Richard Bradley. *Decision Theory with a Human Face*. Cambridge University Press 2017. 350 pp. $105.00 USD (Hardcover ISBN 9781107003217).

A non-expert who struggles to make good decisions and who turns to decision theory for help, might be more than a little surprised by what they find. If they read a standard treatment of the subject, they will find that they are assumed to be *logically omniscient*: they know all the logical facts about the propositions whose truth they have considered. Their beliefs are also assumed to be *logically closed*: if they believe each of a set of propositions S, then they believe everything that can be deduced from S. Finally, they are assumed to be *maximally opinionated*—they have assigned precise probabilities and cardinal utilities to each possible state of the world that can be formulated via S. The normative core of standard decision theory consists of some very weak axioms for these probabilities and cardinal utilities, plus the advice to maximize their expected utility, which is a function of the probabilities and cardinal utilities for each possible state of the world given each possible choice that they can make. The non-expert might understandably react by saying that this theory is too idealized to be useful for human beings. It is this criticism that Richard Bradley addresses with patience, rigour, and ardour in this book.

His focus is firmly on normative decision theory (how we should reason) rather than descriptive decision theory (how we actually reason). Of course, the two cannot be fully separated—what we *can* do has implications for what we *should* do. In Part I, Bradley introduces standard decision theory. These chapters are notable for their philosophical thoroughness. He identifies important assumptions, clarifies them, and links his discussions with relevant controversies. For example, standard decision theory might seem to assume a sort of behaviourism, due to its emphasis on revealed preferences over introspection for the measurement of beliefs and desires. However, Bradley argues that a type of pragmatism also suffices for this emphasis. This version of pragmatism only gives a methodological and explanatory 'priority' for revealed preferences over introspection.

In Part II, Bradley develops and extends Richard Jeffrey's decision theory. Bradley exposits this theory with tremendous clarity in chapter 5, before extending it to conditional attitudes in chapter 6, and to conditional propositions in chapter 7. With these building blocks in place, Bradley constructs a deep analysis of a broad range of issues in Part III. Chapter 8 will benefit readers who are interested in the relationship between possible worlds and decision theory. In chapter 9, Bradley proves that standard decision theory (Bayesian expected utility theory) has a place within his framework as a special case. Finally, in chapter 10, Bradley defends an extension of the Bayesian learning theory. This extension drops some unrealistic idealizations, like the assumption that a decision-maker's evidence consists of propositions that they believe with complete certainty.

Bradley's relaxation of common idealizations in decision theory reaches its fullest extent in Part IV. Here, the face of the decision theory truly starts to look much more human. In chapter 11, he drops the requirement of *full awareness*: that the decision-maker is aware of all possible actions, all possible states of the world relevant to the decision problem they are facing, and all the possible consequences of any choice that they make. As Bradley notes, while *full awareness* might be a desirable thing to have (notwithstanding blissful ignorance!) it hardly seems to be a requirement of rationality. He also drops the requirement to be *maximally opinionated*: that one has precisely determinate belief, desire, and preference attitudes towards all possible states of the world. This requirement does not even seem to be an ideal. The result is 'Imprecise Bayesianism,' in which a person's belief states are represented by a set of probability functions, rather than a single function. He formalises agent's

belief in a proposition A by the interval of all the different values for A given by the different functions in the set. One important feature of this formalism is that we can distinguish between states of equivocation (when you are confident that you should believe A just as strongly as not-A) represented by the degenerate interval [0.5, 0.5], from states of ignorance (when you are uncertain of what degree of belief to have) represented by wider intervals such as [0.2, 0.8].

The application of this theory unites the last few chapters, in which Bradley discusses changes of awareness of possibilities (chapter 12), ambiguous belief attitudes (chapter 13), and the representation of belief states (chapter 14). All of these chapters are stimulating and novel, but I was most intrigued by the last. Bradley notes that, under some circumstances, it is intuitive that an Imprecise Bayesian's beliefs should become more precise as they acquire more evidence. For instance, their belief interval that their 101st toss of a coin of unknown bias will land heads should be closer to [0.5, 0.5] after observing 50 'heads' and 50 'tails' in their 100 earlier tosses of the coin, given suitable prior beliefs in the distribution of 'heads' among the tosses. However, it is surprisingly hard to capture this relativity of belief-precision to evidence in Bradley's theory of imprecise beliefs. To solve the problem, Bradley suggests that an agent's attitude towards various intervals could vary with 'confidence thresholds'. These are standards for second-order confidence in belief intervals: the higher the threshold, the more confidence one needs in order to adopt a particular interval as one's belief state. This opens up the possibility that belief states can change with contextual factors, such as the stakes involved. There are parallels with some other imprecise beliefs theories that are worth pursuing in future research, such as with Henry E. Kyburg's system of Evidential Probability, which has a similar relativization of belief states to confidence thresholds.

Even though Bradley's main text is mostly informal, a background in formal methods (logic, basic set theory, and algebra) is necessary for following all of his explanations and arguments. For technically-minded readers, Bradley provides an excellent appendix of formal proofs. He also begins every chapter with an account of what is to come and finishes it with a concise summary of the chapter's contents. Consequently, despite the intricate thickets of some of his arguments, it is quite easy to find the way again when lost.

Bradley is successful in his goal of weakening many of the idealizations of standard decision theory. Yet this achievement comes at the cost of accentuating another common criticism, which is standard decision theory's thinness. Bradley is a subjectivist—he permits a huge variety of beliefs relative to almost any evidence. Unlike some Imprecise Bayesians, his tolerance extends to leaving the degree of precision in a person's beliefs to that person's subjective whim. This leaves the normativity of the decision theory looking very hollow. Once again, the usefulness of decision theory for human beings comes under threat.

Moreover, Bradley is not consistent in his subjectivism. An evidentialist Imprecise Bayesian, such as James M. Joyce, would require evidential reasons for precision in one's beliefs. Bradley rejects this requirement. However, when it comes to his argument for updating via conditionalization, he apparently leans on the assumption that one's conditional degrees in belief (one's belief in A, given full belief in B) should not change without some reason that requires them to change (191-192). Why should arbitrary precision of beliefs be permissible, but arbitrary changes in conditional degrees of belief be forbidden?

One norm for generating a more restrictive decision theory is the (in)famous Principle of Indifference. This demands that your initial degree of belief in each of a partition of possibilities should be $1/n$, where n is the number of possibilities in the partition. Bradley objects that the value of n will vary with the choice of language, which is a subjective choice. The Principle would thus provide no more real restrictions than subjectivism. I agree with his rejection of the Principle, but

not with his reason. It is true that language choice is exogenous to decision theory (which should be purely formal) but it does not follow that the choice is subjective. Arguably, there can be objective and intersubjective criteria for considering a language as better than another. Perhaps no such criteria exist and the choice is purely arbitrary. However, that is *prima facie* implausible, since languages vary in their possible expressive and illocutionary acts, as well as their complexity and ease of use for human beings. Overall, even if the choice of language is partly arbitrary, the Principle of Indifference would still be a meaningful constraint on our decision-making.

I strongly recommend this book. Bradley's contribution to decision theory will stimulate the subject for years to come. The book shows how rigour, novelty, and clarity can be combined in an accessible book. It is suitable for a wide audience across philosophy, economics, computer science, and other disciplines that use decision theory. The faces of the ideal reasoners in Bradley models are not fully human, but their countenances are a lot more human than what one can find almost anywhere else in decision theory.

William Peden, Università Politecnica delle Marche