Preface

When we reason about what to do we try to include everything that we think might affect the outcome of our decision. When choosing between different career options, for instance, we may take into account things such as the earnings potential, work hours, prestige and social benefit of the career options. But often we fail to include something in our deliberation, even though it could affect the outcome of our decision in an important way. It is this limited appreciation of the full scope of relevant possibilities – dubbed *limited awareness* – that is the topic of this Element. Navigating limited awareness is a pervasive aspect of our reasoning, and yet it has hitherto been relatively little studied.

The most dramatic cases of limited awareness are when we simply lack the conceptual resources to entertain the possibilities in question. For instance, there may be features of our solar system – not captured by our best scientific theories – that bear on whether or not to pursue a career as an astronaut that even the most educated person could not entertain. To give a historical example, when early industrialists reasoned about their actions, they failed to account for the possibility of a 'greenhouse effect' at the global level; this was well beyond their scientific comprehension of the world at the time. (In this historical case, the awareness gap was eventually filled, but there may be some awareness gaps that humanity will never fill.)

Other cases of limited awareness are more mundane. Sometimes we fail to account for relevant possibilities in our decision-making due to a momentary lack of perspective – a failure to consider some otherwise familiar contingency that bears on the decision at hand. For instance, we might reason impeccably about career choices on the basis of a select set of career characteristics but overlook other characteristics, say concerning health, that for some reason happen to be inaccessible to us at the time.

Given that it covers both of the above types of cases, the category of 'limited awareness', as used in this Element, is broad. That is, it covers both dramatic and more mundane sorts of conceptual inaccessibility. This may not be the only way to conceive of an agent's awareness and the limitations thereof; but it is one that suits the context of decision-making, which is the focus of this Element. (Arguably, any study of an agent's belief state and associated awareness must make reference to some function or role that the beliefs play, whether in decision-making or otherwise.)

To be more precise, according to our use of 'awareness', an agent counts as being aware of a possibility in a decision situation just in case he or she could in that situation, without either further evidence gathering or reflection,

factor the possibility into the decision. The reasons why an agent cannot factor a possibility into a decision in a particular situation can therefore be anything from limited conceptual resources to mere absent-mindedness (as one might call it). Indeed, the line between these, as far as decision-making is concerned, is far from sharp.

The main aim of this Element is to introduce the topic of limited awareness, and changes in awareness, to those interested in the philosophy of decision-making and uncertain reasoning. While it has long been of interest to economists and computer scientists, this topic has only recently been subject to philosophical investigation. At first sight limited awareness seems to evade any systematic treatment: it is *beyond the uncertainty* that can be managed. After all, an agent has no control over what contingencies she is and is not aware of at a given time, and any awareness growth takes her by surprise, at least in the sense that she can never predict *what* she might become aware of. On the other hand, agents apparently learn to identify the situations in which they are more and less likely to experience limited awareness and subsequent awareness growth. In other words, agents can predict *that* they will become more aware. How can these two sides be reconciled? That is the puzzle we confront in this Element.

We propose a way of conceiving limited awareness that does justice to its elusive character. While we build on earlier work of others, our analysis departs from this previous work in various ways. We accept that awareness growth can have radical and unpredictable effects on an agent's beliefs. But we argue that this does not preclude *anticipating* awareness growth. Moreover, we argue that, unlike the effects of experiencing 'unexpected' awareness growth, the effects of anticipating awareness growth both are quite predictable and can be captured without too radical a departure from the standard (Bayesian) model of rational preference and belief.

1 Introduction

1.1 Roadmaps to the Unknown

This Element is about our plight as reasoning agents in the world – that is, our plight as agents who seek to understand the world and how we can change it to best align with our ends. This requires some ingenuity because our perspective on the world is inherently limited. Think of it this way: our experience is confined to a more or less tiny patch of the world's history, so we can be certain of relatively little. The best we can do is try to account for all the contingencies – that is, all the ways the world *might* be – in at least as much detail as is relevant

for our purposes. In this way we can build ourselves a roadmap, so to speak, for navigating the unknown.

Consider, for instance, the reasoning of a single-minded conservationist who cares only about eradicating weeds and pests. At a particular juncture, our conservationist deems that she has a limited set of options: she can release a moth that will hopefully eat the non-native cactus plant known as 'prickly pear', or she can continue with the status quo, whereby all resources are devoted to manually uprooting the pear. Our conservationist judges that which of the two options will best realise her ends depends on whether or not the world is such that the moth will eat (and kill) the prickly pear if released, and this she is unsure about. That is, these are the two possible states of the world that the conservationist deems relevant to her decision.

The conservationist's problem is summarised in Table 1, where the columns represent the *states* of the world and the rows represent the available *options*; the interior cells of the table depict the *outcomes* in each state of choosing each option. So, depending on how the world turns out, or on the true state of the world, the options yield different outcomes. In this case, there are trade-offs between the options across the states: releasing the moth is better if the first state is true while the status quo is better if the second state is true.

The prickly pear decision is a highly stylised one, but it exemplifies the general predicament we reasoning agents face, day in and day out. We are condemned to live as gamblers. By our own lights, our choices are nearly always risky ventures – we are not assured that the world will turn out one way or another, and thus whether our ends will be served more or less well by any given choice of option. Not only are we limited by our practical circumstances – the options we have to change the world – but we are limited also by our epistemic circumstances – the ability we have to discern what is true of the world and thus which of our options serve us best and even what our options *are*.

To say that we live as gamblers in fact understates the precariousness of our position in the world as reasoners. For one thing, the gambles we face in the

Table 1 A simple roadmap for navigating the unknown, a.k.a. a decision model

	moth eats pear if released	moth fails to eat pear if released
Release moth	pear eradicated	pear thriving, wasted resources
Status quo	pear thriving	pear thriving

pursuit of our ends are not like games of roulette or dice for which the proba-
bility of the outcomes is typically thought to be objective and easy to calculate.
(We'll briefly return to this point a little later; it has been much studied else-
where.) The other thing – the main topic of this Element – is that we typically
do not have a good grasp of what *are* all the possible outcomes or contingencies
that are relevant to the decision at hand. That is, it is not just that we confront the
world not knowing which of the possible states of the world is actual: we do not
even know what are the pertinent possibilities to begin with. Throughout the
Element we refer to this latter predicament as *limited awareness*. Another form
of limited awareness that will occasionally come up in this Element concerns
the options available to an agent. In addition to not knowing which amongst
the options she considers will best serve her ends, an agent may often not even
know what *are* the options available to her.

The twenty-first-century reader may indeed have been struck by the limited
awareness of our conservationist introduced in Table 1. This stylised example
is in fact inspired by a historical episode in environmental management in Aus-
tralia. (The prickly pear episode in Australia in the 1920s had a happy ending,
as the moth that was released did in fact eradicate the highly invasive cactus.
But there were other prominent cases of introduced biological pest controls in
Australia that did not end so well.) With the benefit of hindsight, we can see that
the conservationist failed to appreciate the complexity of decisions to introduce
a biological pest control; she failed to consider other contingencies that were
relevant to her decision, such as that the introduced moth might itself become
a pest, eating native plant species instead of the target cactus plant. In addi-
tion, modern conservationists may see that our simple-minded conservationist
failed to consider other viable options – say, a targeted chemical pest control
or an alternative biological pest control – which might have served her ends at
a lower risk to the ecosystem.

Let us give another stylised historical example that also highlights – and
perhaps even more clearly – limited awareness due to the body of scientific
knowledge available to the relevant decision-makers at the time. Although
the possibility that human activity could change the climate through the
'greenhouse effect' had been discovered during the nineteenth century, it was
only in the 1970s that it became relatively widely known that greenhouse gas
emissions were wreaking havoc on our planet. Thus, when the first hydropower
plant was built in Iceland in 1904, the country's contribution to climate change
did not figure in the reasoning of the country's decision-makers. At the time,
coal was the most common energy source in Europe, and importing coal instead
seemed to some to be a viable alternative to building hydropower plants. Today,
about 55 per cent of Iceland's energy consumption comes from hydropower

and only about 2 per cent comes from coal, compared to a global average of 6 per cent from hydropower and 25 per cent from coal. So, given the pressing need to tackle the climate crisis – and since the climate impact of hydropower is generally much lower than that of burning coal – the decision to invest in hydropower was arguably right. Nevertheless, today we see that the decision was not based on all the best reasons; after all, the decision-makers in question were unaware of one of the best reasons for choosing hydropower over coal. In addition, these decision-makers were, we can safely assume, unaware of some of the options for generating significant amounts of energy that we are aware of today, such as those harnessing wind and solar power.

Examples from history allow us to witness limited awareness and subsequent growth in awareness. But fast forward now to the present. A little reflection suggests that limited awareness is not something that we reasoners have overcome. We continue to face novel scenarios and have our own forms of limited awareness. An example we focus on (in Section 6) is solar radiation management. This is a technique that could reduce (and perhaps even revert) climate change. But even its proponents admit that predicting the consequences of adopting solar radiation management on a global scale goes beyond today's scientific knowledge. Or consider carbon capture and storage. This is a relatively new technology, which consists in capturing carbon dioxide in the atmosphere and typically storing it underground, and which could be critical in combating the climate crisis. However, important uncertainties remain – for instance, about the impact of long-term underground storage of the captured carbon. Again, since this is a rather new and radical technology, the history of technological innovation would seem to suggest that it could result in consequences of which we are currently unaware.

It is these types of trying epistemic circumstances – the unavoidably parochial view of the world held by agents ranging from private citizens to individual public servants to the global community – that is the topic of this Element. The examples we will appeal to include dramatic cases of limited awareness and subsequent growth that involve novel combinations of concepts (e.g., a 'greenhouse effect' at the global level prior to the twentieth century) or even novel concepts simpliciter (e.g., an 'electron' prior to the late nineteenth century). But we will also appeal to more mundane cases of limited awareness and subsequent growth due to temporary shifts in attention or imaginative ability. The plight of the conservationist, an example we will pursue in this introductory section, arguably lies somewhere in the middle of the spectrum.

There are two things to say about our liberal stance on what counts as limited awareness that may help orient the reader from the outset as to the target of

our inquiry and our approach. The first is that, as the reader may well discern, we examine limited awareness in a practical decision-making context (and as such we will be appealing to and extending the tools of 'Bayesian decision theory'). While we focus on decision-making, we have little to say about an agent's basic values or ends; we simply take them as given. (That is, we leave the analysis of values or ends for others to address.) Nonetheless these ends have an important bearing on our epistemological project. It is not just that we are ultimately interested in how an agent reasons about what to do to further her ends. Our very understanding of her epistemic state and associated (limited) awareness is intimately tied to the pursuit of her ends.

Moreover, we doubt whether an agent's epistemic state and its limitations can even be well understood in the absence of *some* functional role that the epistemic state plays. The functional role that we are interested in is decision-making, but we allow that others may have different projects in mind and may thus understand an agent's epistemic state and her limited awareness in different ways that are moreover less liberal about what is genuine limited awareness and what is a mere mistake. That is, although we think that limited awareness is an important phenomenon that may be explored in a range of guises, by appeal to a variety of models, our particular interest is limited awareness in the context of decision-making. As such, we deem limited awareness to be concerned with whatever is the agent's decision frame at the time. In particular, we take an agent to be aware of a possibility, in a given decision situation, just in case she could in that situation – without either further evidence gathering or, say, overcoming any defects in imagination – factor the possibility into her decision. Her awareness may in this sense be limited and subject to growth.

Our conservationist, for instance, may well have many (at least implicit) ideas about the way the world is, including the weather, her family and friends, and so on. But these ideas are in a sense idle, at least in the context of her current options and ends, which concern the eradication of pests. Generally speaking, there may or may not be a richer story to tell about an agent's epistemic life. This Element, however, aims only to capture a part of this story. When we talk of an agent's epistemic perspective, we mean her current views about the possible contingencies, or ways the world might be, in so far as those contingencies play a role in her reasoning about what to do *now* to further her ends. In other words, an agent's epistemic perspective is relative to a decision problem, as we model it.

The second thing to note is that, while we thus seem to engage with rather hapless agents – including agents that may seem quite far from any ideal

state of awareness, even accounting for the limitations of the best science of the day – our project remains normative. We will examine how an agent *should*, rationally speaking, navigate her limited awareness and awareness growth. In particular, we take as our starting point agents' differing degrees of awareness – and, as we shall later see, the different extent to which agents are aware of their unawareness – and we ask what principles of rationality such agents should satisfy, for instance, when their awareness grows or when they simply predict awareness growth. So, the Element seeks to answer normative questions about agents who are less than epistemically ideal, in that they lack full awareness.

Now, any normative project of this kind will inevitably to some extent be prescriptive. That is, the principles of rationality we discuss are useful not just for *assessing* the rationality of agents but also for *guiding their deliberations*. But our primary aim is the former rather than the latter; we do not set out to offer principles that it would necessarily be wise (or even possible) to *apply* whenever one finds oneself in, say, a state of growing awareness or anticipated awareness growth. Still, we hope that by, for instance, revealing what principles one should ideally satisfy in situations of limited unawareness – and, moreover, illuminating the nature of (un)awareness and the situations where people have previously been demonstrably unaware – this Element can help readers make better decisions and reach more justified conclusions when they find themselves in such situations.

1.2 Internal Consistency and Its Limits

Since we will investigate limited awareness in the context of an agent's decision-making, we will appeal to standard (Bayesian) decision theory as our starting point. In the remainder of this introductory section we will explain how we will build upon standard decision theory – why it does not accommodate limited awareness and what we will seek to fill in, in the Element.

The standard decision-theoretic account of our reasoning goes beyond simple roadmaps such as those we have described. The roadmaps that the theory offers account not only for the supposed possible contingencies or ways the world might be but also, typically, their relative plausibility. To be *rational*, i.e., to reason well, one's judgements of relative plausibility must be *internally consistent*. Another requirement of rationality is that one's judgements of relative desirability be internally consistent. Indeed, decision theory can be understood as a theory of internal consistency. It tells how our epistemic and evaluative

judgements or attitudes must 'hang together' so as to yield clear choices of action that are not self-defeating with respect to our ends.

This Element is about the limits of internal consistency, in particular due to an agent's *(limited) awareness*, or what she perceives to be the possible contingencies or ways the world might be. But we need an understanding of the guidance that internal consistency can provide in order to see what are the shortcomings of this guidance. In what follows, we start by artic-ulating the guidance (1.2.1) before looking more closely at how arguments from internal consistency work (1.2.2) and what are their inherent limitations (1.2.3). One way to understand the point of this Element is to consider that we want to go beyond the uncertainty that decision theory typically deals with – that is, beyond the type of uncertainty that can be treated normatively in terms of internal consistency. However, we acknowledge that we are only taking one step beyond this uncertainty and that further steps may have to be taken.

1.2.1 Introducing Probabilities

We said that agents consider the relative plausibility of the possible ways the world might be. Put differently, agents have varying *degrees of confidence* – also known as *degrees of belief*, or as *credences*, which is the term we shall mostly use – in ways the world might be. It is as if they weigh the competing possibilities on a scale with multiple arms. The common wisdom is that, as an arm gets more weight, the others should collectively get less weight. To be more precise: credences are rational only if they can be represented as probabilities. This norm is often referred to as *probabilism*. For instance, if our conserva-tionist assigns much weight, or has relatively high credence, say, of 0.9, in the moth eating the pear if released, then on pain of inconsistency she must assign little weight, or have relatively low credence, here 0.1, in the moth *not* eating the pear if released.

Let us more thoroughly describe our conservationist's credences, as perti-nent to the choice problem depicted in Table 1. As noted, what matters for determining how well her options realise her ends is whether the released moth will eat the prickly pear or not – that is, which of these states of the world is actual – which we can denote M and $\neg M$ respectively. Strictly speaking, our conservationist is also unsure about what she will do, whether she will release the moth or not, denoted R and $\neg R$ respectively. This yields four rel-evant possibilities for how the world might be: $R\&M$, $R\&\neg M$, $\neg R\&M$ and $\neg R\&\neg M$. We assume that our conservationist has credences in each of these

fine-grained possibilities or *outcomes* that are each non-negative and together sum to one. Her credences in all other propositions involving M and R can be derived in conformity with the probability calculus. Moreover, presumably our conservationist's credences in *M* versus ¬*M* do not depend on her credences in *R* versus ¬*R*. That is, $P(M|R) = P(M|¬R)$, where *P* represents the agent's credences and $P(M|R)$ denotes her *conditional* credence in *M* given *R*. That is, in this case we have *act-state probabilistic independence*, but this need not always be so.[1]

Table 1 is the most economical depiction of our conservationist's choice problem. But note that the view of the world she brings to bear on this choice problem, and her associated credences, may be somewhat more complicated. Perhaps she entertains other potential properties of the world in an effort to form judgements about the relevant states of the world. For instance, perhaps our conservationist recognises that there may or may not be a drought during the year following the potential release of the moth, denoted *D* and ¬*D* respectively. She does not care about droughts. Our assumption is that she cares only about the eradication of weeds and pests. So in a sense whether or not there is a drought does not matter to her. Nonetheless, the consideration of whether there will be a drought may assist our conservationist in forming her credences in *M* and ¬*M*. After all, by the law of total probability, $P(M) = P(M\&D) + P(M\&¬D)$. Plausibly, our conservationist arrives at a settled credence in *M* by considering her 'component' credences in $P(M\&D)$ and $P(M\&¬D)$. This is to say that our conservationist's roadmap may look more like Table 2.

In general, there is a *space of propositions describing ways the world could be* about which the agent has an opinion that bears on her practical reasoning at some given time. This space of propositions about which she has an opinion is assumed to have a certain completeness in structure. In technical terms, it forms an algebra $\mathcal{F}$ with the following characteristics (which means that it is what is called a *Boolean* algebra):

- $\mathcal{F}$ contains a contradictory proposition ($\perp$).
- $\mathcal{F}$ contains a tautologous proposition ($\top$).
- $\mathcal{F}$ is closed under disjunction, conjunction and negation. That is, if *A* and *B* are in $\mathcal{F}$, then $A \lor B$, $A\&B$ and ¬*A* and ¬*B* are also in $\mathcal{F}$.

---

[1] The knowledgeable reader may discern that our presentation of the agent's decision model follows that of Jeffrey (1965), as opposed to Savage (1954).

Table 2 A more detailed roadmap for navigating the unknown

	moth eats pear; drought	moth eats pear; no drought	moth fails to eat pear; drought	moth fails to eat pear; no drought
Release moth	pear eradicated	pear eradicated	pear thriving, wasted resources	pear thriving, wasted resources
Status quo	pear thriving	pear thriving	pear thriving	pear thriving

Table 3 A yet more detailed roadmap for navigating the unknown

	moth eats pear; drought; pest	moth eats pear; drought; no pest	moth eats pear; no drought; pest	moth eats pear; no drought; no pest	...
Release moth	pear eradicated, pest	pear eradicated, no pest	pear eradicated, pest	pear eradicated, no pest	...
Status quo	pear thriving	pear thriving	pear thriving	pear thriving	...

The rational agent has credences in the propositions in $\mathcal{F}$ that can be represented by a probability function P. That is, $P(A) \in [0, 1]$ for all A in $\mathcal{F}$; $P(\bot) = 0$; $P(\top) = 1$; $P(A \vee B) = P(A) + P(B)$ for all mutually exclusive A and B in $\mathcal{F}$.

1.2.2 Rationality as Internal Consistency

Why think that rational credences are probabilities? There are various arguments for this position. A relatively straightforward one is known as the 'Dutch book argument'.[2] It turns on the claim that an agent's credences are effectively her 'betting odds' or the proportion of the stakes she'd be willing to pay for a bet that yields the stakes if the proposition in question turns out true but yields nothing otherwise. It is shown that if and only if her betting odds over the space of propositions conform to the probability calculus, the agent is *not* vulnerable to accepting a set of bets that would guarantee her a sure loss

---

[2] An argument like this was first suggested by Ramsey (1926). For a recent overview of different Dutch book arguments, see Richard Pettigrew's Element in this series (Pettigrew 2020).