Epilogue

I have aimed to rectify what I see as persistent misunderstandings about, as well as a general underappreciation of, a mode of reasoning that, evidence indicates, forms an important part of our cognitive tool box: abduction, which gives special weight to explanatory relations, as when, for instance, the appreciation that our evidence makes perfect sense in light of one hypothesis but not so much in light of its rivals raises our confidence in the former at the expense of our confidence in each of the latter.

That does not mean that there is literally a single rule of abduction that people can apply wherever and whenever they are able to make judgments of explanation quality. I have argued that critics are basically right when they contend that abduction is a slogan but wrong to deem this a shortcoming. Deductive logic offers a set of domain-general and topic-neutral rules, and perhaps this made many think that the rules guiding our nondeductive reasoning must be of that kind as well. On the other hand, chapter 6 looked at a psychologically oriented theory of rationality that makes it from the start reasonable to expect that the rules for nondeductive reasoning will *not* hold uniformly. Even more generally, the *Encyclopædia Britannica* (2006) circumscribes intelligence as the "ability to adapt effectively to the environment." While this is only one among many definitions of intelligence around, most would agree that the said kind of adaptability is at least a key characteristic of human intelligence. And this, too, should make it unsurprising if people are found to reason in ways that can be fine tuned depending on the environment so as to more readily pick up the inductive cues provided by that environment. I looked at rules of inference that appear to have that feature, and those rules were seen to do better than Bayes's rule in environments that were nothing out of the ordinary. The rules, while context specific, were tied together by the slogan that explanation guides inference.

If you had heard about abduction before you read this book, what you heard was probably mostly critical. Current thinking about how we ought to respond to new information, how we should revise our beliefs or degrees of belief in light of new evidence, is so strongly dominated by Bayesian confirmation theory that the idea that we are better served, even if perhaps only sometimes, by any mode of reasoning that could go against Bayesian tenets may be taboo to many. Bayesians have nothing against abduction as a kind of assistive tool, a helper function for determining priors perhaps, or (sometimes) likelihoods, but to the extent that abduction conflicts with Bayesian dogma, we should avoid it like the plague.

Bayesians have been quite explicit about the doom that awaits abductive reasoners: they make themselves vulnerable to exploitation, and they make their degrees of belief unnecessarily inaccurate. Worse yet, they do so willingly, or at least they could have figured out a priori that those things were going to happen! If that is not a token of irrationality, what is?

At first blush, the arguments for these claims can easily seem unassailable, supported even by *proofs*. On closer scrutiny, however, those proofs do not appear to connect so well to Bayesian allegations of irrationality on the part of, for instance, abductive reasoners. That is because, in arguing for the broader claims, Bayesians sneaked in some assumptions that the friends of abduction need not sign up to. We saw that if we think of update rules as packaged together with other epistemic principles and also with decision-theoretic principles, then there are ways to reason abductively free of charge, meaning that, for instance, you can assign bonuses for best explanations and still avoid Dutch books. That such additional principles are irrelevant and thus can be ignored, or that there must be something amiss with the ones that make for a non-Dutch-bookable combination with abduction, has been assumed by Bayesians without argument. Similarly, their inaccuracy-minimization argument assumes, again without any support, that the one and only sense of inaccuracy minimization that should matter to questions of epistemic rationality is that of expected next-step inaccuracy minimization. Moreover, Bayesians present the latter argument as being purely epistemic, assuming, again without argument, that scoring is completely free of pragmatic encroachment. As if that were not enough, we found that, for independently plausible ways of measuring inaccuracy, the claim is false to begin with.

Apart from all that, the dynamic Dutch book and inaccuracy-minimization arguments would still have been successful only had they shown that

using EXPL or any other non-Bayesian update rule incurs costs to no benefit. But the most that they have shown is the "incurs costs" part and not the "to no benefit" part. As said, I strongly doubt that the dynamic Dutch book argument shows much of interest at all. However, cutting the Bayesian some slack (e.g., by granting that the Brier score is the only correct way of measuring inaccuracy), at least the inaccuracy-minimization argument could be said to accomplish *something*, viz., shifting the burden of proof to those wanting to maintain the rationality of forms of non-Bayesian updating, making it incumbent on them to show that there *are*, or at least *can* be, benefits to such updating procedures.

That is what I have sought to do, using various probabilistic instances of abduction. The main point of chapters 6 and 7 was to show that in some contexts the use of such rules has clear advantages over the use of Bayes's rule. I very much doubt that what I have shown generalizes to each and every context. Indeed, we showed in section 7.4 that sometimes Bayes's rule did best in a context of social learning, where teams of doctors were evaluated on the basis of their ability to save their patients' lives. From my perspective, that is perfectly fine. My goal has emphatically *not* been to show that the widespread popularity that Bayesianism enjoys is completely undeserved; adopting Bayesian norms may be the right thing to do in certain contexts. As I put it in section 1.2.1, what is misguided is Bayesian *imperialism*, which deems any deviation of those norms a token of irrationality. In chapter 6, I embraced insights from Elqayam, Gigerenzer, and others, who have argued against a universalist conception of rationality and in favor of a more local and contextual notion that makes room for a plurality of normative criteria. Rationality, in this view, requires a match between, on one hand, a subject's epistemic or practical behavior and, on the other, her goals, her cognitive capacities, and the environment in which she finds herself.

A caveat about the probabilistic versions of abduction that have featured in the foregoing: no claim has been made to the effect that these rules are realistic in any way, that they are actually being employed by people, or that they are at least idealizations of rules that we actually employ. Some of the rules—in particular Good's and Popper's rules—were loosely inspired by the findings from Douven and Schupbach (2015b), which showed that Good's and Popper's measures of explanatory strength came close to capturing human judgments of explanatory strength. It does not follow from this, however, that people use Good's or Popper's rule to accommodate new evidence, which in

turn does not imply that these rules cannot at least sometimes serve as useful approximations of actual human reasoning. It would certainly be worth conducting further experimental work to investigate the extent to which these rules are able to accurately predict people's deviations from Bayesian reasoning. That work might suggest versions of explanation-based update rules that are more realistic than EXPL, Good's rule, or Popper's rule. If these to-be-discovered rules can be formalized—admittedly, a big *if*—we might be able to investigate their normative status much in the way in which we put EXPL and the other rules to the test in chapters 6 and 7. But, as intimated at the start (section 1.2.1), I am committing neither to the descriptive adequacy nor to the normative correctness of EXPL or any of the other explications of abduction that featured in the foregoing. In the present work, these abductive rules serve only dialectical purposes. In particular, they were introduced to help make two main points: (1) that abductive reasoning can in some contexts outperform Bayes's rule in ways that are of immediate importance to us; and (2) that the exact form of abductive reasoning on which we should rely may differ from one context to the next, so that there can be no single best formalization of abduction to which we can turn irrespective of the circumstances. The choice for EXPL was primarily motivated by the fact that it had been used in one of the main *attacks* on abduction, making it seemingly safe to choose the same explication in a *defense* of abduction. Good's and Popper's rule were, as said, introduced as variants of EXPL to make the point that different contexts may call for different explications of abduction; those rules were inspired by empirical findings about how real people update their probabilities, but without the ambition to capture real updating behavior.[1]

For those interested in pursuing follow-up research aimed at providing more realistic versions of abduction, I note here that there is no shortage of alternative explications of abductive reasoning that prima facie merit some atten-

1. Thus I take Climenhaga's (2017b) critique that EXPL and the other probabilistic versions of abduction that we considered may lead to trouble when simultaneously applied at different levels of explanation to be misguided, not only because, as said (and as pointed out in various previous publications), I consider none of those formulations a definite statement of a possible realistic instance of abduction, but also because, even more importantly, different levels of explanation may create different contexts, and there is no reason to hold that whichever explication of abduction is best suited for one level of explanation must also be best suited for any other level of explanation. (Climenhaga does say at various junctures that he believes that it is preferable to have only one version of abduction, without however providing an argument for this claim.)

tion. One could not only consider different measures of explanatory goodness but also related measures, notably measures of coherence (see section 1.2.1) and measures of the degree of competition among hypotheses (Schupbach & Glass, 2017; Glass, 2021a). Moreover, contributions to the debate about explanatory reasoning tend to be noncommittal regarding theories of explanation—as this book has also been—and in general there has been little contact between the debate about rational updating and the debate about explanation. It might be worth investigating whether the various proposals concerning explanation that have been made in the philosophy of science can be compared in the kind of settings developed in chapters 6 and 7. For instance, might a suitably formalized causal model of explanation (perhaps building on Pearl, 2000) do better in the sort of simulations reported in those chapters than an equally suitably formalized unificationist model (perhaps building on the aforementioned measures of coherence)? It would be an interesting finding if, for example, inferring to the best causal explanation turned out to be generally more successful, or in some contexts more successful, than inferring to the best unifying account.

Jeffrey (1992) famously argued that Bayesian epistemology should not be conceived as providing a set of rules that apply mechanically and universally but rather as providing tools whose proper handling requires experience and skill. Indeed, he thought of Bayesian reasoning as constituting an *art*: "the art of judgment." The view on abduction that I have advocated in this work is very much in the same spirit except that it presents the art as being even more encompassing: we also have to be able to pick the, or a, right updating rule for any given context. That may sometimes be Bayes's rule, but it can also be some version of abduction.

Scientific progress has to an important part been due to the fact that we have been able to devise better ways of obtaining data about the world. But, to repeat from chapter 1, it has also been due to the fact that we have learned more about what we can *infer* from the data that we have obtained: that we have learned more about how to learn. We have not just expanded our knowledge; we have improved how knowledge is being expanded. To my eyes, becoming clearer about our inferential practices, in particular about the confidence that these practices permit us to have in our beliefs, constitutes progress—*scientific* progress, even if perhaps not of the most glamorous kind. I have tried to contribute to this progress by, I hope, removing some misconceptions of

abductive reasoning and by showing that, in the right hands and when used judiciously, it can be a powerful tool.

My main concern has been to make a positive case for abduction as a mode of nondeductive reasoning in its own right, not subservient to this or that confirmation theory. Given the dominance of Bayesianism, however, this book may well be seen as a critique of that position. So, in closing, I want to be still a bit clearer about how, in my view, the two accounts relate to each other and what I would regard as helpful responses from Bayesian colleagues and what as unhelpful responses.

Although not concerned with abduction, Giere (1988) voiced serious misgivings about Bayesianism, and he did so largely on the same grounds that motivated the advocates of ecological rationality. Most centrally, he argued that people cannot possibly obey Bayesian norms given their limited capacity to gather, memorize, and process information. This consideration led Giere to reject Bayesianism altogether and to adopt in its stead a Simonian satisficing account of rationality.

It has always appeared to me, however, that there is a tension between Giere's unqualified rejection of Bayesianism and the version of the so-called semantic view of theories that he offers in the same book, a version that many still regard as the best on offer. According to the semantic view, scientific theories should not be thought of as axiomatic systems, connected to the world via certain bridge principles, but rather as sets of models. Giere convincingly argues that there is an important element missing from the standard presentation of the semantic view, to wit, the observation that theories also comprise hypotheses postulating similarity relations between the models that they define and certain worldly systems. To present a theory is to present a set of models *and* to point out which parts of those models reflect which parts of reality to what degree.

While I share many of Giere's qualms about Bayesianism—even if these only lead me to reject Bayesian imperialism—I still believe that precisely his version of the semantic view allows for a more favorable assessment of Bayesianism than it receives in Giere (1988). For in spite of all the criticisms, one could still argue that Bayesianism offers a set of models of reasoning—models for synchronic and diachronic belief management and for decision making—that not only allow us to represent with some degree of precision important aspects of how people do behave cognitively but also captures to some extent our

pretheoretical ideas about what it is to behave in cognitively appropriate ways (e.g., how to choose wisely).

Clearly, though, this is not to say that there is no room for improvement. If Bayesianism captures important aspects of our reasoning to some extent, but we are able to capture those aspects more accurately or more comprehensively by acknowledging the role that explanatory considerations sometimes play in reasoning, then that is progress. If it turns out that by attending to such considerations we sometimes get more of what we want from our reasoning than if we were pure Bayesian agents, then the same proposal is also closer to describing our pretheoretic normative thinking about rationality. The proposal made in this book, at least insofar as it pertains to graded beliefs, is far from a radical departure from Bayesianism. It emphasizes the importance of explanation to reasoning, but it could be rightly remarked that from a neutral perspective the update mechanisms that we looked at, which are sensitive to explanation quality, are but relatively minor deviations from Bayes's rule. At an abstract level, the endeavor undertaken in the foregoing is perhaps best thought of as an attempted de-idealization (McMullin, 1985) or concretization (Nowak, 1980) of an existing account of rationality, meaning that we try to pass from certain idealizations to less idealized assumptions and thereby to get closer to an account of rationality for real people. That is how I hope to have improved on Bayesianism.

Paul Moser (1990, p. 6) urges specifically with respect to decision theory that "adequate principles of rational decision making should be psychologically realistic relative to actual human decision making." The present work has been written in the conviction that the same holds, *mutatis mutandis*, for theories of rationality generally and that our efforts should be geared toward making such theories more responsive to human psychology, in particular to cognitive limitations. We can acknowledge as much without abandoning the project of conceiving a normative theory of rationality. Indeed, it would seem that *only* a theory of rationality that attends to human psychology can have normative force, inasmuch as the time-honored principle that "ought" implies "can" would be totally empty if "can" were not to refer to what is possible for actual human beings to accomplish.

To be sure, theorizing may become hopelessly difficult if we try to do away with idealizations altogether. But our goal can and should still be to see which of the idealizations that we are making can be replaced by more realistic assumptions while nonetheless retaining a theory that *both* accurately

describes people's belief management and decision making *and* reflects our main pretheoretical convictions regarding rationality. This goal is not fostered by taking on board *more* and *crasser* idealizations just to shield one's preferred theory from refutation.[2] In this light, my recommendation to Bayesians wanting to object to the criticisms that this book levels against the dynamic Dutch book and inaccuracy-minimization argument is to refrain from taking on board assumptions that would bring us further away from a theory of rationality for real people, for instance, assumptions that require cognitive capacities entirely unlike any we ordinary mortals possess.

For instance, in section 4.2, I compared Bayes's rule with an instance of EXPL by letting two players compete in a game in which they asserted what they thought was the true hypothesis as soon as that was subjectively highly likely to them. The inspiration for this came from a remark by Frank Jackson about high subjective probability warranting assertion. Bayesians might refuse to go along with that idea and propose instead that we should assert something only if doing so maximizes expected utility. Accordingly, they might say that my comparison was unfair, for a true Bayesian would never have agreed to play by the rules that I was assuming.[3] It seems to me, however, that the just-mentioned Bayesian account of assertion can at best pertain to artificial, computationally unconstrained agents and certainly not to us. The flow of our conversations would be seriously disrupted if we had to constantly perform the computations that the Bayesian account of assertion would demand from us.

Similarly, Bayesians could respond to my critique of their inaccuracy minimization argument, which as I said focuses on next-step inaccuracy, by trying to extend it to cover the accuracy or epistemic utility of *all* future updates. In doing so, they might have to suppose that we can envision all possible data streams that we might receive and all possible decisions that we might have to make after the receipt of each piece of data, that we can estimate how much inaccuracy or epistemic utility will matter to us at any particular

2. Exactly what idealizations we are willing to make may also depend on context. If accepting a couple of extra idealizing assumptions makes your theory much more powerful or elegant, then for some purposes that may be fine. At other times, you may be willing to sacrifice some computational power or elegance for the sake of making your theory more realistic or more predictively accurate.

3. Richard Pettigrew raised this objection in personal communication, for which I thank him.

moment in time, and perhaps more.[4] But why not? Fantasize away—there is a market for philosophical fiction!

Perhaps not all Bayesian colleagues have an interest in developing a theory of rationality for *real* people or see any value in trying to connect with the extant psychological literature on human reasoning. Cognitive limitations are hard to model formally and may be even harder to generalize, given that they can differ from one individual to another. Thus, attending to such limitations could come in the way of an activity that many philosophers like so much, to wit, formal model building. If that comes with the price tag of having to accept some highly unrealistic assumptions, then so be it.

But not all Bayesians are alike. In section 1.2.1, I referred to Lycan's typology of explanationist positions, going from a weak form according to which abduction is *a* (possibly derivative) mode of rational nondeductive reasoning to a very strong form according to which it is ultimately the only rational mode of reasoning. As mentioned there, my favored form is closest to the intermediate position that Lycan calls "Sturdy Explanationism," according to which abduction is a fundamental mode of rational reasoning, but not necessarily excluding the existence of other modes of rational reasoning. Whereas I am not aware of work trying to delineate types of Bayesianism along these same lines, a typology much like the one that Lycan proposed for explanationism would make sense and might be fruitful for Bayesianism as well. In particular, there would seem a clear place for what could be called "Sturdy Bayesianism," according to which Bayesian reasoning is both fundamental and rational but not necessarily the only such form of reasoning. If I am not mistaken, there are already Bayesian philosophers of science who appear to conceive of Bayesianism as providing valuable tools for modeling important parts of scientific reasoning and human reasoning more generally, rather than as a universal theory of rationality (e.g., Sprenger & Hartmann, 2019; Vassend, 2019, in press; Sprenger, 2020).[5] I cannot speak for these philosophers and I

4. For instance, they might suppose this to undercut the first objection to Briggs and Pettigrew's (2020) argument that was raised in chapter 4, footnote 17.

5. Some statisticians in the Bayesian camp definitely seem to qualify as sturdy Bayesians. For instance, Gelman and Shalizi (2012, 2013) have argued that while we are often well advised to analyze our data using Bayesian statistical techniques, in bringing the results of such analyses to bear on our theories we are not committed to Bayesian practices. Indeed, Gelman and Shalizi (2012, 2013) advocate a hypothetico-deductive approach at the level of theory choice. However, as Borsboom and Haig (2013) argue in their commentary on Gelman and Shalizi (2013), this is a place where one might more fruitfully want to invoke abduction. Furthermore,

do not know whether they would want to be identified as sturdy Bayesians, but my impression—and hope—is that they would be prepared to seriously consider abduction as a possible further valuable tool for studying rational reasoning. In section 1.2.2, we saw that Lipton, Psillos, and others argued that explanationism is compatible with Bayesianism. In doing so, however, they were happy to go along with Bayesianism's imperialist aspirations and sought to show how explanationism might be able to effectively serve those. By contrast, Sturdy Bayesianism is evidently compatible with the brand of explanationism defended in this book in a way that does not make one subservient to the other.

This book will have accomplished its mission not if all readers agree with each and every argument I have given—a most unlikely event to happen—but if philosophers and psychologists of reasoning find some motivation in it to start considering abduction as being at least worthy of more attention than it has been given in the last twenty-five years or so. Abduction could certainly benefit from such increased attention. We looked at formal models of abduction, both in an individual and in a group setting, but—to repeat—those were only toy models, serving strictly dialectical purposes. We identified as an important avenue for future research the task of devising models that connect more closely to how actual people, or groups of actual people, update on newly acquired information. To the best of my knowledge, all the relevant data available at the moment was reviewed in chapter 3—and clearly, that was not a whole lot of data. My hope is that we will see further experiments of the kind reported in that chapter and that the outcomes of these experiments will help to refine our models of abduction, which in turn will suggest further experimental work, and so on.

---

there is nothing that would prevent sturdy Bayesians from taking on board Lipton's idea that abductive reasoning may help us fix our priors or Psillos's idea that Peircean abduction may be invoked to narrow down the number of theories that are candidates for further exploration (for both ideas, see section 1.2.2). Thus, Bayesian statistics may be supplemented by abduction in a variety of ways.

Appendices