This is the version of the article accepted for publication in Academy of Management Proceedings published by Academy of Management: https://doi.org/10.5465/AMBPP.2018.141
Accepted version downloaded from SOAS Research Online: http://eprints.soas.ac.uk/29955

UPDATING SMALL WORLD REPRESENTATIONS IN STRATEGIC DECISION-MAKING UNDER EXTREME UNCERTAINTY

ALBERTO FEDUZI
School of Finance and Management, SOAS University of London
10 Thornhaugh Street, Russell Square, London, WC1H 0XG, UK

JOCHEN RUNDE
Cambridge Judge Business School, University of Cambridge

LAURE CABANTOUS
Cass Business School, City, University of London

PHIL FAULKNER
Clare College, University of Cambridge

CHRISTOPH LOCH
Cambridge Judge Business School, University of Cambridge

INTRODUCTION

Strategic decision-makers (henceforth SDMs) typically operate in complicated and messy "large worlds" very different from the simple and sanitized "small worlds" of traditional decision theory (Savage, 1954) in which the decision-maker knows the full list of possible "acts", "states" and "consequences" required to identify optimal solutions in accordance with the subjective expected utility (SEU) theory. In large worlds, in addition to risks that can be identified at the outset, SDMs may face extreme uncertainty in the form of *unknowns* and the potential *Black Swans* that may ensue from them (Feduzi & Runde, 2014). And where there are unknowns, state spaces are likely to be incomplete, classical probabilities are ruled out, and it is generally more difficult for SDMs to identify and select between appropriate courses of action.

The problem of unknowns is especially acute where SDMs face ambiguous and non-definitive information (Forbes, 2007; Gavetti & Rivkin, 2007). In the worst cases, trial and error (Cyert & March, 1963) or learning after the fact may be more effective than thinking in advance (Gavetti & Rivkin, 2005). Yet there remain strategic problems that do require thinking in advance and where boundedly rational SDMs may reasonably attempt to anticipate the broad consequences of broadly defined courses of action by constructing Small World Representations (SWRs) "approximating Savage's (1954) notion of a perfect small world" (Maitland & Sammartino, 2015: 1556, fn 1; Gavetti & Levinthal, 2000).

However, where unknowns are rampant, the risk of committing prematurely to particular courses of action encourages a higher proportion of strategic decisions to be made during implementation. SDMs may wish to fine-tune their strategies as information accumulates, and to be in a position to contemplate courses of action not apparent at earlier points in time (Klingebiel & De Meyer, 2013). Strategic decision-making then becomes an ongoing, adaptive process in which

SWRs may be updated several times over the course of the implementation phase.

While there is a large literature that addresses how SDMs (should) go about *constructing* SWRs (e.g., Maitland & Sammartino, 2015), there is little on the prescriptive question of how they should *update* their SWRs during the implementation phase under extreme uncertainty. We take a step towards addressing this question, drawing on work on reasoning in psychology. We focus on the problem that the lists of scenarios associated with particular SWRs tend to be highly incomplete in strategic decision processes, and investigate how two heuristic methods of inquiry—disconfirmation and counterfactual reasoning—perform in the updating of SWRs in situations characterised by ambiguous and non-definitive information.

SMALL WORLD REPRESENTATIONS, UNKNOWNS AND BLACK SWANS

To capture the relationship between SWRs, unknowns and Black Swans, we draw on Faulkner, Feduzi and Runde (2017). Faulkner et al. define a *known* as "any feature of the world that an individual has knowledge of", and an *unknown* (or *gap in knowledge*) as "any feature of the world that an individual lacks knowledge of" (2017: 1282), where a *feature of the world* is any fact "about past, present and future reality" (2017: 1281). They further distinguish between a *known unknown*—"a gap in knowledge that an individual knows about and is aware of at the relevant time"—and an *unknown unknown*—"a gap in knowledge that an individual is not aware of at that time, either because they do not know about that gap in knowledge or because, despite knowing of it, they are unaware of it" (2017: 1283). Finally, they introduce the concept of *hypothetical values* associated with a gap in knowledge, where a hypothetical value is "any value—outcome, state of affairs, result, quantity and so on—that could conceivably be thought to be a candidate for the actual or true value of the unknown under consideration" (2017: 1283).

Consider a SDM in the process of constructing scenarios that might ensue from some strategic decision. The unknown in this case is how the future will unfold, and where the relevant hypothetical values are all of the scenarios that could conceivably be imagined with respect to that unknown. A scenario is a hypothetical future state of affairs flowing from a set of imagined influences, where an influence is any event or condition that could contribute causally to a state of affairs. Adapting Faulkner et al. (2017), all scenarios that could conceivably be imagined can be assigned to one of the cells in the table below.

Insert Table 1 about here

Cells 1 and 2 contain scenarios imagined and regarded as possible by the SDM, split between those that are correctly (cell 1) and mistakenly (cell 2) regarded as such. Taken together, the scenarios in these two cells constitute the SDM's *subjective scenario space* (henceforth SSS). The scenarios in cells 3 and 4 are those consciously imagined but excluded from the SSS because they are regarded as impossible, correctly so in cell 4 but mistakenly so in cell 3. The scenarios in cells 5 and 6 lie outside the SSS because they are not even imagined.

We can now make the connection with the Black Swan popularised by Taleb (2007). One of Taleb's (xvii–xviii) fundamental conditions for a Black Swan is that it is an event that is "an outlier, as it lies outside of the realm of regular expectations, because nothing in the past can convincingly point to its possibility." A natural interpretation in light of this condition is that Black Swans arise where states of affairs occur *ex post* that did not figure in the SSS *ex ante*. Since a

Black Swan arises only once the relevant event occurs, it follows that the corresponding scenario would have had to be genuinely possible *ex ante*. We can thus disregard as candidate Black Swans the scenarios located in the right-hand column of Table 1. Black Swans arise only in respect of scenarios located in cells 3 and 5 and avoiding Black Swans then means minimizing the number of scenarios in cell 3 and bringing as many of the scenarios in cell 5 into cell 1.

INSIGHTS FROM THE PSYCHOLOGICAL LITERATURE ON REASONING

Since scenarios are hypotheses generated and evaluated by boundedly rational decision-makers, it is natural to ask what the psychology literature on hypothesis–testing behaviour (Poletiek, 2001) might tell us about how SDMs test or should test scenarios.

A popular experimental task used to evaluate hypothesis-testing behaviour is Wason's (1960) 2-4-6 discovery task. Here the experimenter decides a rule that generates particular triples of numbers, e.g., "three numbers in increasing order of magnitude". Subjects are given a triple consistent with that rule (e.g., 2-4-6) and invited to uncover the rule. They do so by generating hypotheses about what the rule might be and testing these by providing triples of their own to the experimenter, who then tells them whether their triples conform to the rule. When subjects feel they have identified the rule, they are instructed to declare this. The task ends when a subject correctly identifies the rule or gives up. Wason found that the hypothesis-testing performance of his subjects was quite poor and attributed this to a bias towards using a *confirmation heuristic* of testing their hypothesized rules by generating triples consistent with those rules.

Wason's task and its variants went on to be used to study the efficacy of two heuristic methods of inquiry used in hypothesis-testing: *disconfirmation* and *counterfactual reasoning*. Disconfirmation recommends that the subject start with a working hypothesis, assume this hypothesis is correct, and then perform a series of tests that involve searching for new evidence inconsistent with the working hypothesis. If such evidence is found, the working hypothesis is eliminated, a new hypothesis generated, and the process begins again. The higher the number of tests passed by any working hypothesis, the greater the subject's confidence in it.

Counterfactual reasoning directs the subject to attempt to falsify any working hypothesis by searching for evidence in favor of alternative hypotheses. Here the subject again begins with a working hypothesis and performs a series of tests, but in this case on the assumption that the working hypothesis is false. The process involves generating an alternative to the working hypothesis that is consistent with the existing evidence, and then searching for new evidence that would be consistent with the alternative hypothesis and inconsistent with the original working hypothesis. If such evidence is found, the original working hypothesis is eliminated and the alternative becomes the new working hypothesis, and the process begins again. The higher the number of tests passed by any working hypothesis, the greater the subject's confidence in it.

Several studies have investigated whether disconfirmation improves performance in the 2-4-6 task and found that its effectiveness depends on whether subjects are able to appeal to an external authority for feedback and on the nature of the rule (Gorman, 1995). Further, some studies (Tweney et al., 1980) have found that disconfirmation may be effective at later stages of the inferential process, after a confirmed hypothesis has been identified. But other studies argue that counterfactual reasoning is the key to improving subjects' performance in the 2-4-6 task and its variants (Farris & Revlin, 1989a,b; Oaksford & Chater, 1994) because it induces subjects (1) to generate and test more than one hypothesis at a time, thereby forcing them to consider more hypotheses than they would otherwise, and (2) to perform negative tests because, when two competing hypotheses are tested simultaneously, the alternative hypothesis testing positive corresponds to the working hypothesis testing negative (Evans, 2016).

DISCONFIRMATION AND COUNTERFACTUAL REASONING

SDMs are also in the hypothesis-testing business, also suffer from the confirmation bias, and also have the option of using disconfirmation or counterfactual reasoning to attenuate it. But situations of strategic decision-making differ from the laboratory in important ways. We focus on two of these differences. The first is that whereas the feedback in Wason-style tasks comes from an external authority and is unambiguous and reliable, strategic decision-making tends to involve *ambiguity*. The second is that while the information in Wason-style tasks typically allows hypotheses to be rejected rapidly and decisively, the information in strategic decision-making is seldom definitive in this sense.

Offsetting The Confirmation Bias: The Ambiguity Effect

While disconfirmation is effective in counteracting the confirmation bias in situations of low ambiguity, it is less so when ambiguity is high. The reason for this is that that newly acquired negative evidence may be false and therefore not to be acted on until more information comes in. This possibility leads to "hypothesis perseveration" (Gorman, 1989) in experiments in which subjects are instructed to use disconfirmation, but know that the feedback they receive has some probability of error. Results show that subjects tend either to immunize their hypotheses against disconfirmation by treating disconfirming results as errors, or to correctly recognize that there was no error but then devote so much time to replicating experiments that they fail to adequately falsify their hypotheses (Gorman, 1986, 1989). SDMs also face the challenge of the threat of false negatives providing grounds to resist giving up on existing scenarios immediately in the face of negative evidence. It follows that, the more ambiguous the information, the less inclined SDMs relying on disconfirmation will be to give up immediately on their existing scenarios, and the less effective disconfirmation will be in offsetting the confirmation bias.

Ambiguity also reduces the extent to which counterfactual reasoning attenuates the confirmation bias, but the reduction is less damaging here. In contrast to disconfirmation, counterfactual reasoning is susceptible to false positives because it is aimed at including alterative scenarios on the basis of positive evidence. False positives are however less disruptive than false negatives because no matter how much evidence in favour of an alternative scenario is collected, it is never possible to verify it conclusively unless and until it occurs. And while false positives may lend alternative scenarios excessive credibility, this is welcome for encouraging SDMs to question the existing set of scenarios and take more seriously than they would otherwise the possibility of there being alternative scenarios. Of course, it may happen that false positives reduce the credibility of existing scenarios to the point of falsely disconfirming them. But in this case hypothesis perseveration works in a positive way by putting a brake on SDMs giving up on their existing scenarios immediately on including new ones. We thus arrive at:

> *Proposition 1: In the face of ambiguity counterfactual reasoning helps SDMs offset the confirmation bias and thereby question their existing set of scenarios more effectively than does disconfirmation.*

Promoting The Exploration Of The Scenario Space

We now compare disconfirmation and counterfactual reasoning with respect to how ambiguity and non-definitive information affect the rate at which new scenarios are generated and the overall number of scenarios contemplated.

Disconfirmation is fundamentally *reactive* method as it moves from the elimination of existing scenarios in the light of negative evidence to the generation of new influences and therefore new scenarios. The generation of new scenarios, and thus the exploration of the scenario space, emerges only as a by-product of updating beliefs in existing scenarios. However, since boundedly rational SDMs faced with ambiguous information tend to stick with their existing scenarios until they feel they can trust market signals, ambiguity reduces the speed at which new scenarios are generated and thereby the number of scenarios considered.

In contrast, counterfactual reasoning is *proactive* as it moves from the generation of new influences (and therefore new scenarios) to attempting to find evidence relevant to them. The generation of new scenarios, and hence the exploration of the scenario space, thus becomes integral to the process of hypothesis testing. Moreover, as SDMs are explicitly directed to look only for positive evidence in favour of alternative possible scenarios, this requirement introduces a virtuous bias insofar as it encourages SDM to take them seriously. Ambiguity amplifies this effect by further increasing the credibility of alternative scenarios and therefore increasing the speed at which SDMs using counterfactual reasoning explore the scenario space and the number of scenarios that they consider. We thus arrive at:

> *Proposition 2a: In the face of ambiguity, SDMs using counterfactual reasoning will contemplate more scenarios than those using disconfirmation.*

While the effects of ambiguity are often significant in the short run, they tend to diminish over time as the evidence mounts and uncertainties are resolved. SDMs will accordingly update their scenarios space anyway over time, whatever method of inquiry they adopt. However, the information in strategic contexts is rarely definitive, the rate at which scenarios are generated and evaluated depends on the method of inquiry used.

SDMs using disconfirmation generate a new scenario only when new negative evidence is sufficient to eliminate an existing one. While this may sometimes occur, new evidence will often only throw doubt on a scenario, and it will take time to acquire sufficient disconfirming evidence for that scenario to be eliminated. This is all the more so the more confident SDMs are that a specific scenario provides a good guide to action. Of course it is possible that negative evidence might build sufficiently to justify the elimination of an existing scenario, and where the SDM is then prompted to explore the scenario space and come up with an alternative scenario. The speed of such exploration, and accordingly the overall number of scenarios contemplated, is however significantly—and negatively—affected by the non-definitive nature of the available evidence.

The non-definitiveness of information also affects SDMs using counterfactual reasoning as the acquisition of new (positive) evidence leads to only gradual increases in the credibility of new scenarios and corresponding decreases in the credibility of existing ones. However, since counterfactual reasoning does not require an existing scenario to be eliminated for a new one to be generated, the non-definitive nature of the information does not impact on the rate at which SDMs explore the scenario space. The set of scenarios will grow until the credibility of any existing scenario is driven low enough for it to be eliminated. But this is not a problem, as the goal here is exactly that of exploring the scenario space rather than merely updating it, and of increasing the overall number of scenarios contemplated. We thus arrive at:

Proposition 2b: In the face of non-definitive information, SDMs using counterfactual reasoning will contemplate more scenarios than those using disconfirmation.

Coping With Black Swans: Taking Action

Finally, we compare disconfirmation and counterfactual reasoning in terms of how well they allow SDMs to mitigate or exploit what would otherwise have been Black Swans.

Disconfirmation invites SDMs to generate new scenarios only when an existing scenario is eliminated in the light of new negative evidence. This limits the extent to which disconfirmation promotes the uncovering of potential Black Swans and being able to respond in a timeous way. In addition, looking only for evidence inconsistent with the existing scenarios might divert attention from alternative influences and scenarios that might usefully be generated. Finally, even if disconfirmation happens to be sufficiently effective to rule out all scenarios that had been contemplated up to that point, SDMs will be left without any hypothesis about where the future is heading, reducing their capacity to anticipate what would then be Black Swans for them.

In contrast, by encouraging the constant generation of alternative scenarios, counterfactual reasoning leads to a larger number of scenarios being tested at any stage, thereby promoting the uncovering of a larger number of potential Black Swans. If sufficient evidence in favor of a new scenario can be found, this opens the way to identifying and implementing completely new courses of action. Moreover, when new evidence indicates that all existing scenarios should be disregarded, SDMs are never left without a hypothesis to fall back on. This leads to:

Proposition 3: In the face of ambiguous and non-definitive information, counterfactual reasoning is superior to disconfirmation in helping SDMs take actions to mitigate or exploit what would otherwise have been Black Swans.

CONCLUSION

In highly uncertain situations, counterfactual reasoning is more effective than disconfirmation with respect to (1) counteracting the confirmation bias, (2) promoting the exploration of the scenario space, and (3) facilitating the adoption of actions to mitigate or exploit what would otherwise have been Black Swans.

REFERENCES AVAILABLE FROM THE AUTHORS

TABLE 1
Classifying The Scenarios Associated With A Strategic Decision

<table>
<tr><th colspan="3"></th><th colspan="2">Scenario is …</th></tr>
<tr><th colspan="3"></th><th>a genuine possibility</th><th>not a genuine possibility</th></tr>
<tr><th rowspan="2">Scenario is …</th><th rowspan="2">consciously imagined and …</th><th>regarded as possible</th><td>1. Scenarios imagined and correctly regarded as possible</td><td>2. Scenarios imagined and incorrectly regarded as possible</td></tr>
<tr><th>regarded as impossible</th><td>3. Scenarios imagined and incorrectly regarded as impossible</td><td>4. Scenarios imagined and correctly regarded as impossible</td></tr>
</table>

not consciously imagined		5. Scenarios not imagined	6. Scenarios not imagined