Arenhart 2018). However, this is not my point here. My point is just that, since the pre-theoretical notion of validity allows for different precisifications, any argument to the effect that one, some, or all of these precisifications are correct with respect to the pre-theoretical notion is unconvincing. The pre-theoretical notion does not provide *evidence* for logical theories. It is only an inspiration, the starting point, but not a criterion for correction or theory choice. As Carnap remarks, given that

in a problem of explication the datum, viz., the explicandum, is not given in exact terms ... it follows that, if a solution for a problem of explication is proposed, we cannot decide in an exact way whether it [the explicatum] is right or wrong [about the explicandum]. Strictly speaking, the question whether the solution is right or wrong makes no good sense because there is no clear-cut answer. The question should rather be whether the proposed solution is satisfactory, whether it is more satisfactory than another one, and the like (Carnap 1950, p. 4–5).

Satisfaction has to do with one's goals. A certain theoretical approach may be satisfactory with regard to some goals and unsatisfactory with regard to others. According to Carnap, a goal shared by any scientific explication is fruitfulness, which he understands as the potential of the explicatum to allow for the establishment of universal laws and connections between the intended phenomenon and other phenomena, a potential that the explicandum may not show to the same degree.

In other words, Carnap's view seems to be that an explication is useful or fruitful when it delivers 'results' that could not be delivered otherwise (or with much more difficulty), i.e. with the explicandum alone. What this suggests is a conception of explication as a method for *discovery* ... The goal is to produce *new knowledge* about the phenomena to which the explicandum pertains (Dutilh Novaes and Reck 2017, p. 206).

This, again, is a reason for accepting a larger degree of mismatch between explicandum and explicatum. The Carnapian view of logical theorizing is diametrically opposed to abductivism and predictivism in this regard; whereas the latter seek conformity between vernacular and theoretical concepts, the former wants fruitful theoretical concepts even at the cost of a mismatch with the vernacular. The explicatum cannot give rise to new knowledge if it remains faithful to the explicandum. "In this way, explication reveals itself as a cognitive tool leading to discoveries and new insights" (Dutilh Novaes and Reck 2017, p. 206). This happens regularly in logical theorizing, when logical theories reveal links between definitions, principles, and theorems that were not previously known nor knowable by means of pre-theoretical notions only.

Exactness and fruitfulness are not the only goals of explication. According to Brun (2016), in Carnap's later presentation of his conception of explication in Carnap (1963), Carnap "takes a decidedly more pragmatic perspective on explication." For the later

Carnap, “choosing an adequate explicatum is a practical decision which has to be taken in view of the specific problems the explicatum is expected to solve and in view of the role it is expected to play in the target theory” (Brun 2016, p. 1225). Here we can draw a closer connection between Carnap’s explication and Haslanger’s ameliorative approach. For Haslanger, as for the later Carnap, an explication is judged more satisfactory than others according to the ends we want the concept to serve. There is no absolutely correct answer when it comes to choosing among different explications; only when a goal is provided, can we select an explication as the most satisfactory with respect to that goal. In the next section, I argue that the satisfaction of one’s investigative goals is a more realistic criterion for theory choice in logic.

7 Goals and theory choice

For abductivists and predictivists, logicians concerned with *philosophical logic* (in the sense defined by Williamson 2017) want to capture the pre-theoretical notion of validity. Philosophical logical theorizing, however, is conducted with a variety of philosophical aims in mind. For example, one may want to identify which rules of inference secure necessary truth preservation when the notion of implication is relevant; or when the modal notions of necessity and possibility are taken into account; or when temporal aspects are considered; or when premises are contradictory; even if some of these aspects are not salient features of an intuitive notion of validity. There are also extra-logical aims. For example, moved by ontological concerns, one may be interested in laying down logical principles and definitions which secure necessary truth preservation when empty names are allowed in, or when only constructible objects are allowed in.

It is uncontroversial that logicians engage in logical investigation with these goals in mind and develop logical systems to meet them. Even so, the discussion about theory choice in logic is usually seen as being about which of these various logical systems are *correct* or *true*. But correct or true about what? Is there any fact of the matter as to, say, whether singular terms must denote? The discussion above should have shown that intuitions about these matters should not be relied on. Free logics can be seen as laying down the conditions under which inferences involving empty names retain the property of necessary truth preservation, rather than describing real facts or intuitions about singular terms. The same goes for other logics; modal logics investigate under which conditions inferences involving modalities retain the property of necessary truth preservation, and paraconsistent logics investigate under which conditions inferences involving contradictions do not trivialize. Putting aside Platonist accounts of logical truth (which anyway seem to be incompatible with anti-exceptionalism), and insofar as facts about vernacular argumentation do not constitute evidence for logical theories, as I have been arguing, the decisive factor for theory choice in logic seems to be one’s investigative goals. The logician selects or develops the theory that best fits her investigative aims.

I submit that this is what logicians and philosophers of logic really do in their actual practice, though sometimes under the guise of “appealing to intuitions.” The following passage by Resnik (2004, p. 181) provides an illustration of this point:

For a case where intuitions play a major role, take the common view among logicians that no formalism should count ‘There are at least two individuals’ as a logical truth. Some logicians base this upon the normative intuition that our inferential practice should not in itself decide questions of existence. While others appeal to the metaphysical intuition that there could be a universe containing fewer than two individuals, and some may appeal to both intuitions.

Is this really a matter of intuitions? That *there are* at least two individuals is intuitively true for most people, as far as I am concerned. What is counter-intuitive is the possibility of not existing at least two individuals; only philosophers entertain this possibility seriously, after years of philosophical training. At any rate, it is reasonable not to count ‘There are at least two individuals’ as a logical truth, but the reason is primarily methodological. Logicians do not want to address such a specific ontological question—how many objects there exist?—*qua* logicians, and therefore it is better not to make such a proposition a consequence of logical theories.

One aspect of Shapiro’s (2006) and Cook’s (2010) account of logical theories is in line with my claim that the logic one chooses depends on one’s investigative aims. As (Cook 2010, p. 500) puts it, “[d]ifferent logics, viewed as models of various linguistic phenomenon, are correct relative to different theoretical goals.” In contrast with the Carnapian view I am defending here, though, Shapiro and Cook see logical theories as *modeling* a previously given phenomenon. In Shapiro’s (2006, p. 49) words, “a formal language is a mathematical model of a natural language.” Here, though, logical theories do not model a previously given phenomenon, but improve pre-theoretical notions, and therefore necessarily depart from them. When a logician is investigating vagueness, for example, she is not (or should be not) concerned with providing a faithful account of how vagueness is treated by speakers in everyday situations, but rather with providing a reliable method of making inferences in contexts where vagueness is present.

If this is so, the relevant data for theory choice are not facts about pre-theoretical notions, but facts about the very logical theories in dispute. For example, if one wants to avoid commitment with the existence of non-constructible objects in mathematics, a logic that allows proofs by *reductio ad absurdum* does not meet her purpose. The relevant data here, then, is whether or not a logical theory admits such proofs. More generally, the relevant data for theory choice includes theorems and meta-theorems that show how the system in question works, which principles and rules of inference fail or hold in the system, how the system can be integrated with other theories, such as mathematical theories and theories of truth, and the like. These are the data that inform the choices of logicians, mathematicians and philosophers, according to their investigative aims. In addition, data about the intended context of application and about what features of a logical theory best serve the intended context are also relevant.

Notice that, in the Carnapian view I am defending here, logics are investigative tools used by specialists with specific purposes in mind. Therefore, the selection of a logical theory among many to fulfill certain investigative goals never has (or should not have) the effect of imposing a way of reasoning as the correct one for all purposes, in all situations. The choice is always instrumental, to fulfill certain investigative purposes

in specific contexts.⁷ For example, the intuitionistic mathematician does not choose (or should not choose) an intuitionistic logic because she thinks that it is the correct logic for every situation, but just because she wants to make sure that, when she is developing intuitionistic mathematical theories, she will not unnoticeably make claims that could commit her with non-constructible objects.

This instrumental use of logics seems to lead to a pluralist stance. But this is not necessarily so. The debate on pluralism *versus* monism is usually understood in representational terms. As a rule, pluralism is conceived of as either the view that there are multiple genuine *representations* of one and the same phenomenon (e.g., substantial logical pluralism as defined in Cook 2010) or, conversely, the view that there are multiple phenomena to be *represented* by logical theories (e.g., Beall and Restall's 2006 logical pluralism). The monist view opposite to these two kinds of pluralism is that there is only one logic that correctly codifies the unique relation of logical consequence in natural language. Either way, both the monist and the pluralist agree that a logical theory aims at representing some previously existing phenomenon or phenomena.⁸

My suggestion here is that logical theories are not representational; they lay down definitions, axioms, and rules of inference with the purpose of securing necessary truth preservation under certain conditions. If these conditions are restrictive—for example, if one wants to restrict the intended domain to constructible objects only—then various logical theories will be needed, each one satisfactory with regard to the specific goals that imposed such restrictions. However, the monist may aim at providing a logical theory that accounts for valid inferences under no restrictions, that is, inferences that are universally valid no matter what. In this case, the goal of the investigation would be exactly the identification of these universally valid forms of inference. Insofar as Priest (2006) claims that deductive validity is necessary truth preservation in all situations, he can be seen as one pursuing such a universal logical theory. Williamson (2017) is another case in point, given his claim that logical laws are unrestricted generalizations, true of absolutely everything. Priest and Williamson, both monists and both taking logical theories to be representational, think that the “one true logic” should include only universally true laws. In the current non-representational account, there is no matter of fact about whether there are universally *true* logical laws. Here the cogency of monism becomes a technical matter: is it possible to provide a logical theory that secures necessary truth preservation in all situations? If the monist can survive the challenge of logical nihilism (Cotnoir 2018) and provide such a logic, even if a very weak one, she can call this *the most* satisfactory logic under the assumption that validity is truth-preservation in all situations.

⁷ According to Steinberger's interpretation of Carnap's view on the normativity of logic, for Carnap “one's chosen logico-linguistic framework does set norms, constitutive norms even, but these norms are not constitutive of thought understood as the psychological activity of individual agents” (Steinberger 2017, p. 156). The norms have *local* normative force only.

⁸ There are exceptions, of course. The sort of pluralism motivated by Carnap's principle of tolerance is perhaps the most notorious one.

8 Conclusion

We have seen that pre-theoretical logical intuitions cannot provide the kind of data that could be relevant to theory choice in logic. Other than what some philosophers claim, competence in deductive reasoning is not a widespread feature of human rationality. Untrained human rationality is non-monotonic, but this should not be seen as a reason to reject deductive logics as wrongheaded. Deductive logical intuitions must be acquired by training, be it in a logical theory or in some sort of Prover–Skeptic dialogues. In the first case, logical intuitions are likely to be reliable, but they should not guide theory choice on pain of biased choices. In the second case, the intuitions acquired by means of experience in Prover–Skeptic dialogues are not biased by any specific logical theory, but, on the other hand, are not reliable. What makes an experienced practitioner of such dialogues have the intuition that a certain inferential step is valid is the absence of counterexamples to that inferential step in her experience. This does not imply, however, that that inferential step is really valid, since the absence of counterexamples in her experience may be due to the fact that the Sceptics she has met have failed to contrive one. These considerations show that both abductivism and predictivism are not viable accounts of theory choice in logic, at least insofar as they take adequacy to the data as a desideratum and take pre-theoretical logical intuitions to be the relevant data.

In order to achieve full certainty about the validity of a certain kind of inference (under certain conditions) one has to investigate it by using regular logical techniques, i.e., by developing a logical theory. This logical theory, however, will not simply capture the pre-theoretical intuitions that motivated it; rather, it will represent an improvement over those intuitions. Relying on the theory, now one can be sure that her pre-theoretical intuitions were right, or can correct them otherwise. In this sense, logical theorizing is ameliorative, rather than descriptive. A general goal of logical theorizing is to prove that certain kinds of inference are valid under certain conditions and to explicate why they are so.

In light of these observations, and putting aside issues concerning the metaphysics of logic, I submit that there is no pre-theoretical data logical theories should account for. Logical theories are not representational, but ameliorative. True enough, amelioration and representation are not mutually exclusive; the natural sciences are both ameliorative (they improve our thoughts and practices) and representational. But my point here is that logic is not like science in this regard: it is ameliorative (improves our reasoning techniques) without being committed to a description of our pre-logical ways of reasoning. This does not turn logic into an exception among other human epistemic undertakings. Logic can still be seen as similar to epistemic activities such as engineering or technological development, in that both logic and engineering aim at improving human practices in response to certain needs. The choice of a logical theory, just as the choice of a technical solution, is (or should be) guided by one's goals and informed by data about the very logical theories or technical solutions in dispute. Just as one chooses the technical solution that best suits her practical aims, one chooses the logical theory that best suits her scientific, mathematical, logical, or philosophical aims.

Acknowledgements I am grateful to Catarina Dutilh Novaes and four anonymous reviewers for valuable comments on earlier drafts of this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ball, L. J., & Thompson, V. A. (2018). Belief bias and reasoning. In L. J. Ball & V. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 16–36). Routledge.
- Beall, J., & Restall, G. (2006). *Logical pluralism*. Clarendon Press.
- Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, *81*, 1211–1241. <https://doi.org/10.1007/s10670-015-9791-5>
- Bueno, O., & Colyvan, M. (2004). Logical non-apriorism and the 'law' of non-contradiction. In G. Priest, J. Beall, & B. Armour-Garb (Eds.), *The law of non-contradiction* (pp. 156–175). Oxford University Press.
- Carnap, R. (1950). *Logical foundations of probability*. The University of Chicago Press.
- Carnap, R. (1963). Replies and systematic expositions. In P. A. Schilpp (Ed.), *The Philosophy of Rudolf Carnap* (pp. 859–1013). Open Court.
- Cook, R. T. (2010). Let a thousand flowers bloom: A tour of logical pluralism. *Philosophy Compass*, *5*(6), 492–504. <https://doi.org/10.1111/j.1747-9991.2010.00286.x>
- Cotnoir, A. J. (2018). Logical nihilism. In J. Wyatt, N. Pedersen, & N. Kellen (Eds.), *Pluralism in Truth and Logic* (pp. 301–330). Palgrave MacMillan.
- da Costa, N., & 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>
- De Cruz, H. (2015). Where philosophical intuitions come from. *Australasian Journal of Philosophy*, *93*(2), 233–249. <https://doi.org/10.1080/00048402.2014.967792>
- Dutilh Novaes, C. (2020a). Carnapian explication and ameliorative analysis: a systematic comparison. *Synthese*, *197*, 1011–1034. <https://doi.org/10.1007/s11229-018-1732-9>
- Dutilh Novaes, C. (2020b). Medieval theories of consequence. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (fall 2020 edn). Metaphysics Research Lab, Stanford University.
- Dutilh Novaes, C. (2021). *The Dialogical roots of deduction: historical, cognitive, and philosophical perspectives on reasoning*. Cambridge University Press.
- Dutilh Novaes, C., & Reck, E. (2017). Carnapian explication, formalisms as cognitive tools, and the paradox of adequate formalization. *Synthese*, *194*, 195–215. <https://doi.org/10.1007/s11229-015-0816-z>
- Elqayam, S. (2018). The new paradigm in psychology of reasoning. In L. J. Ball & V. Thompson (Eds.), *International Handbook of Thinking and Reasoning* (pp. 130–150). Routledge.
- Evans, J. S., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*(3), 295–306. <https://doi.org/10.3758/bf03196976>
- Hahn, H. (1980). The crisis in intuition. In B. McGuinness (Ed.), *Empiricism, logic and mathematics* (pp. 73–102). Springer.
- Haslanger, S. (2012). *Resisting Reality: social construction and social critique*. Oxford University Press
- Hjortland, O. T. (2017). Anti-exceptionalism about logic. *Philosophical Studies*, *174*(3), 631–658. <https://doi.org/10.1007/s11098-016-0701-8>
- Hjortland, O. T. (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>
- Johnson-Laird, P. N., Goodwin, G. P., & Khemlani, S. S. (2018). Mental models and reasoning. In L. J. Ball & V. Thompson (Eds.), *International Handbook of Thinking and Reasoning* (pp. 346–365). Routledge.
- MacFarlane, J. (2004). *In what sense (if any) is logic normative for thought?*. Unpublished.

- Machery, E. (2015). The illusion of expertise. In E. Fischer & J. Collins (Eds.), *Experimental Philosophy, Rationalism, and Naturalism: rethinking philosophical method* (pp. 188–203). Routledge.
- Markovits, H. (1985). Incorrect conditional reasoning among adults: Competence or performance? *British Journal of Psychology*, 76, 241–247. <https://doi.org/10.1111/j.2044-8295.1985.tb01948.x>
- Martin B, & Hjortland, O. T. (2020). Logical predictivism. *Journal of Philosophical Logic*, 1–34. <https://doi.org/10.1007/s10992-020-09566-5>.
- McCauley, R. N. (2011). *Why religion is natural and science is not*. Oxford University Press.
- Mercier, H. (2018). Belief bias and reasoning. In L. J. Ball & V. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 401–414). Routledge.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608–631.
- Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, 71, 305–330. <https://doi.org/10.1146/annurev-psych-010419-051132>
- Priest, G. (2006). *Doubt truth to be a liar*. Clarendon Press.
- Priest, G. (2016). Logical disputes and the a priori. *Principios*, 23(40), 29–57. <https://doi.org/10.1017/s0140525x10000853>
- Ragni, M., Kola, I., & Johnson-Liard, P. N. (2017). The Wason selection task: A meta-analysis. In *Proceedings of the 39th annual meeting of the cognitive science society*. London. <https://cogsci.mindmodeling.org/2017/papers/0192/index.html>.
- Resnik, M. D. (2004). Revising logic. In G. Priest, J. Beall, & B. Armour-Garb (Eds.), *The law of non-contradiction* (pp. 178–194). Oxford University Press.
- Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 91(3), 497–510. <https://doi.org/10.1037/0022-0663.91.3.497>
- Shapiro, S. (2006). *Vagueness in Context*. Clarendon Press.
- Smith, P. (2011). Squeezing arguments. *Analysis*, 71(1), 22–30. <https://doi.org/10.1093/analysis/anq120>
- Steinberger, F. (2017). Frege and Carnap on the normativity of logic. *Synthese*, 194, 143–162. <https://doi.org/10.1007/s11229-015-0880-4>
- Tarski, A. (2002). On the concept of following logically. *History and Philosophy of Logic*, 23(3), 155–196. <https://doi.org/10.1080/0144534021000036683>
- Williamson, T. (2007). *The philosophy of philosophy*. Blackwell Publishing.
- Williamson, T. (2011). Philosophical expertise and the burden of proof. *Metaphilosophy*, 42(3), 215–229. <https://doi.org/10.1111/j.1467-9973.2011.01685.x>
- Williamson, T. (2013). *Modal logic as metaphysics*. Oxford University Press.
- Williamson, T. (2017). Semantic paradoxes and abductive methodology. In B. Armour-Garb (Ed.), *Reflections on the liar* (pp. 325–346). Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.