UNKNOWN UNKNOWNS AND THE TREATMENT OF FIRM-LEVEL ADAPTATION IN STRATEGIC MANAGEMENT RESEARCH

Timo Ehrig and Nicolai J Foss
Department of Strategy and Innovation, Copenhagen Business School,
Kilevej 14A, DK-2000 Frederiksberg, Denmark
Te.si@cbs.dk, njf.si@cbs.dk
UNKNOWN UNKNOWNS AND THE TREATMENT OF FIRM-LEVEL ADAPTATION IN STRATEGIC MANAGEMENT RESEARCH

Abstract

“Unknown unknowns,” that is, future contingencies that lack an ex ante description for some decision-makers for whom the contingency is relevant, are fundamental to strategy theory and practice. And yet, most strategy research is founded on the assumption that the future can be described in terms of “known unknowns,” that is, future contingencies that are known in principle (but whether and how they actually occur is unknown). We discuss the importance of unknown unknowns for strategy, focusing specifically on firm-level adaptation. We also discuss why prior literature has failed to address unknown unknowns, and outline key points that should be addressed by a program of research into the nature and role of unknown unknowns in strategy.
INTRODUCTION

While former World Bank President Jim Yong Kim stated that “[n]o one in the field of infectious disease or public health can say they are surprised about a pandemic” (Maxmen, 2020), there is no doubt that the COVID-19 pandemic took the business world very much by surprise. The pandemic was initially an “unknown unknown” in the now-famous terminology of former US Defence Secretary Donald Rumsfeld. The business world faced the kind of radical uncertainty highlighted by economists such as Frank Knight and George Shackle: Unforeseen events that, at least initially, escape the kind of categorization that makes fast action and adaptation possible. It simply wasn’t clear what rational adaptation to the pandemic entailed, at least initially. In this article, our central concern is to raise the issue of adaptation to unknown unknowns as a key problem in strategy and discuss why it has hitherto been relatively neglected, approach the issue in terms of top manager decision-making, and finally sketch a decision-making framework that may assist adaptation to unknown unknowns. For reasons of space, we do not discuss issues of organizational preparedness to deal with unknown unknowns. We define “unknown unknowns” as future contingencies that lack an ex ante description for at least some decision makers that are later affected by the contingencies. Because such contingencies have not been written down or thought or talked about, they are not present in the minds of the relevant decision-makers (i.e., who are “unaware” of them). Managers are frequently confronted with such contingencies. For instance, the entry of Airbnb confronted the hotel industry with an “unknown unknown”: An event that lacked ex ante description for most strategists in this industry. Such events often shake up industries, change the foundations of competitive advantage by redefining what are relevant strategic factors, create new barriers to imitation and entry, and so on. While previous research has not been entirely silent about how, for example, firm flexibility or team diversity may assist in dealing with such
unforeseen events (e.g., Cannella, Park, & Lee, 2008), we do not have a framework that allows us to understand how strategists cope with them in terms of decision-making. In other words, while unknown unknowns intimately relate to canonical issues of strategy the microfoundations of understanding adaptation to them are highly incomplete.

This is to a large extent a problem of a lack of analytical tools that can help build such microfoundations, that is, a lack of proper analytical “technology.” While we have very well-developed probabilistic tools for analyzing the decisions of individuals, the same is not the case for those decision situations that escape a probabilistic (e.g., Bayesian) treatment. This is reflected in our theorizing, and perhaps partially reflects the reliance of strategy on economics. According to Nobel laureate, Robert Lucas (1981: 224), “In cases of uncertainty, economic reasoning will be of no value.” While management scholars may be less nihilistic in this regard, attempts to deal with decision-making under situations of radical uncertainty, such as “judgment” (Foss & Klein, 2012) or “sense-making” (Weick, 1988, 1993; Gephardt, 1993), are to a considerable extent labels for our ignorance. For example, Foss and Klein (2012) characterize “judgment” as decision-making in situations in which there is no decision rule, and argue that there are no decision rules under radical uncertainty. However, they offer little insight into what judgment actually is. More generally, there is a sense in which we literally don’t know how decision makers cope with radical uncertainty, which has prompted a few attempts at getting to grips with such decision-making by means of qualitative research (e.g., Mintzberg et al., 1976; Klingebiel & Meyer, 2012). Our lack of ignorance here is clearly problematic, given the importance of radical uncertainty to real-world decision-making (as reflected, perhaps, in the success of e.g. Taleb, 2007).

In the following, we explore the role of unknown unknowns in strategic decision-making, particularly with respect to adaptation, that is, how firms react to and cope with the emergence of

---

1 We here use “technology” in the sense of March (2006) who coined the notion of “technologies of rationality” to refer to systematized, potentially-model based ways of making decisions in organizations.
novelty. The premise is that the strategy research community may have been committed for too long to the idea that its ultimate, fundamental purpose is to explain sustained rents in equilibrium (see also Collis, 2019). Strategy also includes the production and adaptation to novelty (Lippman & Rumelt, 2003), but this focus tends to be excluded if strategy’s core models are essentially timeless equilibrium models with heterogeneous, but perfectly adapted firms. In other words, the heterogeneity does not seem to include heterogeneity with respect to how well firms adapt to unknown unknowns. We suggest that making more room unknown unknowns in strategy may further our understanding of the consequences of heterogeneity.

Luckily, we are not in the situation that Lucas thought economists were in--that is, having to stay silent “in cases of uncertainty.” We argue that there are in fact insights in decision-making that potentially allow for a rigorous treatment of unknown unknowns and how decision-makers cope with them. As an example, we sketch cognitive science ideas on belief revision and explain how they allow us to deal with unknown unknowns in terms of firm-level adaptation. We discuss why management theory did not pick up these ideas earlier, end by offering a research program on unknown unknowns, and speculate on its possible implications.

FROM KNOWN KNOWNS TO UNKNOWN UNKNOWNS IN STRATEGY

What Does Strategy Research Seek to Explain?

Part of the reason strategy has had little room for unknown unknown has to do with what strategy seeks to explain, and how it explains this. It is often argued that strategy research ultimately seeks to explain sustained competitive advantage (e.g., Barney, 1991; Peteraf, 1993; Wiggins & Ruefli, 2002; Denrell, 2004). In this view, all corporate, business, functional etc. strategy research ultimately derives its importance to the extent that it improves the understanding of sustained competitive advantage. This approach has the advantage of unifying strategy research around a dependent variable, theoretical structures that organize relevant variables in a transparent
manner (e.g., the various lists of jointly necessary and sufficient conditions for sustained competitive advantage), and the insights and models that articulate this causal structure (e.g., price theory and industrial organization theory from economics, options theory from finance and so on). Still, it has been criticized for making the field concentrate on only the few exceptional outliers that earn substantial rents in equilibrium (Wiggins & Ruefli, 2002).

Our critical perspective, or at least reservation, is different: The view that strategy is first and foremost about explaining rents earned in equilibrium risks missing out on a lot of the real action. Thus, the closed world of the rents-in-equilibrium view makes it difficult to handle dynamics introduced by the innovative and entrepreneurial activities of firms (Schumpeter, 1911; Phelps, 2006). In fact, in such models the sources of rents are typically represented as initial draws from some distribution of productive efficiencies (as in Lippman & Rumelt, 1982). Adaptive, competitive actions are framed as reactions to exogenous changes that can be described in terms of a probability distribution function. Interestingly, these representations also characterize various ostensibly evolutionary representations of the market process that seek analyze this process in terms of a sequence of temporary equilibria and applies the technology of Markov chains to analyze the evolution of the analyzed system (e.g., Nelson & Winter, 1982). It is, we think, highly debatable whether such equilibrium and probabilistic tools are appropriate for the analysis of processes that are open-ended in the sense that they involve creativity and novelty, or, in Kauffman’s colorful metaphor, “the adjacent possible” (Kauffman, 2010). Such analytical technology makes the environment, in principle, knowable, although firms may have to engage in search (again modelled probabilistically) to, for example, find better production routines (Nelson & Winter, 1982).

In our view, strategy research should deal more broadly with actions aimed at creating and appropriating value in dynamic conditions (cf. also Lippman & Rumelt, 2003)—what firms do to temporarily create positions and advantages associated with creating and appropriating more value
than a good part of the relevant competition, how they fight for them, regenerate them, and eventually lose them. This directs the focus towards how the long-term survival of firms is contingent on their innovation and adaptation strategies—whether or not firms can successfully cope with environmental change.

This has, of course, been the key concern of the behavioral theory of the firm (Cyert & March, 1963; Gavetti, Levinthal, Greve, & Ocasio, 2012), but much research in this stream has blackboxed the issue of how relevant decision-makers interpret changes and what makes them pick some responses rather than other ones (e.g., Greve, 1998). An important approach to conceptualize, model, and analyze processes of adaptation is to assume a “landscape” or “phase space” on which firms adapt (Gavetti and Levinthal, 2000; Gavetti, Levinthal, Rivkin, 2005; Gavetti, 2012), typically based on the technology of NK modeling. In this approach, the changing environment is in principle knowable, but firms cannot “see” the entire landscape and need to rely on search strategies (“local” and/or “distant”) to find favorable positions in the changing environment. This approach has been criticized (e.g., Felin, Kauffman, Koppel, & Longo, 2014). In particular, picturing adaptation as a search on given landscapes may mischaracterize the emergence and management of novelty (ibid.). However, we argue that viewing adaptation exclusively as a process of “search over landscapes” has another fundamental shortcoming: Even if firms are not creating novelty, they may still be faced with unknown unknowns. As we will argue dealing with unknown unknowns is very different from a search on given landscapes (or, obviously, from making draws from a distribution of productive efficiencies).

**What Are Unknown Unknowns?**

What are unknown unknowns, and why is dealing with them not captured by, for example, the notion of adaptation as search on a given landscape—or, indeed any of the analytical technologies currently applied in strategy research? The issues here involve much philosophical
subtlety having to do with the ontological and epistemological status of unknown unknowns. This essay is not the place for an extensive discussion of this (neither do we have the philosophical competence for an in-depth discussion), so here we mainly seek to characterize unknown unknowns in an attempt to increase construct clarity.

As a starting point unknown unknowns are naturally characterized as events that lack an *ex ante* description. But, an obvious question is *to whom* the *ex ante* description is lacking. There may be cases of radical unknown unknowns in the sense that even relatively shortly before their emergence, very few individuals had anticipated them. Another obvious observation is that there is a time frame issue: The further the temporal distance between decision makers and future events, the more likely it is that some of these events will be unknown unknowns. This suggests that for the purposes of strategy research, we focus on those unknown unknowns that are so temporally close that they can make a meaningful difference to the actions and reactions of the focal firms.

Moreover, as our opening quotation suggests, there is a case for arguing that the Covid-19 was not entirely lacking an *ex ante* description, as parts of such a description (perhaps rather coarse-grained) may have been present in the world epidemiological community. In fact, there may be few unknown unknowns in the sense of future contingencies *entirely* lacking in any *ex ante* description (which again depends on the temporal perspective). This discussion motivates our above definition of unknown unknowns as future contingencies that lack an *ex ante* description for at least some decision-makers that are going to be affected by the contingency.

This definition is a stronger statement of relative ignorance than a statement about the lack of a probability for an event (i.e., the way “Knightian uncertainty” is sometimes interpreted; cf. e.g., LeRoy & Singell, 1987). It is not just that probabilities for an event are unavailable—rather, the event is not even anticipated or imagined and therefore not *ex ante* categorized, discussed, etc. by

---

2 The invention of the wheel and the 1989 fall of the Berlin wall may be examples.
decision-makers (in fact, to the extent that unknown unknowns are unique they may also be challenging to categorize *ex post*). This implies that Bayesian attempts (e.g., see Wiblin & Harris’ [2019] interview with Philip Tetlock) at reinterpretating Knightian uncertainty in terms of subjective probabilities relating to “unique” events (which may also include Knight’s own discussion, Knight, 1921) do not work, as the notion of formulating a “probability for the category of events that have not yet been imagined” is meaningless.³

Our definition is also a weaker definition than what may be called strong unawareness of a possibility, that is, something that a strategist is unable to conceive of *ex ante*. For instance, in the perspective of early 2020, Covid-19 was an unknown unknown, as it back then lacked a description. However, people (particularly epidemiologists) were not strongly unaware of the possibility of something like Covid-19; the event was in principle conceivable, and indeed the possibility of new viruses and pandemics was known in principle (cf. Gates’ 15. March 2015 warning in a Ted Talk; Gates, 2015). However, the Covid-19 pandemic was not featured in the state-spaces underlying policy-making up until the onset of the pandemic, as the virus and its consequences lacked description for this set of decision-makers. By contrast, climate change was already well described in early 2020.

**Dealing with Surprises**

To make these distinctions may seem to be mainly a philosophical exercise, but it actually is of considerable relevance to strategy theory and practice. If events are unknown unknowns in *ex ante* perspective and only later turn into known unknowns, that means that at some point in time, descriptions of future contingencies need to be revised—a process that British economist G.L.S. Shackle (1949, 1955) framed as one of dealing with surprise. Shackle (1949) built a sophisticated

---

³ To construct subjective probabilities is a demanding task, as a state-space needs to be specified, and numbers attached to events that fulfill the probability axioms (in particular, the probabilities of singular events need to sum up to one). If events have not yet been imagined, this implies that there is not even a state space on basis of which probabilities could be attributed (explicitly or implicitly).
decision-theoretic approach around the notion of surprise and degrees of surprise (see, e.g., Basili & Zappia, 2010). Unknown unknowns are events that involve a very high degree of surprise.⁴

Apart from a few Shackle-inspired economists—sometimes called “radical subjectivists” (e.g., Loasby, 1976; Lachmann, 1986; Littlechild, 1986)—very few economists followed Shackle’s emphasis on radical uncertainty, surprise, and the need for insight into how surprise is dealt with by decision-makers. There are arguably two reasons for this, one substantive and one more methodological.

First, allowing for radical uncertainty and attendant surprise was seen by many as destructive to economics (Coddington, 1982): If the economy can unexpectedly be upset at any time by the emergence of major unknown unknowns, economics seems void of predictive content.⁵

Second, unknown unknowns seem to play havoc with established decision theory. In particular, when decision-makers, including strategists, have to deal with unknown unknowns, Bayesian technologies cannot provide them with “the right way” to update their beliefs: There is nothing in standard probability theory that tells us how to update probability distributions when facing unknown unknowns. While probability theory provides means to understand how to derive a marginal distribution if a known, bigger, joint distribution is available (by conditioning probabilities on an event), the same does not apply when a description of a novel type of event is included into a

---

⁴ Other, related types of surprises are “tail events” (events with a very low ex ante probability), or, relatedly, events that were described but lacked salience for the decision-maker. However, the cognitive scientist Johnson-Laird argued that a human’s mental representations of an environment does typically only contain a number of described elements (Johnson-Laird, 2008). Thus, events that lack salience may, from psychological perspective, be events that lack an ex ante representation and thus description. We opt for our above definition of surprise, for the following reason: While one could argue that unknown unknowns are just low probability events that are excluded from the mental model of a decision-maker due to behavioral constraints, unknown unknowns like Airbnb are, in ex ante perspective, unknown new patterns to arrange social, economic and physical reality. In the same way, as a new animal could in principle be described as a genome sequence that could already be stored as a low probability event in a gigantic ex ante representation of the world that encompasses the entire biological evolution on planet earth (including evolution paths that were not chosen), we argue that this way of conceptualizing unknown unknowns is of little value for strategy research. Unknown unknowns are often undescribed because collectively, sufficient knowledge about them is yet missing.

⁵ Shackle (1972: 76) described this as a “kaleidic” view of the economy, “interspersing its moment or intervals of order, assurance and beauty with sudden disintegration and a cascade into a new pattern.”
state space and thus, no bigger joint distribution that could be updated is available in the first place. State spaces expand, but probability theory can only say something about updated distributions over state spaces that are bigger than the state spaces that were available to a decision maker \textit{ex ante} if a larger state space that already included the unforeseen contingency was somehow available \textit{ex ante}, too. Thus, the only way in which Bayesian technologies can be applied to model coping with surprises is to assume that some group of agents A is lacking an \textit{ex ante} description of a major contingency, but some other group of agents B already anticipated the contingency \textit{ex ante} and anticipated everything that group A anticipated. More generally, Bayesian modelling can only work if a complete, merged, state-space was somehow already available \textit{ex ante}.

More recently, however, economists seem to have warmed to the idea that somehow taking unforeseen events into account is important (e.g., Kreps, 1992, 1996; Grossman & Hart, 1986; Williamson, 1996; Zeckhauser, 2006). One reason is that such events do happen and often have a major impact on the economy (the 1997 Asian meltdown, the 9/11 attacks, the collapse of Lehman Brothers, etc.). Another one is more theoretically specific: Research suggests that the role of firms is to minimize the frictions associated with bargaining in the face of unanticipated contingencies and minimize losses from underinvestment because parties fear being held up by other parties that may use unanticipated contingencies as bargaining levers (Grossman & Hart, 1986; Williamson, 1996).\footnote{In fact, following Hayek (1945), Williamson (1996) holds even more broadly that economic organization should first and foremost be assessed against the extent to which it assists adaptation. We concur.}

In this context, economists seek to make room for the unanticipated by assuming that while unanticipated contingencies are indeed unanticipated, the consequences of something unanticipated happening can be foreseen, at least probabilistically. This allows for the use of technologies such as dynamic programming and for assumptions that agents are still capable of choosing efficient
governance structures (Kreps, 1996). The problem, of course, is that seems at least borderline inconsistent to assume that there are contingencies that agents cannot foresee—but they can anticipate the pay-off consequences of such unanticipated contingencies. Not only unknown unknowns but also their consequences are not foreseeable. Importantly, strategic uncertainty (Ehrig, Jost, Katsikopoulos, & Gigerenzer, 2019), that is, uncertainty about the reactions of others to such events (what is sometimes called “behavioral uncertainty”), prevents anticipating their consequences. For example, the reactions of the multiple stakeholder groups to events such as the Covid-19 pandemic or the emergence of the Airbnb business model were unknown, but obviously matter a great deal for firms’ pay-offs.

Our standard decision-making tools, such as maximization of expected payoffs, are not well-equipped to deal with unknown unknowns and the surprise they introduce. Dealing with these requires new technologies of strategic analysis, in theory and in practice. We argue that while strategy research has occasionally allowed for unknown unknowns, strategy lacks a systematic treatment because there is little awareness of the developments that may allow for such a treatment (Foss & Hallberg, 2014). Thus, while the literature on sensemaking and categorization provides important steps towards an analysis of unknown unknowns, we need to go beyond and study more broadly how strategists revise their representations of environments when they face unknown unknowns. To make this concrete, below we review the potential of one specific approach for dealing with unknown unknowns and the surprises they imply, namely the rather established cognitive science literature on belief revision literature.

**A TECHNOLOGY FOR HANDLING UNKNOWN unknowns:**

**BELIEF REVISION**

**Cognitively Dealing With Surprise**

Dealing with surprise is initially a cognitive operation, as it necessitates the creation of an
initial description for something that was not present in the decision-makers’ representation of the world, that is, not in her state space (Levinthal, 2011). Creating such a representation is different from creating a lower-dimensional representation of reality (Gavetti & Levinthal, 2000; Csaszar & Ostler, 2020), as typically the "dimension"—that is the aspect of the world that the surprise is about—is not given, and therefore also not given in some available higher dimensional description of the environment (see above; the higher dimensional description corresponds to the “bigger” statespace in above argument). Dealing with surprise is somewhat similar to sensemaking. Some new aspects of the world are put into brackets and interpreted (cf. Weick, 1988, 1993), and this process is generating a description of the world. It is also somehow related to the process of categorization (Porac & Thomas, 1990) in the sense that an as yet undescribed phenomenon is labeled, that is, attributed to a category of events or entities, and thereby interpreted. Sensemaking can be understood as a process of abduction (Peirce 1978) in which something hitherto unlabeled becomes labeled, involving a search for the best possible explanation that fits the phenomenon. However, what is not addressed by the sensemaking literature is how exactly these new explanations integrate and potentially alter the beliefs that a firm holds about its environment. There are qualitative studies that speak to the issue (Barr, Stimpert, & Huff, 1992; Tripsas & Gavetti, 2000), but we are lacking a theoretical basis to understand belief change when firms face something that lacked an *ex ante* description.

**Belief Revision**

Dealing with surprise encompasses more than sense-making and categorization, as it inherently *re-organizes* mental representations of the world. For example, surprise often forces strategists to *stop* believing something. Overall, after a new description for the unknown unknown is found (which then by our definition turns into a known unknown), representations and subsequently strategies usually need to be altered. Firms may have been searching for
representations and strategies already before the unknown unknown was encountered, in which case the search may need to be adjusted. This process involves expansion and/or contraction of the set of beliefs that constitute the firm’s representation of the world (Gardenfors, 1988).

Such processes of belief revision are different from probabilistic belief changes (Spohn, 2012). Essentially, conditioning on an event in a Bayesian framework means to narrow the support of the overall probability distribution of the decision-maker. It means to states that are considered to be possible. Belief revision, on the other hand, operates with a set of beliefs that constrain the possibilities that a decision-maker sees. Suppose that you believe that the world has four possible states: (A, B), (A, not B), (not A, B) and (not A, not B). And you believe that A is true. Thus, you believe that the world can be in two states (A, B) or (A, not B). If you now learn that A implies B, you need to rule out the state (A, not B). Thus, while your set of beliefs increased (you believed A is true before, now you believe A is true and A implies B), the set of states that are consistent with your set of beliefs decreased (the only state that is consistent with your reasoning is (A, B). This process is called expansion in Gardenfors’s (1988) framework, if the new belief is consistent with what you believe already. However, if you now learn that B is false, you need to revise your beliefs. This process involves a process of contraction: You need to throw out some beliefs, as you cannot believe A is true, B is false, and A implies B. Importantly, the contraction could be solved in different ways: You could give up the believe that A is true OR that A implies B, or both. Belief revision provides us with criteria about how this should be solved that are not provided in probability theory.

If an unknown unknown is encountered, firms should probably not discard all their previous knowledge about an environment. Of course, adjusting a mental representation by describing an

---

7 The difference between belief revision and probabilistic updating is well known in the statistics community. For instance, Judea Pearl, while being an authority in statistics, also publishes on the topic of belief revision (Darwiche & Pearl, 1997), using formal technologies that are different from probability theory.
unknown unknown will *expand* representations, as a new aspect is considered and included in a representation. However, often expansion creates tensions and contradictions; thus, previously held beliefs need to be discarded. In other words, if the explanation of the unknown unknown contradicts prior wisdom a *contraction* of a current mental representation is necessary (Gardenfors, 1988).

As an example, consider the epistemic situation faced by a hotel chain manager who hears rumors over the summer of 2008 of an entirely new concept, an online platform that serves as an intermediate between suppliers of vacation rentals and consumers of such services. This manager faces the unknown unknown that became known as “Airbnb” in August 2008. The manager will likely engage in sense-making by trying to fit Airbnb into familiar categories, perhaps viewing Airbnb as a portal for “hotel like stays” or “short-term rentals of private apartments.” But that will not be sufficient. After a while, she will need to revise her models to make predictions, for example, about prices or revenues. She cannot simply incorporate Airbnb listings into existing models of competitive pricing: Airbnb listings are not “competing hotels” as such. Rather, the relations of demand, offer descriptions, availability and price will change, as customers become aware of the new opportunity, preferences adapt, competitive conditions change, and so on. In other words, the chain manager will have to re-do important aspects of pricing models to predict profits of her hotel chain in the next season. But, it would not be wise to throw all prior knowledge away (e.g., there should still be peak demand over Easter and in the high season in Summer). So, what should this manager continue to assume when she re-does her pricing models? What should she no longer assume, exactly?

Cognitive science research on belief revision (James, 1907; Harman; 1986; Gardenfors, 1988) is essentially about the art of keeping beliefs if you need to throw some of them away, as exemplified with the hotel manager who faces the entry of Airbnb and has to manage her transition from an unknown unknown to a known unknown that she can learn about systematically. Belief
revision theories rest on a number of assumptions. One is that people typically aim at logical consistency in their beliefs such that the process of revision typically involves restoring consistency which may have been upset by a surprise. This often may not take place in a “global” sense, and there may conceivably be epistemic benefits of the ability to entertain some inconsistent beliefs.\(^8\) However, individuals typically seek consistency within a domain, such as their work life, and therefore they tend to detect logical inconsistencies when thinking about a task and to remove them (Khemlani & Johnson-Laird, 2011). In the domain of strategic management, qualitative research suggests that making sense of something new often means to create consistency with beliefs that a top management team already holds (Tripsas & Gavetti, 2000).

**Normative Criteria for Belief Revision**

Restoring consistency of beliefs when thinking about a task is a useful normative criterion for thinking (Harman, 1986), that is, it a criterion decision-makers ought to follow if they want to think rationally. If consistency would not be restored, connections of beliefs would not be taken into account and thus, thinking may lead to ill-founded outcomes or straightforward mistakes.

A second normative assumption that is typically made in belief revision theories is that we should revise beliefs in a minimal way (James, 1907; Harman, 1986; Gardenfors, 1988), that is, keep as many beliefs as possible when consistency of beliefs is restored. This may be seen as an application of a basic principle of bounded rationality: Our cognitive resources are limited and need economizing. There are typically several ways to restore consistency of beliefs. In the above abstract example, we could throw out the belief A is true, or the belief A implies B, or both. Typically, the minimality principle would now require us to throw out *only one* of the two beliefs. In the abstract context of logical reasoning, consistency of beliefs can for instance always be achieved by radically throwing all beliefs away. In the abstract example, we could throw out A is

---

\(^8\) Cf. the famous aphorism attributed to F. Scott Fitzgerald that “The test of a first-rate intelligence is the ability to hold two opposed ideas in mind at the same time and still retain the ability to function.”
true and the belief A implies B. However, this approach would not make sense in the context of a firm that is adapting to a novel situation. We need to have systematic principles for giving up beliefs, that also allow us to incorporate some knowledge about the origin and purpose of beliefs. For instance, A implies B may be a general law that is worth preserving. But that needs not be the case. A implies B could also just be a pattern that was learned earlier, but A is true could express a core value of a strategist. Thus, we need to account for the meaning of the beliefs in context to say something about how we should revise them.

When firms face unknown unknowns it probably makes sense to “shield” their core values and ideas as much as possible, that is, “shield” some beliefs from being thrown away. In terms of the AirBnb example, it makes sense to shield the belief that prices are determined by supply and demand (implying, e.g., that high season means peak prices). However, it is also well established that shielding beliefs may create bad outcomes—if the wrong beliefs are shielded. For example, the Tripsas and Gavetti (2000) Polaroid study shows that the top management team shielded Polaroid’s belief in the razor and blade model, hindering adaptation to the increasingly digitalized market for photographic images. This attachment made Polaroid basically stick to its core business of photographic film, even though Polaroid had already invested in digital technology and had done so successfully (ibid.). As the example suggests, the question is not so much as to whether shielding per se is useful, but which beliefs should be shielded. This question is of fundamental importance to strategy research. But it seems hard to decide on which beliefs should be shielded in the presence of unknown unknowns because the decision situation may be entirely novel. One way to approach this is to ask if there are ex ante criteria to discern which beliefs of a firm deserve shielding, and which not, when the firm faces an unknown unknown.

Note again that imposing a minimality criterion—keeping as many beliefs as possible—does not in itself determine which beliefs should be kept. As a normative criterion, minimality is
incomplete. In the above abstract example, minimality does instruct us to only give up “A is true” or “A implies B,” but not which one of them. In the business context, suppose that I believe that my company’s current business model is a razor and blade model. Moreover, I believe that a razor and blade business model is the basis for the past success of my company. Now I observe that my company is no longer successful. I could now either stop believing that a razor and blade model is the basis for my company’s success, or I could stop believing that my current business model is a razor and blade model. Both would help to restore consistency in my beliefs. The minimality principle would imply that I throw away only one of these two beliefs, not both, but it does not tell me which of the two should be thrown away. So, the minimality principle needs to be augmented with an idea about which beliefs are worth being shielded.

New Research Questions on Belief Revision

Thus, while strategy research can “import” two useful assumptions for reasonable (perhaps even “rational”) strategies to revise beliefs while facing unknown unknowns (restoring consistency and revising minimally), doing this raises further questions that need to be answered.

First, what are effective principles of ordering beliefs in terms of their value in the context of adaptation? In other words, should strategists stick to some beliefs than others? Is there value, for instance, in keeping certain law-like beliefs which may not be specific to an environment? If yes, what are examples of such laws? Raising these issues is particularly important in light of arguments that emphasize the general importance of commitment under uncertainty (Ghemawat, 1991), and also harmonizes with recent arguments (Van den Steen, 2017, 2018) that strategy revolves around “the smallest set of choices to optimally guide (or force) other choices.”

Consider again the example of Airbnb. It obviously was necessary to question industry wisdom about trust in the booking process, but other, more general laws (certainly the law of supply and demand which implies peak prices in the high season, but also insights about the value of
customer satisfaction and loyalty) from the hotel industry could usefully be applied to think about an accommodation market with Airbnb as a player. Are there ex ante possibilities to discern which laws to stick to and which laws to question when an environment changes?

Is emotional attachment to some beliefs always bad (Tripsas & Gavetti, 2000)? Or can emotional attachment to beliefs be useful? In the cited Polaroid case, the TMT’s emotional attachment to the razor-blade model was obviously harmful in Polaroid’s adaptation attempts. But think about the growth of Wholefoods in the US. Wholefoods became a successful brand in the upper middle class (some people labelled it “whole paycheck”) and Wholefood started to sell luxury food. However, Wholefoods stuck to its beliefs in organic sourcing of the food and local community integration and did not develop into a mere upmarket supermarket, which preserved its USP and positioning (Mackey & Sisodia, 2014). Attachment to beliefs appears to have both positive and negative aspects. Can we discern the former and the latter using ex ante criteria?

COMING TO GRIPS WITH UNKNOWN UNKNOWNS IN STRATEGY RESEARCH

The Neglect of Unknown Unknowns

Unfortunately, strategy research does not provide answers to the above questions. While they would seem to be at the core of behavioral strategy research, much of this research is based on the psychological literature in the tradition of Kahneman and Tversky. Research into belief stickiness under labels such as confirmation bias (Tversky & Kahneman, 1974) investigates when strategists (wrongly) stick to beliefs even though it is clear how beliefs should be revised. Thus, the confirmation bias points to cases in which sticking to beliefs is a straightforward mistake. More generally, the application of the heuristics and biases program to core strategy issues have proven very fertile (for recent applications to strategic factor market and learning arguments, see Leiblein, Chen, & Posen, 2017, and Posen, Leiblein, & Chen, 2018, respectively). However, this broad line of research still has reference to situations in which state spaces and probabilities are well defined,
in other words, in situations of known unknowns. In contrast, belief stickiness is a phenomenon of a fundamentally different nature when firms face unknown unknowns.9

As we observed earlier, the basic problem is that assumptions of well-defined state spaces and probabilities abstract away from the actual challenges of the management of surprise, that is, how contractions of belief sets should be dealt with.10 If probability distributions change qualitatively due to yet unknown types of events, one needs to leave the domain of probability theory and incorporate qualitative factors to anticipate their change (Kay & King, 2020). Indeed, leading econometricians argue exactly along these lines (Castle, Fawcett & Hendry, 2011). But incorporating qualitative factors means to manage state spaces, to exercise judgment. In other words, when facing unknown unknowns, firms need to rely on, and manage, values, beliefs, and judgement.

**Getting Belief Revision Into Strategy Thinking**

As suggested earlier, the reason why strategy research has made relatively little progress when it comes to theorizing unknown unknowns, the surprise they produce, and how decision makers deal with such surprise, likely is that the field has been entrenched in analytical technologies that assume an external reality that is in principle already captured in existing state-spaces. Thus, strategy research as often abstracted away from systematic inquiry into the question how state spaces are constructed and revised. The implication seems to be that, somehow, state spaces are presented to decision-makers.

---

9 More generally, it is questionable whether psychological research results that have been derived for situations with known unknowns (probabilities defined in principle) can be applied for situations in which strategists face unknown unknowns. Thus, for instance, it is questionable whether results around confirmation bias (Tversky & Kahneman, 1974) are applicable to the understanding of belief revision in context of unknown unknowns.

10 For a related reason, we think that it is not enough to analyze uncertainty in entrepreneurial decision-making by leaving probabilities unspecified, but keeping the state and action space of the entrepreneur fixed, as in, for example, Gans, Stern, and Wu (2019). Thus, “adding” a description of something new, that is, expanding a state space, is not modeled.
However, a recent metastudy confirms that it makes a huge difference for human decisions whether state space information that pertain to a given decision task is acquired through description or experience (Wulff, Mergenthaler-Canseco & Hertwig, 2018). There is no reason to think that decisions that pertain to strategy are any different in this respect. Essentially, people construct state spaces when they are faced with new types of events and then use them to estimate probabilities. Humans seem to adapt well by combining state-space construction skills with probability estimation skills. As the Kahneman and Tversky tradition, including its various offshoots in behavioral strategy, only focuses on the latter skill, it may misinterpret human behavior as biased when it is highly functional, given that humans cannot take a description of a situation as given (Wulff et al., 2018.). Thus, the way in which state spaces are constructed may need to be factored into our analysis. Papers such as Gavetti and Levinthal (2000) and Levinthal (2011) already address the problem of the construction of mental representations (e.g., in these papers small world representations are lower-dimensional state spaces), but the problem of expansion and contraction of beliefs over time is not addressed in these papers. The same applies to current attempts to model firms’ shaping of their business context (Gavetti, Helfat, & Marengo, 2019). However, belief revision theories are available to model belief revision, helping us to analyze how strategists should cope with surprises and contradictions (Ehrig & Schmidt, 2020b).

Belief revision theories have been formalized (Gardenfors, 1988; Pearl, 1998; Spohn, 2004) and can in principle be used to analyze learning and choice in strategic contexts in novel ways. However, those who want to follow this route should also be prepared to also change their standards of analysis in modelling. Notably, we may need new ideas about what efficiency or even “rationality” can mean when firms face unknown unknowns (Ehrig & Schmidt, 2020b). Clearly, when descriptions of the world change along the way (when unknown unknowns turn into known unknowns) any notion of optimal choice given pre-specified alternatives seems meaningless.
However, it may still be possible to speak of optimality in a more restricted sense. For example, "optimal" learning may mean to optimize the trade-off between the speed of learning and the expected prediction error (cf. Ehrig & Foss, 2021). Ehrig and Schmidt (2020b) propose that “learning quicker if one’s subjective assumptions are wrong” is a useful learning efficiency criterion that does not break down when descriptions of the world are in flux; but perhaps there are yet more important undiscovered efficiency criteria for managing unknown unknowns.

In sum, building up the capability to systematically analyze the management of unknown unknowns as the strategic management field will not come “cheap” as a mere augmentation or tweak of the analytical technologies that are already being applied in the strategy field. It will require us as a field to search for efficiency and rationality criteria that are applicable when the conditions outlined by Savage (1956) are not met. In particular, it will require us to re-think what optimality can mean in adaptation processes.

**CONCLUSIONS: HOW THE PROPOSED AGENDA MAY INFLUENCE STRATEGY**

Our key argument is that strategists need specific cognitive skills when they encounter unknown unknowns that cause a need for adaptation. In particular, they need to explain the unknown unknown, that is, they need to connect the unknown with the known (Barr, Stimpert, & Huff, 1992). The ideas in this essay are in line with prior approaches to adaptation, but points to an omitted important aspect of adaptation. While we, like prior qualitative research (Tripsas & Gavetti 2000) point to the importance of cognition in adaptation processes, we argue that for a systematic analysis of such cognition we need 1) a formal “technology” for organizing our thinking about it and 2) normative criteria. Moreover, we need more rigorous theorizing than the extant research on sensemaking and categorization. In particular, we need formal models to analyze how firms can successfully cope with unknown unknowns. Such formal research should go hand in hand with the development of useful normative criteria to shield beliefs in *ex ante* perspective. For instance, we
observed that we need to study the art of discerning which beliefs deserve shielding in an *ex ante* perspective, and which beliefs a strategist should be ready to let go—again, taking an *ex ante* perspective—unknowns are still “unknown.” We argue that prioritizing the shielding of different beliefs is a key aspect of adaption to unknown unknowns. A systematic treatment of this problem may have consequences for several strands of strategy research.

First, while the literature on dynamic capabilities has indeed pointed to the importance of belief revision and the corresponding emotional processes (Teece, 1997; Hodgkinson & Healey, 2011: 1504), this analysis is in its infancy. Both theoretically and practically, facing unknown unknowns forces firms to reconfigure their resources. Thus, belief revision processes are key to sensing and seizing (Teece, 2007), because the resulting revised beliefs are the basis for a re-evaluation and thus reconfiguration of resources. But, as we have argued current research offers no normative criteria on how beliefs should change when firms face unknown unknowns. In particular, no normative criteria for shielding beliefs are known in the strategy field. Thus, we currently don’t understand what may lead to advantages when firms face unknown unknowns and resources need to be reconfigured. It seems obvious that differences in how existing beliefs are altered and in particular, which beliefs are given up when contradictions arise, may have a significant impact on firm adaptation.

To revise beliefs effectively matters for sensing and seizing opportunities, but also for shaping the business context (as the beliefs of other stakeholders may need revision while a market is shaped) and for firm-internal reorganization processes (as the beliefs of employees may be subject to revision when firms re-organize their structure). Thus, belief revision processes provide important micro foundations for most, if not all, dynamic capabilities. Current research on exploration and exploitation (Brusoni, Laureiro-Martínez, Canessa, & Zollo, 2020) provides experimental settings that could be extended to study the questions we outline above systematically.
Thus, our argument implies an agenda for the study of the origins of advantages in resource reconfiguration processes.

Our argument implies that we may question some key aspects of the by now established strand of literature labeled “behavioral strategy” (Powell, Lovallo & Fox, 2011; Hambrick & Crossland, 2018), at least as it applies to situations of unknown unknowns and to the extent that this literature makes use of the Kahneman and Tversky line of research (ibid.), which is fundamentally about experimentally measured deviations of human behavior from a rational standard of behavior. The implicit assumption is that the world could in principle be knowable in probabilistic terms and strategic advantages come about by avoiding mistakes that others make (Denrell & Fang, 2010). Mistakes are well-defined as there are objectively correct ways to update probability distributions.

But, as argued above, objectively correct ways to update probability distributions are not available when firms face unknown unknowns. In this situation, we need to fundamentally revisit efficiency and rationality criteria: What does it mean to rationally adapt to unknown unknowns? Are the two criteria we proposed above (consistency and minimality of revision of beliefs) useful? Before we have such rationality criteria, we cannot speak of “behavioral” in the sense of a deviation from a rational standard, at least not in the kind of decision situations we are addressing here. We thus agree with Levinthal’s (2011) rhetorical question “What is the alternative to bounded rationality?,” but note that Levinthal did not exhaustively list bounded rational technologies available for adaptation. As we have argued, technologies do exist for situations of decision-making under radical uncertainty, and these need to be brought into strategic management theory.

Once we accept that we can usefully incorporate ideas from the belief revision literature, we can add important points to the research agenda of the “behavioral strategy” community:

1. Can we unpack the above-mentioned ex ante criteria for shielding beliefs? We probably can, as there are laws, regularities, and mechanisms that are likely to prevail, even
when industries transform radically. Thus, at the most fundamental level, the law of
supply and demand, should prevail as long as markets are an accepted institution, and as
long as scarcity exists.

2. New qualitative research questions could be posed, such as, Under which conditions did
firms shield beliefs that deserved shielding, in an *ex ante* perspective? Under which
conditions did organizational processes hinder the shielding of beliefs that appeared
worthwhile shielding in an *ex ante* perspective?.

3. We need new formal models for the analysis of learning and adaptation. The above-
mentioned belief revision principles have been formalized (Gardenfors, 1988, Darwiche
and Pearl, 1997; Spohn, 2012). There is also relevant research on “unawareness” in
games (e.g., Heifetz, Meier, & Schipper, 2008). While, the application of the available
formal modeling possibilities to strategy remains in its infancy (Ehrig and Schmidt,
2020b), formal treatments of belief revision in strategy contexts can help to address
questions such as: When do firms with different *ex ante* representations of contingencies
converge to shared beliefs when they jointly face the same unknown unknown? When
do beliefs not converge? When can entrepreneurs convince financiers to incorporate a
yet unforeseen contingency into their representation of the world?

Addressing and resolving such questions will impact our key theories of competitive
advantage. As already suggested, reactions to unknown unknown often redefine what are key
strategic factors to different firms. For instance, Airbnb turned empty private apartments into a
relevant strategic factor, and local governments’ regulations of Airbnb rentals turned into a relevant
policy parameter for the hotel industry. Neither this relevant factor, nor the latter parameter was
likely to be found in the mental models of strategists in the hotel industry before Airbnb became
known. Additionally, if a firm is quicker in reacting to an unknown unknown, it may be able to
acquire strategic factors to cope with the changes set in motion by the emergence of the unknown unknown at a lower price than competitors. Likewise, if a firm is better in anticipating reactions of others to an unknown unknown, it may have advantages in predicting price changes in factor markets. As managing belief revision processes is a key skill to cope with unknown unknowns, firms with an edge in managing belief revision processes should thus be able to acquire factors below their value in equilibrium.\(^{11}\)

In sum, a better understanding of ways to cope with unknown unknowns is equally important for practitioners as for strategy scholars. As we argued in this essay, this calls for an in-depth study of normative criteria to manage belief revision processes. We hope that other scholars will join us in this important investigation.

REFERENCES


\(^{11}\) When the unknown unknown turned into a known unknown for most strategists, the equilibrium logic of Barney (1986) applies again.


Wiblin, R. and Harris, K. 2019. Accurately predicting the future is central to absolutely everything. Professor Tetlock has spent 40 years studying how to do it better. 80,000 hours Podcast.: https://bit.ly/33HaeNS


