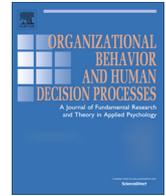




Contents lists available at ScienceDirect

## Organizational Behavior and Human Decision Processes

journal homepage: [www.elsevier.com/locate/obhdp](http://www.elsevier.com/locate/obhdp)

## Uncovering unknown unknowns: Towards a Baconian approach to management decision-making

Alberto Feduzi <sup>a,b,c,\*</sup>, Jochen Runde <sup>b,1</sup><sup>a</sup> Department of Financial and Management Studies, SOAS University of London, Thornhaugh Street, Russell Square, London WC1H 0XG, UK<sup>b</sup> Judge Business School, University of Cambridge, Trumpington Street, Cambridge CB2 1AG, UK<sup>c</sup> Department of Economics, University of Rome III, Via Silvio D'Amico 77, Rome 00145, Italy

## ARTICLE INFO

## Article history:

Received 29 April 2013

Accepted 7 April 2014

Available online 5 May 2014

Accepted by Francesca Gino

## Keywords:

State space construction

Unknown unknowns

Black swans

Inductive methods

Organisational and management

decision-making

Cognitive biases

## ABSTRACT

Bayesian decision theory and inference have left a deep and indelible mark on the literature on management decision-making. There is however an important issue that the machinery of classical Bayesianism is ill equipped to deal with, that of “unknown unknowns” or, in the cases in which they are actualised, what are sometimes called “Black Swans”. This issue is closely related to the problems of constructing an appropriate state space under conditions of deficient foresight about what the future might hold, and our aim is to develop a theory and some of the practicalities of state space elaboration that addresses these problems. Building on ideas originally put forward by Bacon (1620), we show how our approach can be used to build and explore the state space, how it may reduce the extent to which organisations are blindsided by Black Swans, and how it ameliorates various well-known cognitive biases.

© 2014 Elsevier Inc. All rights reserved.

“How often have I said to you that when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth?” (Sherlock Holmes to Watson, Doyle, 1890).

## Introduction

9/11, the Gulf oil spill and, more recently, the uprisings in the Middle East and North Africa and the Tohoku earthquake, tsunami and subsequent problems at the Fukushima nuclear plant in Japan, are all examples of events that disrupt the lives of millions and which, before they occur, are simply not on the radar of many of those affected. In the same way, if somewhat less dramatically, organisations are often buffeted by events that they had not even registered as possibilities prior to their occurrence, and which may have a considerable impact on their fortunes.

Interest in such events—sometimes called Black Swans or, prior to their occurrence, unknown unknowns—is currently running high in the organisational and risk-management literature

(Cunha, Clegg, & Kamoche, 2006; Lampel, Shamsie, & Shapira, 2009; Lampel & Shapira, 2001; Loch, De Meyer, & Pich, 2006; McGrath & MacMillan, 2009; Mullins, 2007; Pich, Loch, & De Meyer, 2002; Rerup, 2009; Sommer & Loch, 2004; Sommer, Loch, & Dong, 2009; Starbuck, 2009; Weick & Sutcliffe, 2007), some authors going so far as to argue that the domain of unknown unknowns is one “to which much of contemporary business has shifted” (Snowden & Boone, 2007, p. 74) and others issuing stark warnings to the effect that “companies that ignore Black Swan Events will go under” (Taleb, Goldstein, & Spitznagel, 2009, p. 79). The notion of “unknown unknowns” is far from new, however, and was already familiar in engineering and project management circles well before entering the popular consciousness via US Defence Secretary Dennis Rumsfeld’s (2002) famous press conference (Wideman, 1992). And it has never been far from the surface in discussions of the problems of arriving at a complete list of “states of the world” in decision theory, that is the problems of generating and evaluating hypotheses about how the future will unfold and, more generally, the issues associated with the framing and structuring of decision problems (Bazerman & Moore, 2009; Miller, 2008).

The two traditions that have contributed most to these discussions are the Carnegie School (Cyert & March, 1963; March & Simon, 1958; Simon, 1947, 1955) and Behavioural Decision Theory (Edwards, 1954, 1961; Einhorn & Hogarth, 1981; Fischhoff, Slovic, &

\* Corresponding author at: Department of Financial and Management Studies, SOAS University of London, Thornhaugh Street, Russell Square, London WC1H 0XG, UK. Fax: +44 (0)20 7898 4829.

E-mail addresses: [af19@soas.ac.uk](mailto:af19@soas.ac.uk) (A. Feduzi), [j.runde@jbs.cam.ac.uk](mailto:j.runde@jbs.cam.ac.uk) (J. Runde).

<sup>1</sup> Fax: +44 (0)1223 339701.

Lichtenstein, 1977; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1979a; Tversky & Kahneman, 1974). While distinct in many ways (Shapira, 2008), both take the form of powerful critiques of the “canonical model” in individual decision-making—classical Bayesianism as represented by Bayesian conditionalization and decision theory à la Savage (1954)—as a description of what practicing decision-makers do.<sup>2</sup> These critiques have led in turn to a growing body prescriptive work, mostly in psychology and management science, offering tools and techniques to help decision makers counteract cognitive biases, broaden decision frameworks and actively search for unknown unknowns (Hirt & Markman, 1995; Larrick, 2009; Loch et al., 2006; Lord, Lepper, & Preston, 1984; McGrath & MacMillan, 1995; McGrath & MacMillan, 2009; Schoemaker, 2002, 2004). Many of these tools and techniques have been used by organisations to “de-bias” practicing decision-makers (Heath, Larrick, & Klayman, 1998; Larrick, 2009).

Taken together, these different bodies of work provide a significant contribution to our understanding of the practicalities of framing and structuring of decision problems in general and the problems of state space construction and unknown unknowns in particular. There is however rather less on these topics from a normative perspective. We will argue that this situation can be attributed to the continuing influence of the canonical model in its normative capacity, which is largely silent on the problems of state space construction and uncovering unknown unknowns. There is an important gap to be filled here since normative models provide the necessary standards for comparison and evaluation that are fundamental to the progress of both descriptive and prescriptive work (Baron, 2004, 2012).

Our aim in this paper is accordingly to introduce a specific normative approach, Bacon's (1620) method of eliminative induction, and to use this to develop a prescriptive approach to state space construction and uncovering unknown unknowns. While we recognise that there are many competing methods of enquiry in the philosophical and wider literature (e.g. Mill, 1843; Peirce, 1898; Popper, 1959), we focus on Bacon's for the central role it assigns to hypothesis generation in the process of hypothesis evaluation. This feature makes it especially suited to dealing with the specific problems that will concern us in this paper. However, since Bacon's account was developed with ideal experimental situations and relatively simple and well-defined hypotheses in mind, it needs to be adapted for use in the non-experimental, complex and often ambiguous situations faced in management. This is what our prescriptive approach seeks to do. We will show that, apart from the ways in which it may facilitate state space construction and the uncovering of unknown unknowns, it also encapsulates many of the de-biasing techniques that have been proposed in the literature, and to this extent provides a unified and implementable approach to offsetting many well-known cognitive biases.

Our argument begins with a literature review and a section that fixes terms and introduces some useful distinctions. This is followed by a section in which, following the same general strategy used by authors like Simon (1982) and March (1991) to tackle problems associated with the canonical model, we first show why Bayesianism does not address the problem of state space construction and is structurally unsuited to dealing with unknown unknowns, and then outline Bacon's original method and why it promises the resources to address these issues. We then propose

our prescriptive version of his method for use in managerial context, and show how this can be applied in practice and what its virtues and limitations are. We close with brief discussions of some theoretical aspects of our approach, some managerial and organisational implications, possible future work, and a short conclusion.

## The literature

Although the problem of unknown unknowns has only come to the fore significantly in the management literature over the last decade or so (Cunha et al., 2006; Lampel & Shapira, 2001; Lampel et al., 2009; Loch et al., 2006; McGrath & MacMillan, 2009; Mullins, 2007; Pich et al., 2002; Rerup, 2009; Snowden & Boone, 2007; Sommer & Loch, 2004; Sommer et al., 2009; Starbuck, 2009; Taleb et al., 2009; Weick & Sutcliffe, 2007), it has a long history in a variety of disciplines including economics (Shackle, 1979, 1983), the decision sciences (Keller & Ho, 1988) and the psychological literature (see Bazerman & Moore, 2009; Miller, 2008). The problem is closely related to the practicalities of constructing the state space, namely the generation and evaluation of candidate hypotheses about how the world might turn out, and, more generally, to wider issues relating to the framing and structuring of decision problems.

These issues have received considerable attention in the literature on management decision-making, starting with the Carnegie School represented by Simon (1947, 1955), March and Simon (1958), Cyert and March (1963) and more recently, Levinthal (1997), Gavetti and Levinthal (2000, 2001), and Gavetti, Levinthal, and Ocasio (2007).<sup>3</sup> The story begins with early critiques of the canonical model focusing on the idea that choice behaviour cannot be reduced to the optimisation of a well-specified choice set (Simon, 1955). The broad argument was that practicing decision-makers are typically not presented with decision problems already neatly broken down into exhaustive lists of acts, states and consequences, that there are limits on their capacity to acquire the necessary information and make reasoned judgements on its basis even if it were available, and that they accordingly do not always act as the canonical model predicts. What practicing decision-makers tend to do instead, according to Simon and his followers, is consider only a few alternatives at a time, assess them (semi-) sequentially rather than simultaneously, and stop searching when they identify an alternative that satisfies some kind of performance criterion. This in essence is the theory of satisficing behaviour for which the Carnegie School is famous. Operating under conditions of bounded rationality as they are, moreover, satisficers are likely to be vulnerable to unknown unknowns. Post decision surprises, pleasant or otherwise, are an unavoidable fact of life (March, 1994, p. 6).

While the Carnegie School has made a lasting contribution in drawing attention to the cognitive limits on human decision-making and learning, its original concern was primarily with establishing that practicing decision-makers regularly deviate from the canonical model and with developing models that relaxed one or more of the strictures associated with it. It paid rather less attention to the nature of those deviations, and, more generally, to the many particular directional biases that affect decision-makers'

<sup>2</sup> We adopt the conventional distinction between descriptive, normative and prescriptive models in the study of decision-making (Baron, 1985; Bell, Raiffa, & Tversky, 1988). Whereas descriptive models aim to portray what practicing decision-makers actually do and normative models aim to portray what decision-makers should do in ideal circumstances, prescriptive models aim to provide tools and techniques to help practicing decision-makers come closer to achieving normative ideals.

<sup>3</sup> There are of course contributions to organization theory that examine other influences on management decision-making (see the review by Hodgkinson and Starbuck (2008)). Moreover, members of the Carnegie school have themselves proposed models such as the garbage can model of organizational decision-making (Cohen, March, & Olsen, 1972), that represent a movement away from the paradigm of individual decision making adopted by Simon and his followers. We concentrate on the original Carnegie approach here because its focus on the role of information in decision-making comes closest to our concerns in the present paper.

judgement (Bazerman & Moore, 2009, p. 5). This gap has been filled by Behavioural Decision Theory over the last 30 years or so, greatly amplifying the Carnegie school critique, and throwing significant light on how decision-makers gather and use information (Bazerman & Moore, 2009; Camerer, Loewenstein, & Rabin, 2004; Edwards, 1954, 1961; Einhorn & Hogarth, 1981; Fischhoff et al., 1977; Kahneman & Tversky, 1979a; Kahneman et al., 1982; Tversky & Kahneman, 1974).

That human reasoning is subject to systematic biases is the guiding theme in Behavioural Decision Theory. Amongst the many forms of human reasoning that it has investigated under this aspect are ones that will concern us in this paper, namely those that lead decision-makers to produce overly narrow decision frames (Bazerman & Moore, 2009; Larrick, 2009). In particular, we focus here on the shortcomings that lead decision-makers to produce overly narrow views of the future by affecting the ways in which they come up with, and collect and use evidence to evaluate, hypotheses about how the future might unfold (for a comprehensive review, see Heath et al., 1998; Larrick, 2009).

With respect to hypothesis generation, Behavioural Decision Theory has shown empirically that people tend to look for hypotheses that put them in a favourable light (Muller & Riordan, 1988), stop searching as soon as they find a plausible candidate hypothesis (Gregory, Cialdini, & Carpenter, 1982; Hoch, 1984), fail to generate alternative hypotheses (Gnepp & Klayman, 1992; Mynatt, Doherty, & Dragan, 1993) and, where they do, generate hypotheses that are not sufficiently different to each other (Fischhoff, Slovic, & Lichtenstein, 1978; Gettys, Pliske, Manning, & Casey, 1987). With respect to hypothesis evaluation, Behavioural Decision Theory has shown that people tend to rely on unduly small samples of information because they underestimate the benefits of larger samples (Tversky & Kahneman, 1971), consider only the most readily available information (Tversky & Kahneman, 1973), look for evidence that confirms pre-existing hypotheses, and consider only part of the information acquired (Anderson, 1995; Klayman, 1995; Kunda, 1990; Wason, 1960; Zuckerman, Knee, Hodgins, & Miyake, 1995).

The work of the Carnegie School and Behavioural Decision Theory is largely descriptive in nature, concerned with capturing what practicing decision makers actually do and how this tends to deviate from the canonical model. However, it has precipitated a body of prescriptively oriented work in the organisational and psychological literature concerned with developing techniques to assist decision-makers improve on their performance. Some of these techniques bear on the issues that will concern us below, including techniques to aid decision-makers broaden their decision frames—and, in particular, avoid the problem of generating overly narrow views of the future (Bazerman & Moore, 2009; Heath, Larrick, & Klayman, 1998; Larrick, 2009)—and actively search for unknown unknowns. Some are relatively formal in nature, and include scenario analysis (Schoemaker, 2002, 2004), trial-and-error learning (Pich et al., 2002; Sommer & Loch, 2004) and discovery-driven planning (McGrath & MacMillan, 2009). Others, sometimes referred to as “cognitive repairs”, are more informal and include simple procedures such as “consider the opposite” (Lord et al., 1984), “consider an alternative” (Hirt & Markman, 1995), and the use of checklists for gathering information and evaluating alternatives (Larrick, 2009). Both varieties of these techniques have been deployed in organisations to “de-bias” practicing decision-makers (Heath et al., 1998; Larrick, 2009).

The literature surveyed above has made a significant contribution to our understanding of the issues involved in generating, evaluating and then accepting or rejecting hypotheses about how the world might turn out. However, the progress it has shown on descriptive and prescriptive fronts is not matched by its progress on the normative front. The reason for this, in our view, is that

the canonical model is still widely regarded as the state of art from a normative point of view, and that this has led to a reluctance to look beyond the Bayesian inductive method associated with it. Unfortunately, Bayesianism has little to say about state space construction and uncovering unknown unknowns for reasons we explain below, and the contributions that reject or ignore Bayesianism—which many of the prescriptive contributions mentioned above do—often propose *ad hoc* procedures and recommendations that are not founded on a coherent inductive method. There is accordingly a gap for normative work on this subject, which might then inform prescriptive work of the kind we pursue below.

## Definitions

Although the term “unknown unknowns” has entered the jargon of management decision-making, there are differences in the literature over exactly what they might be (Loch et al., 2006; Mullins, 2007; Snowden & Boone, 2007; Sommer et al., 2009; Wideman, 1992). In particular, there are differences over whether they are possibilities or actualisations, whether they refer to events or states, and where use of the term extends variously to Black Swans, unpredictable surprises, unimagined events, unexpected events, unforeseen events, unforeseeable events and rare events (Runde, 2009). It is therefore necessary to fix terms.

In what follows an *unknown* is understood as a hypothetical event that may or may not go on to occur. From the point of view of a decision-maker, an unknown may be known or unknown. A *known unknown* is one the decision-maker imagines and regards as having a real possibility of occurring. Thus in the simple case of a toss of a classical die, the relevant known unknowns would generally be taken to be the elementary events 1, 2, 3, 4, 5 and 6. An *unknown unknown* is one that the decision-maker does not imagine and therefore does not even consider. Thus if the decision-maker is unaware of the existence of exotic dice and that the die being rolled is in fact seven-sided, then the event of a 7 would be an unknown unknown from her perspective. We can then further define a *Black Swan* (Taleb, 2007) as an unknown unknown that has gone on to occur, that is, an event that the person who goes on to be surprised by it did not even imagine as a possibility prior to its occurrence.

Note that we have defined unknown unknowns in a way that makes them subjective to the decision-maker. This means that an event experienced as a Black Swan by one person may not come as even a mild surprise to the next person. While someone who did not know about the existence of seven-sided dice would be extremely surprised by a 7, for example, someone who knew the game was being played with a die of this type would not. Further, and contrary to many peoples’ intuition, neither unknown unknowns nor Black Swans need necessarily be rare or low frequency events. Taking again the case of our seven-sided die, the underlying relative frequency of the 7 our naïve or unlucky gambler experiences as a Black Swan first time around may actually be relatively high (1/7 if the die takes the form of a lat-long polyisohedron for example, and higher in some of its non-symmetric versions).

Two key questions arise at this point. The first is whether unknown unknowns refer to isolable events or to what decision theorists call “states of the world”. In decision theory it is usually assumed that decision-makers are orientated towards the latter, that is, possible unfoldings of the world described in sufficient detail to determine the relevant consequences of each of the possible courses of action they might take. When people refer to unknown unknowns and Black Swans, however, they generally seem to have isolable events in mind, that is, particular occurrences described in ways that fall short of exhausting all decision-relevant features of the situation in which they may arise.

We will follow this usage, but bearing in mind that the existence of unknown unknowns in this sense implies that the states of the world in which they might arise must be unknown unknowns too. From a decision-theoretic perspective, the possibility of unknown unknowns in the form of isolable events that might occur implies an incomplete state space.

The second question is what it is about the world (which includes human actors and their activities) that gives rise to unknown unknowns and, therefore, to people being periodically surprised by Black Swans. Two ideas that often come up in this connection are emergence and epistemic constraints (Runde, 2009). The concept of emergence locates the problem in the world, namely that the world itself may be a source of novelty in periodically throwing out novel events, new forms of existence, phase changes and so on that are “emergent” in the sense of not being reducible to a fixed set of prior causes and therefore not foreseeable *ex ante* even in principle, on the basis of existing evidence about initial conditions, laws, and so on (Harper & Lewis, 2012; Lawson 2012).

Emergence in this sense is however neither a necessary nor a sufficient condition for the existence of unknown unknowns. It is not a necessary condition since all that unknown unknowns require is limits on what the decision-maker can imagine, due to epistemic constraints on her ability to collect and process evidence. It is not a sufficient condition because it is at least conceivable that a particularly prescient decision-maker—as futurists do from time to time—might be able to imagine emergent possibilities and their consequences, even if these can’t be directly inferred from the existing evidence. Unknown unknowns require no more than an inability to imagine some or other possibility, no matter what the source of this inability may be.

However, the distinction between the two possible sources of unknown unknowns suggests that, at a theoretical level, it is possible to distinguish between: (i) *knowable unknowns*, unknown unknowns that could have been transformed into known unknowns at some point in time in the absence of epistemic constraints; and (ii) *unknowable unknowns*, unknown unknowns that are emergent and therefore could not have been transformed into known unknowns at some point in time, even if it were possible to amass and process all information there was to know at that point. Thus the example of 9/11 with which we began falls into the category of knowable unknowns. While the events of that day were likely a Black Swan for most of us, they were not so for everyone. The idea of airliners being used as missiles had already been considered by the North American Aerospace Defense Command (NORAD) two years before 9/11, and NORAD had even run simulations of the World Trade Centre and the Pentagon being attacked in this way ([http://www.usatoday.com/news/washington/2004-04-18-norad\\_x.htm](http://www.usatoday.com/news/washington/2004-04-18-norad_x.htm)).

The possibility of knowable unknowns implies that, at least in principle, a subset of what would otherwise remain unknown unknowns could be “uncovered”, that is transformed into known unknowns, by overcoming the epistemic barriers that engender them. This is the idea we will pursue in what follows.

We now turn to what different inductive methods have to say about constructing the state space and uncovering unknown unknowns. From here on we will focus principally on knowable unknowns, henceforth referred to simply as unknown unknowns. However, the approach we propose below is not restricted to static situations and can also be used in dynamic/changing situations in which new, formerly unavailable information emerges over time, and there is the possibility that what were unknowable unknowns at one point in time become knowable at a later point in time. Since decision-makers never know what they do not know, and since this is so irrespective of whether or not the relevant unknown unknowns are knowable in principle, our approach applies in the

presence of unknowable as well as knowable unknowns, and, in addition to facilitating the constructing of the state space at a point in time, can be used to monitor the state space dynamically over time.

## Bayes and Bacon

Is there anything decision-makers can do to reduce the number of unknown unknowns they are likely to encounter? To answer this question, it is useful to begin with Bayesian decision theory, which serves as the benchmark for much of the literature on management decision-making. This will allow us both to locate and frame the problem of unknown unknowns with reference to the familiar canonical model, and to pinpoint why Bayesianism has so little to say about this problem. Once done, we will introduce the Baconian alternative.

### *Bayesian decision theory in the “small” and in the “large”*

Decision theory typically assumes that the decision-maker has to choose between competing “acts” leading to different “consequences” depending on which of a set of possible “states of the world” obtains. The decision problem can thus be modelled as a function  $F:AxW \rightarrow C$ , where  $A$  is the set of available acts  $a_i$  ( $i = 1, 2, \dots, m$ ),  $W$  the set of possible states of the world  $w_j$  ( $j = 1, 2, \dots, n$ ), and  $C$  the set of possible consequences  $c_{ij}$  ( $c_{ij} = F(a_i, w_j)$ ). In this setting, an act is any function  $a:W \rightarrow C$ , and the decision-maker chooses between acts on the basis of his “desires” and “beliefs”. Desires are generally expressed by a utility function defined over the set of possible outcomes, and beliefs by a probability function defined over the set of possible states.

The purpose of much of decision theory is to provide the decision-maker with the means to translate the foregoing information into an ordering of acts. Expected utility theory is by some distance the most widely accepted decision theory of this sort, and recommends that acts be ranked in terms of the sum of the probability-weighted utilities of their consequences. Expected utility theory with subjective probabilities is commonly called Bayesian decision theory, and is based on the following tenets:

- *Probabilistic beliefs (1)*: the Bayesian subject is always willing to assign a degree of belief to any proposition, event or hypothesis (de Finetti, 1937; Ramsey, 1926).
- *Probabilistic beliefs (2)*: the degrees of belief assigned by a Bayesian subject are always *coherent* in the sense of conforming to the laws of the probability calculus (de Finetti, 1937).
- *Bayesian updating*: when new evidence is acquired, the Bayesian subject modifies her probabilistic beliefs in accordance with Bayes’ updating rule.
- *Expected utility*: when facing a decision problem, the Bayesian subject maximises the expected utility of an action with respect to her Bayesian beliefs and *chooses* the action that leads to the highest expected utility.

Although these tenets are often represented as relatively innocuous, and as Herbert Simon and his Carnegie School colleagues saw immediately, they actually involve strong assumptions about the information available to the decision-maker. We will concentrate on one of these assumptions, namely that the decision-maker possesses an exhaustive list of the mutually exclusive possible states of the world relevant to a decision problem (sometimes called the “grand state space” assumption (Gilboa, Postlewaite, & Schmeidler, 2012) or, in Savage’s (1954) terminology, the “small world” assumption (Binmore, 2009)). On this assumption, if  $W$  represents the possible states of the world assumed in Bayesian

decision theory, then each element  $w$  of  $W$  is taken to describe one way the world might turn out in enough detail to determine the relevant consequences of each act (Savage, 1954, p. 9).

One of the consequences of always starting with a small world is that Bayesian decision theory effectively precludes “genuine” learning in the sense of uncovering new, formerly unimagined, possibilities. That is to say, any “genuine” learning must take place in advance of receiving any information that may lead to probabilities being updated, so that decision-makers have already eliminated the possibility of future surprises in the model they use to construct their beliefs. Note that we are not denying that there may be situations in which the small world assumption is justified, that is, where decision-makers are able to arrive at an exhaustive list of possible states of the world and nothing can ensue that is not on that list. However, in practice, management decision problems rarely present themselves in the sharp and comprehensive form assumed in Bayesian decision theory, that is, in a way in which it is immediately obvious what states of the worlds should be entertained (Gilboa & Schmeidler, 1995, pp. 605–608). Surprise is an unavoidable fact of life, and the assumption that decision-makers can anticipate every eventuality that might befall them is highly demanding.

What then about Bayesian learning and inference? After all, learning in the form of updating prior beliefs using Bayes’ rule is one of the cornerstones of Bayesianism. The difficulty here is that while new information, or indeed the mere exercise of imagination, can bring into view hitherto unrecognised states, the resulting shifts in the decision-maker’s beliefs cannot be described by Bayesian conditionalization. To see what is involved here, take the example of someone attempting to estimate the proportion of red balls by drawing from an urn she believes contains only red and black balls, and who, after drawing some red and black balls, proceeds to draw a yellow ball. The Bayesian would grind to a halt at this point, because Bayesianism precludes adding new states or updating a zero probability to a positive probability. The reason for this is that conditionalizing on information that a previously unarticulated possibility has been introduced is literally nonsensical, since such conditionalization presupposes there was a well-defined prior probability for that possibility in the first place (Earman, 1992). The Bayesian would thus be obliged to start over with a reformulated state space, re-specify her priors, and begin sampling and updating again. This process would have to be repeated whenever she encounters a state she had not previously considered. Crucially, none of such learning would be “Bayesian learning”, that is via updating priors using Bayes’ rule.

The upshot is that Bayesian decision theory and inference have no place for unknown unknowns. Further, to the extent that it treats the mind of the decision-maker as a “black box”, always with a given small world and in which all kinds of information about possible states is automatically and unproblematically translated into point-valued probabilities, Bayesianism has next to nothing to say about how to go about constructing the state space or what kind of evidence should be taken into account when doing so. We therefore now turn to an inductive approach proposed by Bacon (1620), which we believe offers resources to address these issues.

#### *The Baconian method of eliminative induction*

A perennial theme in the philosophy of induction is whether it is the *multiplicity* of evidential instances or the *variety* of evidential instances that matters most in the evaluation of hypotheses (Keynes, 1921). On the first view, induction proceeds by simple *enumeration*, that is, a generalisation is supposed to acquire support that varies in strength with the number of *positive instances* that verify it: from some observed evidence of properties,  $P$ , of some object,  $O$  (“ $O_1$  is  $F$ ,  $O_2$  is  $F$ ... $O_n$  is  $F$ ”), we infer that “All

$O$ —observed and not observed—are  $F$ ”. The multiplicity of instances thus gives ground for believing a hypothesis, and the intuition here is that the belief in the truth of that hypothesis ought to rise as confirming instances increase.

Yet it is clear that induction based on simple enumeration of individual instances cannot establish the truth of any hypothesis even if all evidential instances examined to date have been consistent with that hypothesis. The reason for this is that an instance being consistent with a hypothesis is not the same thing as that instance confirming the hypothesis. Francis Bacon (1620) argued that hypotheses about how nature works can never be justified merely by collecting favourable instances, and repudiated as “childish” the method of induction by simple enumeration. He gave two reasons for this (Cohen, 1970, 1977, 1989; Schum, 1994). The first is that, regardless of how many favourable instances have been observed in the past, it takes but one negative instance to undermine a generalisation. The second is that it is not the mere number of instances that should count, but also the *variety of circumstances* in which instances of the phenomenon under investigation are present.

This emphasis on negative and variative instances led Bacon to propose a new form of induction, *eliminative and variative induction*. On this method, (i) as a hypothesis can be eliminated on the basis of a single negative instance, evidence should be gathered and hypotheses tested with an explicitly eliminative mindset; and (ii) as the variation of circumstances may be regarded as a method of eliminating alternative hypotheses, experiments should be structured so as to yield instances that have the capacity to exclude them.

An investigator adopting Bacon’s method starts with an initial hypothesis to explain some observed phenomenon and then tests it against a series of alternative hypotheses that might also explain the same phenomenon. She conducts the tests by systematically *varying* the circumstances under which the experiment is performed, in order to *eliminate* each of these alternative hypotheses. The higher the number of tests passed by the initial hypothesis, the greater the investigator’s confidence in it, the intuition being that observed evidence of properties  $P$  of some object  $O$  that has been found under a greater variety of circumstances makes for a more severe test.<sup>4</sup> Alternatively, if the outcome of one of the experiments is inconsistent with the initial hypothesis, a modified hypothesis is then substituted and the process can begin again.

Bacon emphasised the role of what he called *instantiae crucis* or what are nowadays called “crucial experiments” in this process, for having the power to determine the direction of the investigation. Contrary to some modern interpretations, however, he was not suggesting that crucial experiments always lead to the decisive rejection of a hypothesis and proof of another (Cohen, 1980a; Hacking, 1983, p. 250). What he had in mind was rather a gradualist view of experimenters performing series of *successive* crucial experiments, always using evidence to eliminate rather than amass support for rival hypotheses, and where the hypothesis that resists these efforts is the one in which they should have most confidence (Cohen, 1970, 1977, 1989; Hacking, 1983; Platt, 1964). Indeed, Bacon’s method cannot produce conclusively certain results. Even if a hypothesis has passed many crucial tests and it is therefore supported by a high number of variative instances, a new variation of circumstances might eliminate it and confirm another (previously overlooked) hypothesis.

Von Frisch’s (1950) famous work on the behaviour of bees provides a good example of this method in action (Cohen, 1977, 1989). Von Frisch’s approach was to start with an initial hypothesis and

<sup>4</sup> Note that this is very different from the idea that, given a complete possibility space, the elimination of a hypothesis alone leads to an increase in the level of confidence attached to the remaining hypotheses.

then proceed by attempting to eliminate alternative explanations of the phenomena revealed by his experiments. For example, on the basis of observations of bees returning repeatedly to a transparent source of food (sugar-water) on a piece of blue card, he formulated the hypothesis that they discriminate between blue and other colours. He then proceeded to evaluate this hypothesis by running a series of tests of various alternative hypotheses:

1. that bees are colour-blind and identify their feeding-place by its shade of greyness, a possibility eliminated by surrounding the blue card with grey cards of all shades from white to black, all cards carrying food-containers but no food, and observing that bees continue to return to the blue card;
2. that bees recognise the relative location of the blue card, a possibility eliminated by rearranging the cards in many different ways, and observing that bees continue to return to the blue card;
3. that bees recognise the smell of the blue card, a possibility eliminated by observing that bees continue to the blue card even if the card is covered with a plate of glass;
4. and so on.

Von Frisch's method thus involves testing the initial hypothesis by varying experiments in a systematic way. If the outcome remains consistent with the initial hypothesis and the alternative hypotheses are eliminated, the initial hypothesis is regarded as less and less open to reasonable doubt. If the outcome fails to accord with the initial hypothesis, a modified hypothesis is then substituted.

The example demonstrates clearly why instantial variety is superior to instantial multiplicity. Von Frisch's hypothesis received greater support from bees returning to a blue-coloured source of food that was moved around several different locations, than it would have received from the same number of bees returning to a blue-coloured source of food that remained in the same place on an equal number of different occasions. The reason is that relative location was known to be a potentially relevant factor in studies of bees' recognition-capacities, that is, that bees have a good memory for places, and varying the location of the food therefore served to eliminate the not wholly implausible hypothesis that memory of place rather than colour was operative. Moreover, the example sheds light on the question of what evidence can logically be considered as confirming evidence. On this approach, mere multiplicity of instances is significant only for the replicability of test-results and not for the strength of support they provide. By running the same eliminative test over and over again, we might strengthen our belief about the reliability of this single test. But we do nothing to strengthen our belief about the extent to which any hypothesis holds up in different circumstances.

Since an initial hypothesis gains more and more evidential support as alternative possible hypotheses are eliminated, this method also constantly pushes the experimenter to actively explore the space of possibilities. To be sure, Von Frisch approached the problem of generating hypotheses and selecting evidential tests by referring to the available information (for instance, that bees have good memory and that different species of insects and birds were colour-blind and relied on recognition by scent). But he also made important progress during his investigations when new relevant variables were *discovered*, such as that of variation from a broken to an unbroken shape (Cohen, 1977, p. 131). The latter phenomenon was discovered because, in the process of eliminating hypotheses, any test involving shape discrimination tended to produce contradictory results until the manipulation of that variable was introduced into the explicit structure of the test. In short, on the method of eliminative induction, the discovery and evaluation of hypotheses are part of the same process, something that has

occasionally also been suggested in contributors to disciplines ranging from artificial intelligence (Buchanan, 1985) to chemistry (Leeson, 1977) and the philosophy of science (Kitcher, 1993; Norton, 1995; Platt, 1964). They are part of the same process because the evaluation of any hypothesis requires generating and testing possible alternatives to that hypothesis, and so driving the evaluator to think up ("discover") new hypotheses. Further, and contrary to the received view that the process of discovery is something that resists logical analysis (Popper, 1959; Reichenbach, 1951), Bacon's method is clearly one of systematic, reasoned investigation.

### **Towards a Baconian approach to management decision-making**

Unfortunately, management decision-makers are seldom in a position to perform controlled experiments of the kind Baconian eliminative induction was designed for, something that has a strong bearing on the extent to which they are able to perform "crucial experiments", decisive or otherwise. There are various issues here. First, the hypotheses in question are no longer possible explanations of an observed phenomenon, but hypothetical future states of the world. Second, unlike scientific experiments in which the hypotheses usually concern some or other property of an isolated and relatively stable mechanism or substance about which knowledge can improve over time, hypothetical states of the world are complex things that occur only once if ever they do. Consequently, third, states of the world are not as easily and clearly individuated as possible experimental outcomes. Finally, because of the noisy nature of the business environment, it is often difficult to identify evidence that unambiguously implies the rejection of a specific state.

There are also differences in respect of the actors involved, and specifically that management decision-makers may be more susceptible than are laboratory scientists to the kind of cognitive issues highlighted by Behavioural Decision Theory. Prominent here are the tendencies to come up with overly narrow ranges of possible states, to fail to produce states that differ in a significant way, to focus on preferred states, and to concentrate on evidence that is readily available and confirms initial states.

Bacon's method therefore needs to be adapted for use in the non-experimental situations typically faced by management decision-makers and with the aforementioned tendencies in mind. To this end we now propose the following "Baconian algorithm" for decision-makers engaged in collecting evidence and generating hypotheses about how the future will unfold.

#### *The Baconian algorithm*

Take the familiar situation in which a decision-maker is deciding whether to introduce a new product, knows that the success of doing so will depend on the future state of the world, but does not know what this state will be. She accordingly proceeds by constructing and evaluating hypothetical states, each of which corresponds to a particular combination of influences she believes may be in play. By an influence we mean simply any event or state of affairs that she believes would contribute, causally or by forming part of it, to the realisation of any state of the world she is contemplating. For instance, influences likely to be relevant to whether or not to introduce the product might include events such as competitor responses and regulatory changes, and states of affairs such as the prevailing state of technology and existing market demographics.

The algorithm we propose provides a means for elaborating the state space that encourages the decision-maker to "think outside the box" and potentially uncover what were formerly unknown

unknowns. We assume a sequential learning process that begins once the decision-maker has already individuated one or more possible states on the basis of her prior knowledge about possible influences. We also assume that at every stage of the process she is able to order those states in terms of *inductive support*, that is, on a qualitative basis in terms of the balance of the evidence for and against a state being realised (Keynes, 1921).<sup>5</sup>

Let  $\Omega$  represent the set of all possible mutually exclusive states of the world relevant to the decision at the beginning of the process, with  $h_j$  as a generic element. Assume that the decision maker has the ability to order the states in terms of how favourable they are to the project, running from the least to the most favourable. Note that this ordering is not the same as the ordering of states in terms of inductive support.

Suppose the decision-maker does not know all of the members of  $\Omega$ , that is, that some of them are unknown unknowns, and that she is elaborating her state space by collecting evidence and generating and evaluating hypotheses about how the future will unfold. Let  $\{E_i\}_{i \leq n}$  be the evidence she has collected up to and including the ( $n$ )th stage of investigation, and  $H_{(n)} = (h_1, h_2, \dots, h_j, \dots, h_m)$  be the set of possible future states already included in her personal state space (the known unknowns at that point). Finally, let  $h_{j^*(n)}$  be the “base point” state that, on the basis of the body of evidence accumulated up to the ( $n$ )th stage of investigation, enjoys inductive support at least as high as that of any other. This state is used as the point of departure or base point in generating and testing alternative states of the world. The algorithm then proceeds in alternate stages from the following two hypotheticals:

- (1) The state that will be realised *ex post* lies to the left of the base point state on the favourability scale.
- (2) The state that will be realised *ex post* lies to the right of the base point state on the favourability scale.

It is arbitrary whether the procedure begins with hypothetical (1) or hypothetical (2). Suppose our decision-maker begins with hypothetical (1), in which case she is directed to do two things. The first is to imagine a possible influence, however unlikely, consistent with but not supported by her current body of evidence  $\{E_i\}_{i \leq n}$  and that, if it were in play, would give rise to a new state  $h_{m+1}$  that lies as far as possible from the base point state  $h_{j^*(n)}$  on the negative side of the “favourability” scale (the *variative phase*). The second is to look for additional evidence  $E_{i+1}$  that, if found, would grant  $h_{m+1}$  inductive support at least as high as  $h_{j^*(n)}$  (the *eliminative phase*).

The new evidence acquired might lead to the inclusion or rejection of the new state, the elimination of the base point state and other members of the original state space, and suggest entirely new states. The decision-maker is accordingly directed to re-define her state space  $H_{(n+1)}$  in the light of  $\{E_i\}_{i \leq n+1}$ , rank the states in terms of inductive support, and individuate the new base point state  $h_{j^*(n+1)}$ . So long as the additional evidence acquired is insufficient to make  $h_{m+1}$  the new base point state, then the process is

repeated as before. First, the decision-maker is required to imagine another influence, however unlikely, consistent with but not supported by her current body of evidence  $\{E_i\}_{i \leq n+1}$  and that, if it were in play, would give rise to a state  $h_{m+2}$  that lies as far as possible from the new base point state  $h_{j^*(n+1)}$  on the negative side of the “favourability” scale. Second, she is required to look for evidence  $E_{i+2}$  that, if found, would grant  $h_{m+2}$  inductive support at least as high as  $h_{j^*(n+1)}$ . Once done, she is directed to re-define her state space  $H_{(n+2)}$  in the light of  $\{E_i\}_{i \leq n+2}$ , rank the states in terms of inductive support, and individuate the new base point state  $h_{j^*(n+2)}$ .

Again, so long as the additional evidence acquired is insufficient to make  $h_{m+2}$  the new base point state, then  $h_{m+2}$  is included in the state space or discarded depending on the now expanded body of evidence  $\{E_i\}_{i \leq n+2}$ , and the process is repeated as before.

If at any stage of the process the decision-maker runs out of ideas and is no longer possible to perform the *variative phase*, she is directed to consider the least attractive state already included but not yet directly tested in her state space and look for additional evidence that, if found, would grant this state an inductive support at least as high as the base point state. Once done, she is directed to re-define her state space in the light of the additional evidence, rank the states in terms of inductive support, and individuate the new base point state. So long as the additional evidence acquired is insufficient to make this state the new base point state, then she is directed to restart the process.

The process continues until, at some stage of the process, say the ( $n+k$ )th stage, and on the basis of the accumulated evidence  $\{E_i\}_{i \leq n+k}$ , one of the following points is reached:

- either the newly postulated state or an already included state becomes the “new” base-point state  $h_{j^*(n+k)}$  to be tested. Note that this may happen at the first attempt, that is where  $k = 1$ , and is consistent both with  $h_{j^*(n+k-1)}$  being knocked out or retained; or
- it is no longer possible to perform the *variative phase* and all the states already included in the state space that are less attractive than the base point state have been directly tested.

In both cases, the algorithm then shifts to hypothetical (2), where attention turns to the positive side of the “favourability” scale. In this case the decision-maker is required to imagine possible influences, however unlikely, consistent with but not supported by her current body of evidence and that, if they were in play, would give rise to states that lie as far as possible from the base point state on the positive side of the “favourability” scale, and to look for additional evidence that, if found, would grant those alternative states inductive support at least as high as the base point state. Since the procedure is perfectly symmetrical with the one outlined for hypothetical (1), we will refrain from spelling it out again.

The process alternates between hypotheticals (1) and (2) until the decision-maker feels she is unable to get any further or decides to stop the process of generating previously unconsidered states, and all alternative states already included in the state space have been directly tested. The greater the number of eliminative tests performed, the greater the weight of evidence in favour of the remaining states, and the greater the confidence that the imagined future states provide appropriate guides to action.<sup>6</sup>

That completes our brief formalisation of what might be called a Baconian algorithm for non-experimental management situations. The key feature of the procedure we have described is that, by (i)

<sup>5</sup> We are not assuming any specific measure of the level of inductive support here, only that decision-makers are able to make intuitive qualitative judgments of this kind and that these are sufficiently finely-grained to allow them to rank states in terms of inductive support. Whatever the measure chosen, however, it would not conform to the axioms of probability calculus. This is because it would need to capture the idea that, on the Baconian algorithm, it is always possible to introduce additional states at each stage of the learning process (and where the non-inclusion of a state in the state space at some point in time does not mean that its probability of occurrence was zero but simply that, at that point, there was no evidence to support its inclusion). There are various measures that might satisfy this requirement by relying on one or another technical feature including Cohen's (1977) notion of Baconian probabilities, and Shafer's (1976) beliefs functions. See also Rottenstreich and Tversky (1997).

<sup>6</sup> Following Keynes (1921), the weight of evidence represents a measure of the absolute amount of evidence (the sum of the favourable and unfavourable evidence) in support of a hypothesis.

requiring the decision-maker to consider new negative/positive influences consistent with but not supported by her current body of evidence and that suggest states that are as distant as possible from the base point state at each stage, and then (ii) to endeavour to make these new states the new base point state (and not merely to include them in the state space) by collecting evidence of sufficient quality and quantity, her chances of uncovering something that was hitherto an unknown unknown, and which has the potential to eliminate the base point state and suggests new states, are enhanced. That is to say, it is by encouraging the decision-maker to expand her horizons by “thinking outside the box” to arrive at possible outlier influences consistent with but not supported by her current body of evidence and then requiring her to find evidence in support of those outliers, that the algorithm increases her chances of transforming unknown into known unknowns.

While what we have proposed differs from our own presentation of Bacon’s original method, it remains Baconian in spirit in two important ways. First, the decision-maker is encouraged constantly to generate alternative states of the world via the identification of influences that she had not considered before (variation). Second, where the decision-maker succeeds in finding evidence that supports the inclusion of new influences, the new states she generates will often throw doubt on or even disconfirm states that she was considering previously (elimination). The Baconian algorithm thus preserves Bacon’s idea of succession of “crucial experiments”, albeit in a non-experimental situation.

### Benefits

Applying the Baconian algorithm is relatively straightforward and offers immediate benefits by:

1. potentially reducing exposure to Black Swans by bringing to light states of the world that might not have been uncovered otherwise;
2. increasing the chances of discovering evidence that bears significantly on whether the states of the world already under consideration should be retained in the state space;
3. counteracting the confirmation bias, peoples’ tendency to favour evidence that confirms their preconceptions (Nickerson, 1998); and
4. counteracting various other cognitive biases.

To show how the Baconian algorithm works and the aforementioned benefits accrue we will run through a hypothetical example.

**Project:** Kate, a freshly minted MBA, has just started her first job at a prestigious Italian coffee retailer, and is tasked with scoping the possible outcomes of opening a chain of coffee shops in key centres in the Middle East.

### Stage 1

Kate conducts a risk/opportunity assessment and, on the basis of her current evidence  $E_1$ , arrives at:

**Influences:** {consumer demand, regulatory environment, level of competition}.

**State Space  $H_{(1)}$ :** three possible states of the world  $h_1 = \{\text{unfavourable}\}$ ,  $h_2 = \{\text{moderate}\}$ ,  $h_3 = \{\text{favourable}\}$ ;  $h_2$  the most likely (base point) state (see Fig. 1).

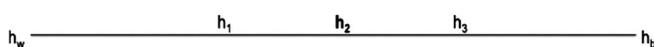


Fig. 1. State space at stage 1 based on  $E_1$ .

### Stage 2

Kate elects to test her views by applying the Baconian algorithm and, as she is keen to guard against her initial assessments having been too optimistic, begins with hypothetical (1). She comes up with:

**New negative influence consistent with but not supported by  $E_1$ :** {Possible adverse shock to coffee supply over medium term}.

**State under consideration:** new state  $h_4 = \{\text{highly unfavourable}\}$ .

**Search for new evidence  $E_2$ :** reports from public and private institutions on factors affecting coffee supply on world markets.

**Result of search:** discovers recent research that warns of Indigenious Arabica Coffee, one of the two main varieties of commercial coffee, becoming extinct due to near term effects of global warming (Davis, Gole, Baena, & Moat, 2012). On the basis of  $E_1$  and  $E_2$ , all original states are rejected as too optimistic and three new states are included.

**State Space  $H_{(2)}$ :** three new states  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ ;  $h_5$  the new most likely state (see Fig. 2).

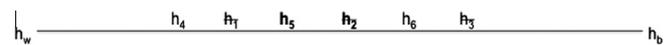


Fig. 2. State space at stage 2 based on  $E_1$  &  $E_2$ .

### Stage 3

Kate continues to attempt to imagine negative influences as directed by the algorithm and comes up with:

**New negative influence consistent with but not supported by  $E_1$  &  $E_2$ :** {Possible economic sanctions that would prevent all trading}.

**State under consideration:** new state  $h_7 = \{\text{catastrophic}\}$ .

**Search for new evidence  $E_3$ :** public debates about the possibility of a new era of protectionism, reports from public and private institutions on the latest introduction of international trade tariffs and non-tariff barriers to trade around the world, changing importance of anti-globalisation movements in the Middle East, etc.

**Result of search:** finds nothing that specifically suggests that sanctions are imminent. On the basis of  $E_1$  &  $E_2$  &  $E_3$ ,  $h_7$  is rejected and all the existing states are retained.

**State Space  $H_{(3)}$ :** same three states  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ , and  $h_5$  is still the most likely state (see Fig. 3).



Fig. 3. State space at stage 3 based on  $E_1$  &  $E_2$  &  $E_3$ .

### Stage 4

The evidence collected in Step 3 leads Kate to realise that sanctions might come from different directions and, specifically, to imagine that there may be ways in which resistance to Italian goods might develop in her target market.

**New negative influence consistent with but not supported by  $E_1$  &  $E_2$  &  $E_3$ :** {Possible resistance to Italian goods}.

**State under consideration:** new state  $h_8 = \{\text{catastrophic}\}$ .

**Search for new evidence  $E_4$ :** recent trends in Italian companies' exports, past cases of consumer boycotts of Italian products around the world, Italy's relationship with the Middle East, past international diplomatic incidents affecting the business of European companies in the Middle East, etc.

**Result of search:** discovers that, a few years back, the publication of a Danish newspaper of a series of caricatures depicting the Prophet Muhammad as a terrorist (also re-published by the Italian newspaper *La Stampa*) led to a consumer and retailer boycott that drastically affected the business of dairy company Arla Foods, Denmark's biggest exporter to the Middle East (Jensen, 2008). On the basis of  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$ ,  $h_8$  is now included and all the existing states are retained.

**State Space  $H_{(4)}$ :** four states  $h_8 = \{\text{catastrophic}_2\}$ ,  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ ;  $h_5$  still the most likely state (see Fig. 4).



Fig. 4. State space at stage 4 based on  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$ .

#### Stage 5

Kate tries but fails to imagine with another negative influence consistent with but not supported by  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  and, following the Baconian algorithm, she shifts attention to testing the least attractive of the untested states already included in her state space. As  $h_8$  has just been tested, she tests  $h_4$ .

**State under consideration:** already included state  $h_4 = \{\text{highly unfavourable}\}$ .

**Search for new evidence  $E_5$ :** further evidence that supports  $h_4$ .

**Result of search:** no further evidence specifically supporting  $h_4$  can be found. On the basis of  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$ , all existing states are retained.

**State Space  $H_{(5)}$ :** four states  $h_8 = \{\text{catastrophic}_2\}$ ,  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ ;  $h_5$  still the most likely state (see Fig. 5).



Fig. 5. State space at stage 5 based on  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$ .

#### Stage 6

Kate tries but fails to generate an additional influence consistent with but not supported by  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$ , and as all of the states already included in the state space that are less attractive than  $h_5$  have been tested, she moves to hypothetical (2). However, she fails to imagine a positive influence consistent with but not supported by  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$  and therefore, as directed by the algorithm, shifts her attention to testing the most attractive state already included in the state space.

**State under consideration:** already included state  $h_6 = \{\text{favourable downgraded}\}$ .

**Search for new evidence  $E_6$ :** further evidence that supports  $h_6$ .

**Result of search:** no extra evidence specifically supporting  $h_6$  can be found. On the basis of  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$  &  $E_6$ , all existing states are retained.

**State Space  $H_{(6)}$ :** four states  $h_8 = \{\text{catastrophic}_2\}$ ,  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ ;  $h_5$  still the most likely state (see Fig. 6).



Fig. 6. State space at stage 6 based on  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$  &  $E_6$ .

As no new positive influences can be generated and all the states already included in the state space that are more attractive than  $h_5$  have been tested, Kate moves back to hypothetical (1). However, as no additional (negative or positive) influences consistent with but not supported by the existing body of evidence can be generated and all alternative states already included in the state space have been tested, she stops the process. On the basis of  $E_1$  &  $E_2$  &  $E_3$  &  $E_4$  &  $E_5$  &  $E_6$ , Kate is confident that the project will face one of the following states of the world:

**Final state space  $H_{(6)}$ :**  $h_8 = \{\text{catastrophic}_2\}$ ,  $h_4 = \{\text{highly unfavourable}\}$ ,  $h_5 = \{\text{moderate downgraded}\}$ ,  $h_6 = \{\text{favourable downgraded}\}$ ;  $h_5$  the most likely state.

The four benefits of the Baconian algorithm noted above will be immediately apparent. First, in requiring the decision-maker to adopt a variative strategy by considering new influences that are consistent with but not supported by the current body of evidence, the algorithm induces the formulation of states of the world that might not have been surfaced otherwise. Further, in directing her to think of influences that give rise to alternative states that are as distant as possible from the base-point state at each stage, the common tendency to come up with states that are overly similar is likely to be counteracted. Finally, as shown by the example, by trying to collect enough evidence to make an alternative state the new base point one, the decision-maker is induced to imagine additional influences. The higher the number of alternative states considered ex ante, the higher the number of unknown unknowns uncovered, the higher the possibility of reducing exposure to Black Swans.

The second benefit of the Baconian algorithm demonstrated by our example is that it promotes the constant acquisition of evidence that has the capacity to disconfirm or even eliminate states already included in the state space. This effect is a by-product of requiring the decision-maker to attempt to find sufficient confirming evidence to convert the newly generated "outlier" state into the new base point state at each any stage of the cycling process. While the evidence will often not be sufficient to achieve this conversion, it will come from places that the decision-maker will likely not have looked before and which might well lead to the elimination of prior states. This is a crucial part of the story, since there is often a premium on eliminating irrelevant states as early on in the game as possible.

The third benefit of the Baconian algorithm is that it counteracts the confirmation bias in both the search for and evaluation of states. This effect is a consequence of the algorithm inducing the decision-maker to come with up alternative states of the world that are as distant as possible from the base point state at any stage of the cycling process, actively search for enough confirming evidence to make those alternative states the new base point ones, and thereby increasing the prospect of disconfirming initial states.

Finally, the algorithm provides a means of ameliorating many other cognitive biases that we alluded to in the review section (see Heath et al., 1998). In particular, in relation to the generation of states, it induces the decision-maker to look for states beyond those that might merely make her look good, to continue searching

for states even after finding one that appears plausible, and to generate alternative states when she might not have done so otherwise. In relation to the evaluation of states, it induces the decision-maker to collect and consider larger samples of information than she might have otherwise, and to look for new information when she might otherwise have restricted herself to only the most readily available information. Further, on that basis that cognitive repair strategies such as ‘consider the opposite’ or ‘consider an alternative’ have been effective in these cases, we suggest that the more elaborate procedure of the Baconian algorithm would also help mitigate judgmental errors that we have not mentioned so far, including anchoring, overconfidence and hindsight bias (Arkes, 1991; Fischhoff, 1982; Hoch, 1985; Koriat, Lichtenstein, & Fischhoff, 1980; Larrick, 2004; Mussweiler, Strack, & Pfeiffer, 2000; Russo & Schoemaker, 1992; Slovic & Fischhoff, 1977).

## Discussion

Here we consider three themes that have come up in discussion of the ideas we are advocating.

### *Bayes and Bacon compared: complements or substitutes?*

Getting to grips with the relationship between Bayesianism and Baconianism is difficult, not only because the philosophical literature on this subject is far from settled, but also because they are not entirely co-extensive in what they designed to do. A major difference, of course, is that whereas Bayesian inductive inference is exclusively about hypothesis evaluation, Baconianism extends to hypothesis discovery as well as hypothesis evaluation. This difference opens up the possibility of a complementary relationship between elements of the two philosophies. In a management decision-making context, for example, there is nothing to prevent something like the Baconian algorithm being used at the information-acquisition stage when the state space is being constructed and then, once the state space has been determined, decisions being made in accordance with the rules of Bayesian decision theory. There is no conflict here, at least at a general level. The Baconian algorithm is after all not a decision theory *per se* and is about the nuts and bolts of establishing hypothetical eventualities that Bayesian decision theory does not address.

However, in the context of hypotheses testing specifically, Bayesianism and Baconianism are customarily regarded as substitutes. Here it is useful to consider some parts of the philosophical discussion of the relationship between Bayesianism and the method of eliminative induction that are particularly relevant in a management context. Take the view reflected in some parts of the literature that, while Bayesian inductive inference proceeds in a more roundabout way, it is itself a (probabilistic) form of induction by elimination (e.g. Hawthorne, 1993, p. 99). According to the Bayesian account, the decision-maker assigns prior probabilities to each of the mutually exclusive alternatives that comprise the possibility space and then revises these in the light of new evidence using Bayes’ rule. Whenever the probability of one alternative increases, there is a corresponding decrease in respect of one or more of the other alternatives, and whenever a possible alternative is eliminated its probability mass is redistributed over the remaining alternatives. If the learning process eliminates all but one of the alternatives, then the posterior probability of the survivor is 1 (Earman, 1992; Hawthorne, 1993). In all other cases, in the presence of evidence that disconfirms/confirms hypotheses only probabilistically, false alternative hypotheses will be *highly refuted* (their posterior probabilities will be arbitrarily close to zero) and the true hypothesis will be *highly confirmed* (its posterior probability

will be arbitrarily close to unity) (Hawthorne, 1993; Savage, 1954). On this view, the Bayesian apparatus might be seen as a “tally device” for eliminative induction, one that “keeps a numerical tally of how hypotheses are faring in the eliminative process” (Hawthorne, 1993, p. 101).

The question, then, is how this Bayesian eliminative process compares with a more direct eliminative strategy. The philosophical literature is divided on this point, and turns on whether the convergence theorems that underpin the eliminative aspect of Bayesian inference apply in real-world scientific inference (see the debate between Earman (1992), Hawthorne (1993) and Kitcher (1993)).

Since Bayesianism requires no more of initial beliefs than that they conform to the probability calculus, it is possible that a rational decision-maker might assign extremely low prior probabilities to alternatives that, intuitively, are quite plausible. This is especially so when there are many possibilities on the table when priors are assigned. The idea is that these initial probability assignments will be “washed out” during the learning process and that, no matter how small the prior probability of the true alternative, given enough evidence of sufficient quality, its probability will converge to unity (Earman, 1992, p. 141).<sup>7</sup>

Supporters of eliminative Bayesianism argue that the convergence theorems that underpin the “washing out” mechanism are robust at a theoretical level and can also support inductive practice (Hawthorne, 1993). Against this, critics of Bayesianism argue that the conditions that have to be satisfied for these theorems to apply in many important cases of inductive inference are seldom met (Earman, 1992; Hesse, 1975), are overly permissive “of bizarre assignments of prior probabilities whose effects cannot be overcome sufficiently quickly” (Kitcher, 1993, p. 293) and that “convergence might take almost forever” (Hawthorne, 1993, p. 100). In these cases, a more direct eliminative strategy may allow someone to reject alternatives that his Bayesian counterpart, not yet free of the residue of his initial probability assignments, is still forced to countenance.

Hence, with respect to management situations, which often demand rapid decisions, adhering strictly to the rules of Bayesian updating in the process of evaluating hypotheses may be difficult. Yet it remains possible that, given its emphasis on finding eliminative evidence, using the Baconian method might lead to the discarding of a hypothesis that later resurfaces in the process (Hacking, 1983, p. 251) and proves to be the true one. In this respect, and in the short run, Bacon’s method might therefore be more vulnerable to error than the Bayesian one. As there might be a trade off between the speed of the testing process and its accuracy, where and when to use a Bayesian or a Baconian method in the process of hypothesis testing remains an empirical question.<sup>8</sup>

However, all of this ignores what is probably the most important count against the Bayesian form of eliminative induction in situations of management decision-making. As is well known, Bayesianism rests crucially on the assumption that decision-makers are always able to come up with numerically definite priors and conditional probabilities, and to update their priors in accordance with Bayes’ rule. The trouble is that there is overwhelming empirical evidence that they do not do so (Camerer, 1987; Grether, 1980; Grether, 1992; Holt & Smith, 2009; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1974, 1981), that they generally “do not trust, do not understand, or simply do not much use precise prob-

<sup>7</sup> Note that it is this process of convergence to unity that typically leads to the emergence of a consensus among Bayesian agents who hold different initial opinions (Earman, 1992, p. 141).

<sup>8</sup> We thank an anonymous referee for these points.

ability estimates” (March & Shapira, 1987, p. 1411). One response to this problem is to urge better training and to persuade people to try harder. But another is to go for an entirely different approach that does not depend on the assignment and updating of numerical probabilities. One of the advantages of Baconianism is that it does not assume that decision-makers rely on any particular form of quantitative reasoning in the process of testing hypotheses. It is therefore better suited to guiding and providing rigour to the qualitative kind of reasoning more often employed in testing and evaluating hypotheses in management decision-making.

### *The stopping problem*

Other things equal, the efficacy of the Baconian algorithm in elaborating the state space and reducing exposure to Black Swans depends on the number of eliminative tests performed. The question that then arises is whether it is possible to determine an optimal number of such tests. This question is an instance of the “stopping problem” of when to stop acquiring additional information before making a decision, and where there are costs as well as potential benefits of acquiring additional information.

The traditional “economic” approach to the stopping problem involves postulating a cost of search function, and then calculating the stopping point by equating the marginal cost of search with the expected marginal benefit (Stigler, 1961). However, in decision situations with unknown unknowns, the old conundrum of when to stop amassing information cannot be solved simply by assuming that the costs and benefits of additional information can be determined in advance of that information being received (March, 1994; Simon, 1978). When different instances contribute differently in terms of the amount and quality of information they provide, and can potentially lead to a radical change in the existing representation of the world, the benefits of another round of search cannot be quantified with any degree of precision.

This is not to deny the substantial statistical literature on stopping problems, many of which do have solutions (Ferguson, 2008). However, this literature is predominantly concerned with problems of choosing when to take a given action based on sequentially observed (independent and identically distributed) random variables whose joint probability distribution is assumed to be known. The problem of an incomplete state space therefore doesn't even arise in these cases. There are a few examples in which the assumption of a known probability distribution is relaxed, such as Rasmussen and Starr's (1979) “adaptive” stopping model for searching for new species, where the assumption of full knowledge of the proportions of the different species in the population required to derive the optimal stopping rule is replaced by a predictor that, at least under some circumstances, performs relatively well by comparison. However, rather than reflecting the kind of complexities and ambiguities typically faced in management, the setups that permit this kind of solution to the stopping problem tend to be highly artificial.

In particular, the kind of observations that typically inform management decisions cannot be likened to random drawings from an urn. Management decision-makers accordingly tend to be limited in what they can infer from samples of the available evidence (they cannot access the whole domain via the technique of random sampling), and where they often have to make decisions expeditiously by looking specifically for instances that have the power to discriminate between competing hypotheses (rather than making randomized observations). Further, in management contexts, the probability of observing a particular piece of evidence is often affected by the evidence already collected. That is to say, rather than the learning process being based on a large number of independent tests of the same hypothesis, it is one in which hypotheses may themselves be modified as a result of falsifying

evidence obtained, and new tests devised for new hypotheses in the light of evidence newly acquired.

In summary, as long as it is possible to think of new alternatives right up to the point at which the outcome of some decision-problem is revealed and all uncertainty is resolved, users of the Baconian algorithm have no option but to exercise their judgement on when to stop generating and eliminating hypothetical states. It then follows that the actual number of experiments performed by management decision-makers will invariably be suboptimal, and that it is impossible to establish the sign of this suboptimality until after the fact. This is an unavoidable feature of the kind of behavioural / satisficing approach we are proposing. But the Baconian approach is hardly unique in this regard. All inductive methods are afflicted by the stopping problem, and which can only be resolved under very special conditions.

### *Cognitive biases and the efficacy of the Baconian algorithm*

Critics might argue that the stopping problem is just one way in which the Baconian algorithm brings in subjective judgement, and that there are others—not least the judgements of inductive support on which it depends—that leave it vulnerable to cognitive biases of its own.

However, this criticism afflicts just about every approach aimed at improving the performance of decision-makers that involves subjective judgements, and obscures what we believe the real criterion should be: whether or not the approach in question does indeed have the capacity, on balance, to lead to an overall improvement on performance. The main problem our algorithm is supposed to solve is decision-makers' tendency to come up with overly narrow decision frames and, in the process, to improve their chances of uncovering unknown unknowns they would not have countenanced otherwise. Our contention is no more than that, given the same epistemic constraints, abilities, behavioural blinkers and so on, a decision-maker who follows the prescriptions of the Baconian algorithm is more likely to enjoy the benefits we claim above than one who does not.

For example, take the critic who asks why using the Baconian algorithm should render the decision-maker more inclined to run crucial experiments than she was before. Now it is clear that the algorithm cannot solve the problem of “lazy” decision-makers who break off testing too early during the learning process. However, while laziness may sometimes be a factor, the evidence suggests that a fundamental reason why decision-makers stop collecting evidence too early is that the information that comes to their minds at any point seems complete and coherent (Anderson & Spellman, 1995; Sanbonmatsu, Posavac, Kardes, & Mantel, 1998). The consequence is that they tend to get stuck in one way of thinking about a situation and cannot generate alternative views, thus reducing the feeling of uncertainty that might encourage the acquisition of further evidence (Larrick, 2009). Therefore, the real question for the Baconian algorithm is whether it has the capacity to help decision-makers overcome this kind of inertia by forcing them to broaden their decision frames.

In the same spirit, take the critic who asks why the Baconian algorithm should make the decision-maker more imaginative than she was before adopting it. Here again, it should be clear that we are not claiming that the algorithm alters the decision-maker's innate powers of imagination, and that the real question is whether it induces the decision-maker to use her given and perhaps quite limited powers of imagination in a more effective way.

We maintain that it is possible to give a positive answer to both questions. In both cases, the effectiveness of the Baconian algorithm derives from it being hypothesis-led, running from the generation of hypotheses to the collection of evidence and the evaluation of that evidence, rather than from the collection of

evidence to the generation of hypotheses and their evaluation. It is this feature that puts the decision-maker's imagination into the driving seat, and that, in constantly being urged to find evidence that supports alternative hypotheses, encourages her to broaden her decision frame.

A more general criticism might be that it is unrealistic to assume a decision-maker who is sufficiently self-aware to understand when to use the Baconian algorithm to “de-bias” herself. This is again an instance of a wider problem that afflicts all techniques that have been proposed to counteract behavioural biases, and one that can be located in the wider debate about the extent to which peoples' cognitive abilities can be affected via training (Larrick, 2004). It may well be that decision-makers cannot be relied upon to correct their own biases and may even resist being de-biased (Arkes, 2003; Kleinmuntz, 1990), and consequently, that the adoption of the algorithm might have to be encouraged or even imposed by a third party such as peers or superiors in an organisation.

### Managerial and organisational implications

The organisational and psychological literature offers various formal and informal techniques to help decision-makers avoid relying on unduly narrow views of the future, and some of those techniques are routinely adopted by organisations (Larrick, 2009). The question we address here is how the Baconian algorithm as a prescriptive tool relates to these contributions and what it can add to the “de-biasing” literature.

#### Formal techniques

Standard risk management techniques such as risk identification, prioritization, mitigation, prevention and contingent response (Chapman & Ward, 1997; Smith & Merritt, 2002) have been criticised for assuming a complete state space and so precluding unknown unknowns *ab initio* (McGrath & MacMillan, 1995; Pich et al., 2002). A number of formal techniques urging the active search for unknown unknowns and the preservation of flexibility have been proposed in response to these criticisms (McGrath & MacMillan, 2009; Miller & Lessard, 2000; Pich et al., 2002; Schoemaker, 2002; Schoemaker, 2004; Sommer & Loch, 2004; Sommer et al., 2009). Three broad recommendations emerge in these contributions. First, once management decision-makers have formed their initial views of how the future will unfold, they should acknowledge explicitly the existence of unknown unknowns by recognising that little is known and much assumed at the start of a new venture (McGrath & MacMillan, 1995). Second, in the presence of unknown unknowns, management decision-makers should not make the mistake of uncritically accepting their own views about how the future will unfold, and should engage in dynamic monitoring to identify when these no longer hold (Schoemaker, 2002). Third, in order to transform unknown unknowns into known unknowns as soon as possible, management decision-makers should engage in trial-and-error learning, that is “actively searching for new information and flexibly adjusting activities and targets to this new information, applying new and original problem solving... as new information becomes available” (Sommer et al., 2009, p. 119).

To the extent that they depart from the Bayesian decision-theoretic foundations of standard risk management methods in highly uncertain settings (Schoemaker, 2002, pp. 9–10; Sommer & Loch, 2004, p. 1337, fn. 6) and recommend that decision-makers constantly question their existing views of the future and actively search for unknown unknowns, these contributions share the premises and goals of the Baconian approach. However, they differ from the Baconian approach in one crucial respect, which is that they are essentially concerned with *disconfirming existing hypotheses* about

how the future will unfold. Thus we read that in the presence of unknown unknowns, genuine learning results from signals that are incompatible with decision-makers' prior predictions (Pich et al., 2002); that “intelligent failures” (Sitkin, 1992) are beneficial in exposing knowledge gaps in a timely way (Loch et al., 2006; McGrath, 2011; Sommer & Loch, 2004); that decision-makers should test systematically whether their initial assumptions still hold or contradictions have emerged (McGrath & MacMillan, 1995, 2009); and that they should look for “data that no longer fits” with their current views of the future (Schoemaker, 2002).

While these recommendations do potentially counteract the confirmation bias, they are unlikely to be much help in uncovering unknown unknowns and broadening decision frames. There are three reasons for this. First, the process of collecting new evidence is driven by hypotheses about how the future will unfold that are already included in the state space. The learning strategy is therefore about updating existing views of the future rather than generating new hypotheses and broadening decision frames. Second, it is difficult to apply a disconfirming strategy in practical choice situations. It is often impossible to distinguish between negative instances and false negatives in noisy environments, and in which case it may be rational for decision-makers to refrain from rejecting existing hypotheses about how the future will unfold in the light of ambiguous signals (Feduzi, Loch, & Runde, 2013). Finally, in the case in which disconfirming evidence can be found, decision-makers are primarily called to refine the existing state space by eliminating existing hypotheses about how the future will unfold rather than necessarily generating new ones. A disconfirming strategy is therefore more about narrowing than it is about broadening a decision framework.

From the perspective of the Baconian approach, the problem with the above techniques is ultimately that, like Bayesianism, they are about hypothesis testing rather than hypothesis generation. Now clearly it may happen that a disconfirming strategy is so effective that all—or a significant number of—hypotheses are eliminated, and that the decision-maker is then forced to generate new alternatives (Schoemaker, 2002, p. 166). But this situation is hardly likely to arise very often, and, rather than being a necessary element of the process, the generation of new hypotheses emerges only as a by-product of the activity of narrowing the existing decision frame. In contrast, on the approach we are proposing, the generation and testing of hypotheses proceed concurrently, and it is this that makes it possible to disconfirm and eliminate alternative hypotheses without running into the limitations of a direct disconfirming strategy. Rather than looking for disconfirming data that might directly lead to a revision of their existing hypotheses, decision-makers are urged to take a proactive approach by formulating alternative hypotheses and looking for evidence that, if found, has the potential to confirm them and potentially disconfirm the existing views. The Baconian approach thus turns the *reactive* logic reflected in the techniques discussed above (update existing beliefs in response to disconfirming evidence) on its head by proposing the *proactive* generation and investigation of possible alternatives.<sup>9</sup>

<sup>9</sup> It is worth mentioning that the Baconian approach is also consistent with Kuhn's (1962) view that working hypotheses be abandoned only when better alternatives are available (Wason & Johnson-Laird, 1972). This point is relevant in the present context because the disconfirming strategies mentioned above—and the strategy of “consider-the-opposite” that is sometimes recommended as a means of stimulating thinking about alternative possibilities (Hoch, 1985; Koriat et al., 1980) is another of these—suffer from the drawback that, if initial states are eliminated, the decision-maker is left with nothing to work with. The Baconian approach, which implicitly employs the strategy of ‘consider-an-alternative’ (Hirt & Markman, 1995), does not suffer this shortcoming.

### Informal techniques

There are also parallels between the Baconian approach and techniques of the informal, “cognitive repair” variety used by organisations to improve hypothesis generation and evaluation and to remedy unduly narrow frames (Heath et al., 1998; Larrick, 2009).

With respect to hypothesis generation, informal techniques that have been proposed include prompting decision-makers to consider alternative hypotheses for success so as to counteract the “self-serving” bias to favour hypotheses that reflect well on them (Walton, 1990); reducing decision-makers’ tendency to stop searching for alternative hypotheses too early by encouraging them to analyse a problem themselves, by asking them “why their favourite answer might be wrong”, by asking “Why?” five times before they have stopped generating hypotheses (see Imai (1986) on the “five whys” at Toyota) and by confronting them with experts or simply outsiders who can ask searching questions (Bazerman & Moore, 2009; Cusumano & Selby, 1995; Kahneman & Lovallo, 1993; Lovallo & Kahneman, 2003); by generating alternatives that truly differ from one another by grouping people who have different perspectives on the problem at hand (see the how the “kokai watch” technique at Bridgestone Tire relies on the use of non-experts, Walton (1990)); and by forcing different individuals to develop hypotheses independently of each other (in project design, Motorola composes cross-functional teams to ensure different perspectives on objectives, alternatives, scenarios, and then rotates team members when the task is completed so that the problem of shared views is avoided (Larrick, 2004)).

With respect to hypothesis evaluation, informal techniques that have been proposed include encouraging members to collect larger samples by using checklists for gathering and evaluating information (Larrick, 2009) and promoting slogans such as “lets talk with data” (Ishikawa, 1985, p. 200) or “Management by Fact” (Walton, 1990, p. 37); avoiding biased sampling by instituting processes to ensure that information is collected more systematically than it might be otherwise (e.g., Motorola’s Feature Prioritization Process (Heath et al., 1998)); prompting individuals to consider missing information, unusual events and counterfactuals (March, Sproull, & Tamuz, 1991); training individuals to avoid asking questions that are likely to elicit the answers they expect and thereby inhibiting the discovery of relevant information (e.g. asking open-ended questions rather than questions that can be answered with a simple yes or no (Heath et al., 1998; Mullins, 2007)); various variants of “red teaming” that involve setting up teams within organisations to challenge assumptions and explore alternative outcomes in order to reduce risks and increase opportunities (MOD, 2010); and by providing schemata that remind individuals what kind of information they need in such-and-such a situation (e.g., the Federal Reserve Bank of New York ensures that regulators give attention to a full set of relevant attributes by using CAMEL, a rating system that prompts them to consider information about Capital adequacy, Asset quality, Management, Earnings, and Liquidity (Heath et al., 1998; Larrick, 2004)).

Space precludes our going through each of these techniques individually. However, it will be evident that the Baconian approach encapsulates most of them. The advantages of the Baconian approach here are that it is at once grounded in a coherent inductive method, and that it offers a unified and implementable means of inducing decision-makers to become more systematic when considering alternative hypotheses, to vary their hypotheses and eliminate those that do not stand up to scrutiny, and then all in a way that encourages the constant acquisition and processing of new information.

### Future work

Our account suggests various avenues for future research. First, there is scope for improving our conceptions of the “states of the world” that decision-makers construct and contemplate when making decisions. One of the many contributions of Savage’s (1954) path-breaking book is the acts/states/consequences framework, which is routinely taken for granted by advocates of the canonical model. However, simple and incisive as this framework may be, most management decision-making cannot be squeezed into it. Even simple “what if” scenarios of the kind performed on spreadsheets typically involve interdependencies—for example acts affecting the probability of states or utilities of consequences depending on the state in which they occur—that the Savage framework is designed to exclude. There would therefore be considerable value in developing alternative conceptions of states of the world, building perhaps on the literature on scenario construction (Goodwin & Wright, 2010; Schoemaker, 1995, 2004; Wright & Goodwin, 2008). Research on these lines would likely also be useful in helping explore and further developing distinctions such as those between predictable and unpredictable surprises (Watkins & Bazerman, 2003), unforeseen and unforeseeable uncertainties (Sommer & Loch, 2004), gray and black swans (Taleb, 2007), and so on.

Second, there is scope for investigating non-standard “Baconian” measures of inductive support associated with the work of authors like Cohen (1970, 1977, 1989). Cohen’s work has been most influential in law and especially with respect to what it is sometimes called “the new evidence scholarship” in the law of evidence (Jackson, 1996; Lempert, 1988). He has also contributed to Behavioural Decision Theory, however, albeit mainly as a critic of the practice of taking the canonical model as the standard of rationality when defining “biases” (Cohen, 1977, 1979, 1980b, 1981; Einhorn & Hogarth, 1981; Heath et al., 1998; Over, 2004). The impact of his views in this regard has been limited so far, possibly due in part to the rejoinder by Kahneman and Tversky (1979b), but his constructive work on measures of inductive support stands in its own right, and, we suggest, is worth exploring more fully from the perspective of management decision-making.

Third, to the extent that the Baconian approach offers a rational procedure for mapping the space of possibilities, it might be applied to inform the processes of searching for and identifying superior (but cognitively distant) entrepreneurial opportunities as discussed by researchers working in the evolutionary and Carnegie traditions, and the entrepreneurship literature (Baron & Ensley, 2006; Denrell, Fang, & Winter, 2003; Gaglio, 2004; Gavetti, 2012). As Bayesianism cannot be used to support the process of discovery, there is room for a method that systematises the search for business opportunities rather than relying purely on chance and serendipity.

Finally, the proposals we have made call for empirical investigation. There are many possibilities here. For example, it could be tested experimentally whether the Baconian algorithm does indeed reduce the confirmation bias and facilitates the uncovering of unknown unknowns. This could be achieved by designing an experiment that tests whether subjects’ confirmation bias and tendency to focus on a single hypothesis about how the future will unfold are reduced after being given specific instructions to adopt the Baconian algorithm. Further, it would be useful to test whether the algorithm reduces judgmental errors such as overconfidence (Hoch, 1985; Koriat et al., 1980) and hindsight bias (Fischhoff, 1982; Slovic & Fischhoff, 1977), and, in view of its emphasis on evidential coverage, the extent to which it counteracts the problem of overreliance on overly narrow samples of evidence (Larrick, 2004).

## Conclusion

The terms “unknown unknown” and “Black Swan” have gained currency in the management literature, but tend to be used in different and sometimes conflicting ways. A key element of the present project was therefore to do some ground clearing, to distinguish between knowable and unknowable unknowns, and to clarify how unknown unknowns relate to the acts/states/consequences framework associated with the canonical model of individual decision-making as represented by Bayesian decision theory.

The problem of uncovering unknown unknowns is closely connected with the practicalities of state space construction, namely the activities of generating, evaluating and then accepting or rejecting candidate hypotheses about how the world might turn out. While there is already a fair amount work on these issues from a descriptive and prescriptive perspective, there is rather less from a normative perspective. We have argued that this situation can be traced to the continuing influence of the canonical model in its capacity as a normative standard, and where the Bayesian philosophy of inductive inference associated with it is largely silent on the mechanics of hypothesis generation, and therefore on the problems of state space construction and uncovering unknown unknowns. This situation has tended to sustain the received view that hypothesis generation is resistant to systemisation and logically and practically distinct from hypothesis evaluation, and that unknown unknowns are intractable almost by definition.

We have argued that Francis Bacon’s method of eliminative induction provides a firm basis for challenging the received view on each of these fronts. First, it offers a rational procedure for generating and exploring possibility spaces. Second, it shows how the processes of generating and evaluating hypotheses can play complementary and mutually supporting roles. Finally, it provides the resources for developing prescriptive approaches to state space construction such as the Baconian algorithm that we have proposed in this paper, and that may help make inroads on the problem of organisations being blindsided by Black Swans. At the very least, any event that would have remained an unknown unknown *ex ante* but for something like the Baconian algorithm, is one less Black Swan for an organisation to have to contend with *ex post*.

## Acknowledgments

For discussion and comments on earlier versions of this paper we are grateful to Eddie Anderson, Laure Cabantous, Massimiliano Corradini, Manuel Dries, Tony Endres, Philip Faulkner, Ben Hardy, Martin Kilduff, Paul Lewis, Christoph Loch, Adam Martin, John Roberts, Alessio Sancetta, Stefan Scholtes, Georg von Krogh, George Wright, two anonymous referees and Francesca Gino.

## References

- Anderson, J. R. (1995). *Cognitive psychology and its implications*. New York: W.H. Freeman & Co.
- Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, 102, 68–100.
- Arkes, H. R. (1991). The costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, 110, 486–498.
- Arkes, H. R. (2003). The nonuse of psychological research at two federal agencies. *Psychological Science*, 14, 1–6.
- Bacon, F. (1620). *Novum Organum* (English translation). In R. L. Ellis, & J. Spedding (Eds.), *The philosophical works of Francis Bacon* (pp. 212–387). London: Routledge (1905).
- Baron, J. (1985). *Rationality and intelligence*. New York: Cambridge University Press.
- Baron, J. (2004). Normative models of judgment and decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 19–36). London: Blackwell.
- Baron, J. (2012). The point of normative models in judgment and decision making. *Frontiers in Psychology*, 3, 1–3.
- Baron, R. A., & Ensley, M. (2006). Opportunity recognition as the detection of meaningful patterns: Evidence from comparisons of novice and experienced entrepreneurs. *Management Science*, 52, 1331–1344.
- Bazerman, M. H., & Moore, D. A. (2009). *Judgment in managerial decision making*. Hoboken, NJ: Wiley.
- Bell, D. E., Raiffa, H., & Tversky, A. (1988). Descriptive, normative, and prescriptive interactions in decision making. In D. E. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: descriptive, normative, and prescriptive interactions* (pp. 9–30). New York: Cambridge University Press.
- Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.
- Buchanan, B. G. (1985). Steps toward mechanizing discovery. In K. F. Schaffner (Ed.), *Logic of discovery and diagnosis in medicine* (pp. 94–114). Berkeley and Los Angeles: University of California Press.
- Camerer, C. (1987). Do biases in probability judgment matter in markets? Experimental evidence. *American Economic Review*, 77, 981–997.
- Camerer, C., Loewenstein, G., & Rabin, M. (2004). *Advances in behavioral economics*. Princeton, NJ: Princeton University Press.
- Chapman, C., & Ward, S. (1997). *Project risk management*. Chichester: John Wiley & Sons.
- Cohen, J. L. (1970). *The implications of induction*. London: Methuen.
- Cohen, J. L. (1977). *The probable and the provable*. Oxford: Clarendon Press.
- Cohen, L. J. (1979). On the psychology of prediction: Whose is the fallacy? *Cognition*, 7, 385–407.
- Cohen, L. J. (1980a). Some historical remarks on the Baconian conception of probability. *Journal of the History of Ideas*, 41(2), 219–231.
- Cohen, L. J. (1980b). Whose is the fallacy? A rejoinder to Daniel Kahneman and Amos Tversky. *Cognition*, 8, 89–92.
- Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317–331.
- Cohen, J. L. (1989). *An introduction to the philosophy of induction and probability*. Oxford: Clarendon Press.
- Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17, 1–25.
- Cunha, M. P., Clegg, S. R., & Kamoche, K. (2006). Surprises in management and organization: Concept, sources and a typology. *British Journal of Management*, 17, 317–329.
- Cusumano, M. A., & Selby, R. W. (1995). *Microsoft secrets*. New York: Free Press.
- Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Davis, A. P., Gole, T. W., Baena, S., & Moat, J. (2012). The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): Predicting future trends and identifying priorities. *PLoS One*, 7(11). <http://dx.doi.org/10.1371/journal.pone.0047981>.
- de Finetti, B. (1937). Foresight: Its logical laws, its subjective sources. In H. E. Kyburg & H. E. Smokler (Eds.), *Studies in subjective probability* (pp. 93–158). New York: Wiley (1964).
- Denrell, J., Fang, C., & Winter, S. (2003). The economics of strategic opportunity. *Strategic Management Journal*, 24, 977–990.
- Doyle, C. (1890). *The sign of the four*. Oxford: Oxford University Press (1993).
- Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge: MIT Press.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51, 380–417.
- Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology*, 12, 473–498.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, 32, 53–88.
- Feduzi, A., Loch, C. H., & Runde, J. (2013). *Learning through Baconian Creative Destruction in projects with unknown unknowns*. Unpublished manuscript, University of Cambridge.
- Ferguson, T. S. (2008). *Optimal stopping and applications*. Los Angeles: UCLA.
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). New York: Cambridge University Press.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 552–564.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 330–344.
- Gaglio, C. M. (2004). The role of mental simulations and counterfactual thinking in the opportunity identification process. *Entrepreneurship Theory and Practice*, 28, 533–552.
- Gavetti, G. (2012). Toward a behavioral theory of strategy. *Organization Science*, 23, 267–285.
- Gavetti, G., & Levinthal, D. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly*, 45, 113–137.
- Gavetti, G., & Levinthal, D. (2001). Bringing cognition back in and moving forward”. *Journal of Management and Governance*, 5, 213–216.
- Gavetti, G., Levinthal, D., & Ocasio, W. (2007). Neo-Carnegie: The Carnegie School’s past, present, and reconstructing for the future. *Organization Science*, 18, 523–536.
- Gettys, C. F., Pliske, R. M., Manning, C., & Casey, J. T. (1987). An evaluation of human act generation performance. *Organizational Behavior and Human Decision Processes*, 39, 23–51.
- Gilboa, I., Postlewaite, A., & Schmeidler, D. (2012). Rationality of belief. *Synthese*, 187, 11–31.

- Gilboa, I., & Schmeidler, D. (1995). Case-based decision theory. *Quarterly Journal of Economics*, 110, 605–639.
- Gnepp, J., & Klayman, J. (1992). Recognition of uncertainty in emotional inferences: Reasoning about emotionally equivocal situations. *Developmental Psychology*, 28, 145–158.
- Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting and Social Change*, 77, 355–368.
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology*, 43, 89–99.
- Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95, 537–557.
- Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organization*, 17, 31–57.
- Hacking, I. (1983). *Representing and intervening. Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Harper, D. A., & Lewis, P. (2012). New perspectives on emergence in economics. *Journal of Economic Behavior and Organization*, 82, 329–337.
- Hawthorne, J. (1993). Bayesian induction is eliminative induction. *Philosophical Topics*, 21, 99–138.
- Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. *Research in Organizational Behavior*, 20, 1–37.
- Hesse, M. (1975). Bayesian methods and the initial probability of theories. In G. Maxwell & R. Anderson (Eds.), *Induction, probability and confirmation* (pp. 50–105). Minneapolis: University of Minnesota Press.
- Hirt, E., & Markman, K. (1995). Multiple explanation: A consider-an alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, 69, 1069–1086.
- Hoch, S. J. (1984). Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 649–662.
- Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 719–731.
- Hodgkinson, G. P., & Starbuck, W. H. (2008). Organizational decision making: Mapping terrains on different planets. In G. P. Hodgkinson & W. H. Starbuck (Eds.), *The Oxford handbook of organizational decision making* (pp. 1–29). Oxford: Oxford University Press.
- Holt, C. A., & Smith, A. M. (2009). An update on Bayesian updating. *Journal of Economic Behavior and Organization*, 69, 125–134.
- Imai, M. (1986). *Kaizen: The key to Japanese competitiveness success*. New York: McGraw-Hill.
- Ishikawa, K. (1985). *What is total quality control? The Japanese way*. Englewood Cliffs, NJ: Prentice Hall.
- Jackson, J. D. (1996). Analysing the new evidence scholarship: Towards a new conception of the law of evidence. *Oxford Journal of Legal Studies*, 16, 309–328.
- Jensen, H. R. (2008). The Mohammed cartoons controversy and the boycott of Danish products in the Middle East. *European Business Review*, 20, 275–289.
- Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39, 17–31.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kahneman, D., & Tversky, A. (1979a). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–292.
- Kahneman, D., & Tversky, A. (1979b). On the interpretation of intuitive probability: A reply to Jonathan Cohen. *Cognition*, 7, 409–411.
- Keller, L. R., & Ho, J. (1988). Decision problem structuring: Generating options. *IEEE Transactions on Systems, Man, and Cybernetics*, 18, 715–728.
- Keynes, J. M. (1921). A treatise on probability. (Reprinted) In *The collected writings of John Maynard Keynes* (Vol. VIII). London: Macmillan (1973).
- Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford: Oxford University Press.
- Klayman, J. (1995). Varieties of confirmation bias. *The Psychology of Learning and Motivation*, 32, 385–418.
- Kleinmuntz, B. (1990). Why we still use our heads instead of formulas: Toward an integrative approach. *Psychological Bulletin*, 107, 296–310.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning, and Memory*, 6, 107–118.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Lampel, J., Shamsie, J., & Shapira, Z. (2009). Experiencing the improbable: Rare events and organizational learning. *Organization Science*, 20, 835–845.
- Lampel, J., & Shapira, Z. (2001). Judgmental errors, interactive norms, and the difficulty of detecting strategic surprises. *Organization Science*, 12, 599–611.
- Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–337). Oxford: Blackwell Publishing Ltd.
- Larrick, R. P. (2009). Broaden the decision frame to make effective decisions. In E. A. Locke (Ed.), *Handbook of principles of organizational behaviour* (pp. 461–480). Chichester, UK: John Wiley & Sons.
- Lawson, T. (2012). Ontology and the study of social reality: emergence, organisation, community, power, social relations, corporations, artefacts and money. *Cambridge Journal of Economics*, 36, 345–385.
- Leeson, G. (1977). *A chat with Linus Pauling*. California living (San Francisco Sunday examiner and chronicle (17 July)).
- Lempert, R. (1988). The new evidence scholarship: Analyzing the process of proof. In P. Tillers & E. Green (Eds.), *Probability and inference in the law of evidence: The uses and limits of Bayesianism* (pp. 61–102). Dordrecht: Kluwer Academic Publishers.
- Levinthal, D. (1997). Adaptation on rugged landscapes. *Management Science*, 43, 934–950.
- Loch, C. H., De Meyer, A., & Pich, M. T. (2006). *Managing the unknown*. Hoboken, NJ: John Wiley & Sons.
- Lord, C., Lepper, M., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47, 1231–1243.
- Lovallo, D., & Kahneman, D. (2003). Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review*, 81, 56–63.
- March, J. G. (1991). How decisions happen in organizations. *Human-Computer Interaction*, 6, 95–117.
- March, J. G. (1994). *A primer on decision making: How decisions happen*. New York: Free Press.
- March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science*, 33, 1404–1418.
- March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.
- March, J. G., Sproull, L. S., & Tamuz, M. (1991). Learning from samples of one or fewer. *Organization Science*, 2, 1–13.
- McGrath, R. G. (2011). Failing by design. *Harvard Business Review*, 89, 77–83.
- McGrath, R. G., & MacMillan, I. C. (1995). Discovery driven planning. *Harvard Business Review*, (July–August), 44–54.
- McGrath, R. G., & MacMillan, I. C. (2009). *Discovery-driven growth. A breakthrough process to reduce risk and seize opportunity*. Boston: Harvard Business Press.
- Mill, J. S. (1843). A system of logic. In J. M. Robson (Ed.), *Collected works of John Stuart Mill* (Vols. 7 and 8). Toronto: University of Toronto Press (1974).
- Miller, C. C. (2008). Decisional comprehensiveness and firm performance: Toward a more complete understanding. *Journal of Behavioral Decision Making*, 21, 598–620.
- Miller, R., & Lessard, D. L. (2000). *The strategic management of large engineering projects*. Cambridge: MIT Press.
- MOD (2010). *A guide to red teaming*. <<http://www.mod.uk/NR/rdonlyres/B0558FA0-6AA7-4226-A24C-2B7F3CCA9A7B/0/RedTeamingGuiderevised12Feb10Webversion.pdf>>.
- Muller, B., & Riordan, C. A. (1988). Self-serving attribution for performance in naturalistic settings: A meta-analytic review. *Journal of Applied Social Psychology*, 18, 3–22.
- Mullins, J. W. (2007). Discovering unk-unks. *Sloan Management Review*, 48, 17–21.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26, 1142–1150.
- Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information relevance, working memory, and the consideration of alternatives. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 46, 759–778.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175–220.
- Norton, J. D. (1995). Eliminative induction as a method of discovery: How Einstein discovered general relativity. In J. Leplin (Ed.), *The creation of ideas in physics* (pp. 29–69). Dordrecht: Kluwer.
- Over, D. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 3–18). Oxford: Blackwell Publishing Ltd.
- Peirce, C. S. (1898). Reasoning and the logic of things. In K. L. Ketner (Ed.), *The Cambridge conferences lectures of 1898*. Cambridge, MA: Harvard University Press (1992).
- Pich, M. T., Loch, C. H., & De Meyer, A. (2002). On uncertainty, ambiguity and complexity in project management. *Management Science*, 48, 1008–1023.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347–352.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Ramsey, F. P. (1926). Truth and probability. In P. Gärdenfors & N.-E. Sahlén (Eds.), *Decision, probability and utility* (pp. 19–47). Cambridge: Cambridge University Press (1988).
- Rasmussen, S. L., & Starr, N. (1979). Optimal and adaptive stopping in the search for new species. *Journal of the American Statistical Association*, 74, 661–667.
- Reichenbach, H. (1951). *The rise of scientific philosophy*. Berkeley: University of California Press.
- Rerup, C. (2009). Attentional triangulation: Learning from unexpected rare crises. *Organization Science*, 20, 876–893.
- Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, 104, 406–415.
- Rumsfeld, D. (2002). *DoD news briefing – Secretary Rumsfeld and Gen. Myers*. <<http://www.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636>>.
- Runde, J. (2009). Dissecting the black swan. *Critical Review*, 21, 491–505.

- Russo, J. E., & Schoemaker, P. J. H. (1992). Managing overconfidence. *Sloan Management Review*, 33, 7–17.
- Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin and Review*, 5, 197–220.
- Savage, L. J. (1954). *The foundations of statistics*. London: John Wiley & Sons.
- Schoemaker, P. J. H. (1995). Scenario planning: A tool for strategic thinking. *Sloan Management Review*, 36, 25–40.
- Schoemaker, P. (2002). *Profiting from uncertainty. Strategies for succeeding no matter what the future brings*. New York: The Free Press.
- Schoemaker, P. J. H. (2004). Forecasting and scenario planning: The challenges of uncertainty and complexity. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 274–296). Oxford: Blackwell Publishing Ltd.
- Schum, D. (1994). *The evidential foundations of probabilistic reasoning*. New York: John Wiley & Sons.
- Shackle, G. L. S. (1979). *Imagination and the nature of choice*. Edinburgh: Edinburgh University Press.
- Shackle, G. L. S. (1983). The bounds of unknowledge. In J. Wiseman (Ed.), *Beyond positive economics* (pp. 28–37). London: MacMillan.
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.
- Shapira, Z. (2008). On the implications of behavioral decision theory for managerial decision making: Contributions and challenges. In G. P. Hodgkinson & W. H. Starbuck (Eds.), *The Oxford handbook of organizational decision making* (pp. 287–304). Oxford: Oxford University Press.
- Simon, H. A. (1947). *Administrative behavior*. New York: Macmillan.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99–118.
- Simon, H. A. (1978). Rational decision-making in business organizations. *Nobel Memorial Lecture*, 8(December).
- Simon, H. A. (1982). *Models of bounded rationality* (Vols. 1 and 2). Cambridge, MA: MIT Press.
- Sitkin, S. B. (1992). Learning through failure: The strategy of small losses. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 231–266). Greenwich: JAI Press.
- Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544–551.
- Smith, P. G., & Merritt, G. M. (2002). *Proactive risk management: Controlling uncertainty in product development*. New York: Productivity Press.
- Snowden, D. J., & Boone, M. (2007). A leader's framework for decision making. *Harvard Business Review*, 85, 69–76.
- Sommer, S. C., & Loch, C. H. (2004). Selectionism and learning in projects with complexity and unforeseeable uncertainty. *Management Science*, 50, 1334–1347.
- Sommer, S. C., Loch, C. H., & Dong, J. (2009). Managing complexity and unforeseeable uncertainty in startup companies: An empirical study. *Organization Science*, 20, 118–133.
- Starbuck, W. H. (2009). Cognitive reactions to rare events: Perceptions, uncertainty, and learning. *Organization Science*, 20, 925–937.
- Stigler, G. (1961). The economics of information. *Journal of Political Economy*, 69, 213–225.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. London: Penguin books.
- Taleb, N. N., Goldstein, D. G., & Spitznagel, M. W. (2009). The six mistakes executives make in risk management. *Harvard Business Review*, 87, 78–81.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
- von Frisch, K. (1950). *Bees, their vision, chemical sense and language*. Ithaca: Cornell University Press.
- Walton, M. (1990). *Deming management at work*. New York: Putnam Publishing.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, 12, 129–140.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Watkins, M. D., & Bazerman, M. H. (2003). Predictable surprises: The disasters you should have seen coming. *Harvard Business Review*, 81, 72–80.
- Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the unexpected. Resilient performance in an age of uncertainty* (2nd ed.). San Francisco: John Wiley & Sons.
- Wideman, R. M. (1992). *Project and program risk management: A guide to managing project risks and opportunities*. Newton Square, PA: Project Management Institute.
- Wright, G., & Goodwin, P. (2008). Structuring the decision process: An evaluation of methods. In G. P. Hodgkinson & W. H. Starbuck (Eds.), *The Oxford handbook of organizational decision making* (pp. 534–551). Oxford: Oxford University Press.
- Zuckerman, M., Knee, C. R., Hodgins, H. S., & Miyake, K. (1995). Hypothesis confirmation: The joint effect of positive test strategy and acquiescence response set. *Journal of Personality and Social Psychology*, 68, 52–60.