

Here (Very Likely) Be Dragons:

The Challenges of Strategic Forecasting

Rex Brynen

McGill University

Forecasting has much to contribute to strategic analysis. It is, however, a difficult—and sometimes thankless—task. This chapter addresses the challenges of policy-relevant prediction, and how they might best be addressed. In doing so it draws in part on a major study of forecasting accuracy by Defence Research & Development Canada (DRDC)—the largest unclassified study of its kind ever looking at actual intelligence output.¹ This reviewed scores of reports and more than two thousand specific predictive judgments produced by the Intelligence Assessment Secretariat (IAS) of Canada’s Privy Council (or cabinet) Office. The chapter also draws extensively upon the broader literature on strategic analysis, intelligence, forecasting, early warning, cognition, and organizational responsiveness.

Prediction and Strategic Analysis

Strategic forecasting—and hence the production of policy-relevant predictions²—is a key function for much of the intelligence community, and an implicit or explicit component of most strategic analysis. Policymakers want strategic warning, and hope thereby to avoid surprises that adversely affect important national interests.³ They also want some sense of the possible future strategic landscape, so as to inform the decisions they make now and in the future. The identification of political, social, economic, technological, and military trends is a key part of this. This is not to say that such assessments are necessarily the basis for policy decisions, of course—that is, after all,

A version of this paper will appear as a chapter in Thomas Juneau and Jean-François Morel, eds. *The Pursuit of Relevance: Strategic Analysis in Support of International Policy-making* (Lanham, MD: Rowman and Littlefield, forthcoming 2016). The author can be contacted at rex.brynen@mcgill.ca.

rightly the purview of policy-makers and political leaders, not of analysts. Rather, predictive assessments should be seen as one of many inputs into the policy-making process, and an intrinsically uncertain but potentially very useful one at that.

When the analytic community fails to anticipate major developments, this often spurs debate about “intelligence failure,” what went wrong, and how it might be put right. The Cuban Missile Crisis, the 1973 Arab-Israeli War, the Iranian Revolution, the 9/11 attacks have all been paradigmatic cases within the modern literature on intelligence.⁴ More recent still, the “Arab Spring” and subsequent instability across the Middle East have raised question about whether academics and strategic analysts alike underestimated, and hence under-predicted, the volatility of the region.⁵ Indeed, one review of Canadian intelligence assessments admitted “the wave of protests and regime changes that swept the Middle East in 2011 had not been anticipated.” It added that there was no reason to believe that the US, UK, or others had done any better either.⁶

At the same time, there can be a peculiar resistance among strategic analysts to admitting that prediction is a central part of what they do.⁷ While some analytical units may consider the identification of trends and possible future developments as a central component of the products they produce, others may be content with largely describing the present while dropping vague and heavily-caveated hints to what the future implications of all this might be.

Part of the reluctance to admit and embrace the challenge of prediction is an understandable concern that the term itself overstates the degree of foresight and precision that is possible. Reluctance almost certainly arises too from a fear at being proven wrong by events, and being held (fairly or unfairly) held responsible for things that were not anticipated, or predicted developments that failed to occur. Policy-makers, after all, are particularly inclined to remember inaccurate predictions or false alarms, even more so than those assessments that proved to be prescient. While being wrong sometimes is inevitable, the bureaucratic costs of inaccurate judgment can be high.

Consequently, a host of other, fuzzier terms are sometimes used to obfuscate the inherent predictive character of future-oriented judgments: horizon scanning, trend analysis, futures exercises, course of action analysis, campaign planning, and so forth all sound rather less definitive than “prediction” or “forecasting,” helping thereby to deflect

blame for when things go wrong. Yet even the use of scenarios in strategic planning, which advocates are sometimes quick to assert is not predictive—*is* usually inherently predictive, in that it builds upon some assessment of what is possible, plausible, or likely in the future.⁸ Similar, many assessments of capabilities—for example, the acquisition of a new weapon system or the deployment of military forces—often contain substantial inherent future-oriented judgment about how such a capability might be employed, and the way in which doctrine, training, command and control, and other factors might shape its future effectiveness.

Countries with major global responsibilities (say the United States) or which might face sudden serious threats (for example, Israel) tend to place greater explicit emphasis on processes and products that offer explicit forecasting and strategic warning. Other countries, less vulnerable to strategic surprise and where decision-makers likely face less urgency to act or react, may devote comparatively fewer resources to the task.

Bureaucratically, analysts may find it safer to refrain from making forward-looking judgments, and instead provide semi-processed intelligence from which clients can draw their own conclusions. Clients may like this too, especially when the information provided is highly classified. However, the failure to offer some sense of what such data means is an abdication of analytic responsibility. It also means that the preconceived views of clients are even less likely to be challenged—an invitation to the sort of uncritical thinking and confirmation bias that is (as will be shown below) particularly associated with poor predictive judgment.

Overcoming the Obstacles to Policy-Relevant Prediction

As has already been noted, the challenges to effective forecasting are many.⁹ Herein they have been grouped into four broad, and sometimes overlapping, categories: data availability, issue complexity, analytic process, and institutional constraints. The first two of these largely relate in large part to the epistemological challenge of whether we can, indeed, predict the future.¹⁰ The latter two generally address how, if prediction is possible and desired, we might do it better.

Data

Data is the prerequisite of all prediction: without it, it is impossible to assess trends or identify future possibilities. Key information may be missing or imperfect, however, especially when dealing with intelligence collection targets that are difficult to access or seek to deny and deceive. Deception, moreover, is not limited to rival governments, economic competitors, and hostile sub-state actors. In violent or repressive settings, even members of society at large may conceal their true views from others for the purposes of self-protection. Such widespread preference falsification makes it even more difficult for both locals and outsiders to gauge the political situation.¹¹

Some of these deficiencies may be recognized. Others—the “unknown unknowns,” where analysts are not aware that they lack key information to begin with—present an even greater problem.

Addressing these sorts of challenges is, to some degree, one of developing appropriate collection priorities and techniques. This is especially true of military secrets in the classic sense: the performance of a new weapons system, for example, or the location of military assets. While the intelligence “great powers” (the US, Russia, China, and to a significant extent countries like Britain, France, Germany, Israel, and India) are able to collect substantial amounts of secret intelligence, many other countries are rather more limited in the quantity and quality of strategic information they can directly obtain, above and beyond that generated by open sources and regular diplomatic reporting. The smaller “Five Eyes” partners—Canada, Australia, New Zealand—are in the unusual position of having only limited national foreign secret intelligence collection capability and only small analytical communities, but having access to vast amounts of raw and processed intelligence by virtue of their intelligence partnership with the US and UK.

However, much of the key information necessary for strategic forecasting is not secret at all, but can be found in open sources (OS): economic data, public reports on political and social developments, media content, academic publications, and so forth. Modern information technology makes more information of this type more available faster than ever before.

Current conflicts and security challenges illustrate how extensive such open source information can be, and the extent to which it increasingly provides data of a type and quality that was once unobtainable, or obtainable only through classified sources and

methods. In the case of the Syrian civil war, for example, more detailed information on available on casualties than in any civil conflict in history. It was open source investigators who did much of the work of identifying chemical weapons usage by Syrian regime.¹² Information on the identity and attitudes of thousands of violent extremists can be accessed via their online presence.¹³ Individuals, scholars, and non-governmental organizations now routinely make use of overhead imagery and other IMINT tools that were once the exclusive purview of governments.¹⁴

Indeed, the problem has often become one of more data than an analyst has the time to read and contemplate: As Thomas Fingar, former Chair of the US National Intelligence Council, has warned, “Many analysts feel overwhelmed because they attempt to—and cannot—“read everything” that collectors push at them and they know is available in unclassified ...[t]he days when an analyst could, or could be expected to, read everything are long gone.”¹⁵

Efforts to address the flood of strategically-relevant open source information include aggregation and monitoring services (such as the US intelligence community’s Open Source Center) as well quantitative visualization and analysis tool designed to help extract trends and other items of significance from the “big data” that is now available. In the area of technology forecasting and horizon scanning, quantitative tools that track scientific publications, patents, and other indicators are increasingly used to identify emerging technologies.¹⁶

Since 9/11 there has also been a recognition that strategic analysis needs to make better use of both the information and insights possessed by those outside the intelligence community, whether in other parts of government, among academics, or with other subject matter specialists. Such outside experts not only represent a source of outside analytical perspectives, but also often have large amounts of domain-specific knowledge.

However, effectively integrating OS material into strategic forecasting involves far more than collecting the data and making it available. Potential users—especially those in the national security and policy community who sometimes confuse the classification level of a report with its veracity or utility—need to understand its usefulness.¹⁷ Analysts also need both the time and skills to make use of it. Often they have neither—whether due to inadequate training, limited resources and heavy workload,

or poor management. Indeed, part of the reason why the CIA established a dedicated Open Source Center was the view that effective monitoring and communication of OS materials was a specialist function, not something that could be left to analysts and their spare time. Bess Puvathingal and Donald Hantula, go a step further, and note that attention also needs to be given to the psychology of information “foraging,” suggesting that this provides insight into the conditions under which analysts are most likely to gather and use such information productively.¹⁸

Complexity

Strategic political and security assessments often examine systems that are not so much “complicated” as they are “complex,” dynamic, and adaptive—that is to say, more like a natural ecosystem than like a mechanical device like an automobile. Complex adaptive systems consist of dense networks of non-linear relationships and feedback loops. They may be prone to homeostasis, whereby external shocks may be dampened by the system. They may also be prone to cascading changes and failures, whereby small changes in conditions create the conditions for subsequent dramatic transformations. Nassim Nicholas Taleb has warned that “history does not crawl, it jumps”—that is, it is characterized by sudden rapid changes. He warns that “black swans”—extremely rare events of high impact, that only appear predictable in retrospect—have become increasingly important in shaping the trajectory of societies in an increasingly interconnected and globalized world.¹⁹ Such events may even have disproportionate effect precisely because they were *not* anticipated.

Consequently, some have questioned the ability of either scholars or government analysts to ever predict certain types of systemic political and strategic change. Charles Kurzman, for example, has argued that the Iranian revolution could not have been anticipated because even most Iranians considered it “inconceivable” until it was underway, and only at that point did they began to “think the unthinkable”.²⁰ Merouan Mekouar has similarly suggested that the revolts of the Arab Spring were spurred by an informational cascade characterized by rapid changes in citizens’ perceptions of costs, benefits, and opportunity.²¹ Highlighting the magnitude of the challenge, Philip Tetlock’s

work on expert political prediction indicates that the ability of most political pundits to predict real-world events is no better than that of straight-line predictions, simple statistical algorithms, or even chimps predicting outcome by random chance.²²

However, review of the accuracy of actual intelligence product provides somewhat greater grounds for optimism. In the case of IAS assessments, some 94% of analyst predictions were assessed to have been generally accurate.²³ In addition to impressive discrimination scores (that is, correctly separating events that would occur from those that would not), analysts did well in calibration too (assigning probability terms to predicted events that matched their actual rate of occurrence).²⁴ In contrast to Tetlock's findings, which suggested that formal qualifications and expertise had little impact on predictive accuracy, analysis of IAS assessments also found that more experienced analysts do somewhat better, that predictions that receive the most intense analytical attention (key judgments) fare somewhat better, and that assessments conducted as part of an interagency process did somewhat better too. In short, for all the myriad challenges of policy-relevant prediction, it is possible to improve performance.

Moreover, even if one accepts that cascading changes can make it difficult to determine the onset, scope, and consequences of certain types of events, strategic analysts can and ought to pay more attention to identifying the sorts of environments that might give rise to rapid systemic shifts. A useful analogy would be the sort of fire risk indicators widely used in the meteorological and forestry sectors: because many of the causal events that contribute to a major forest fire are exogenous (unpredictable careless human activity) or difficult to predict (lightning strikes), analysts focus on contextual conditions that contribute to rapid ignition and spread, and issues warnings on that basis.²⁵ In the same vein, a qualitative review of Canadian IAS assessments suggested that "for some time to come there may be a particular need in Middle East assessment to flag wildcards and low probability/high impact developments that could result in rapid and substantial shifts in otherwise apparently stable political trajectories." This, the report, suggested, should not be confined merely to the "dogs that bark"—that is, evidently unstable areas that were already the focus of collection and analytical efforts—but also the "dogs that didn't bark" too. In other words, to avoid surprise it is important that one not become too distracted by current headlines and maintains some analytical

attention to “underlying medium-and long-term trends in countries without ongoing protests or civil violence.”²⁶

Analysis

Information—even if it were to be fully accurate, complete, and accessible—does not generate forecasts all by itself. Rather, it must be interpreted and analyzed. The intellectual and cognitive barriers to doing so are substantial, especially when trying to understand the sorts of complex political and social processes described above.

Two different sorts of challenges lurk here. The first concerns the adequacy and predictive value of the models that analysts consciously bring to bear on an issue. All sense-making and social prediction is essentially a functioning of modeling, whether implicit or explicit, since it fundamentally rests upon assumptions regarding what drives events and how these variables interact. Typically analysts pick up some of this from prior education, a little from formal analyst training, and a lot from experience and observation.

The degree to which they are acquainted with alternative models—for example, through interaction with scholars and others working on similar issues—can vary substantially. Although there has been marked improvement in interchange between the analytical and academic communities post-9/11, it still remains uneven, constrained by resources, available time, security classification, and organizational culture. Differences in the sorts of problems they address, the context for analysis, and the urgency and time-frames involved constrain the influence of scholars on strategic analysis too.²⁷ All too often the strategic analysis community engages such expertise simply by holding academic outreach sessions at a broad level of generality. Embedding visiting scholars within the strategic assessment community is rare indeed, especially, outside the US. This is partly because of the complications of security clearances, as well as the associated risk that academics will inadvertently utilize classified information in their public lectures or scholarly writings. It is also because it is simply not standard bureaucratic experience, and there is rarely an existing procedure for making it happen.

In my own personal experience, such interaction pays long-term dividends for both the analytic community and the scholar involved. Scholars acquire a much more nuanced sense of how information is acquired and used within the bureaucratic process, and experience which tends to heighten the value of their advice and input. Analytical units gain access to both domain knowledge and differing perspectives. The fact that scholars have careers outside the bureaucracy may also mean that they are less constrained by hierarchy, and can more safely and easily offer critical thinking that challenges current wisdom.

The second challenge to better forecasting is perhaps even more serious than this. It involves not the deliberate and thoughtful choice of analytical models, but rather the many characteristics of human cognition that impair effective analysis. These problems have been recognized within the intelligence community for decades. As Richards Heuer noted in *The Psychology of Intelligence Analysis*, human beings tend to perceive what they expect to perceive; form and adhere to implicit mental models that they are reluctant to change; and remember—imperfectly at that—what they want to know.²⁸ They often handle imperfect information poorly, have difficulty linking cause and effect, and “satisfice” by uncritically accepting the first explanation they encounter that appears to fit the facts. They also rather poorly understand their own analytical process, such that “the analyst is typically unaware not only of which variables *should* have the greatest influence, but also which variables *actually* are having the greatest influence.”²⁹

To a substantial extent, modern analytic tradecraft has been designed to acquaint analysts with their perceptual biases, structure their analyses in such a way so as to clarify their assumptions and encourage exploration of competing hypotheses, identify indicators of change so as to minimize surprise, and generally promote more critical thinking.³⁰ The CIA’s own tradecraft primer generally groups these techniques into three categories: “diagnostic techniques are primarily aimed at making analytic arguments, assumptions, or intelligence gaps more transparent; contrarian techniques explicitly challenge current thinking; and imaginative thinking techniques aim at developing new insights, different perspectives and/or develop alternative outcomes.”³¹ The continuing challenge with many of these techniques—even in cases where analysts receive sufficient training with them—is the amount of time and resources that they might take to utilize. In a busy

assessment organization, time and money may be a luxuries in increasingly short supply. Urgent analyses may be least likely to benefit from alternative analytical methods, while lower-priority analytical tasks may be least likely to benefit from the investment of material resources.

In recent years, growing work in cognitive psychology and decision science has contributed further insight into how analysis—and hence forecasting—might be further improved.³² Several key themes emerge from this, with implications not only for the process of analysis, but also for analyst recruitment and training; writing and communication of assessments; and organizational structure and process.³³

Tetlock's research on political prediction has highlighted the substantial effect that cognitive style has on predictive accuracy. Specifically, those that he terms *foxes* ("those who 'know many little things,' draw from an eclectic array of traditions, and accept ambiguity and contradiction as inevitable") fare much better than *hedgehogs* ("those who know 'one big thing,' toil devotedly within one tradition, and reach for formulaic solutions to ill-defined problems").³⁴ Research by the subsequent Good Judgment Project (supported by the Intelligence Advanced Research Projects Activity of the US Office of the Director of National Intelligence) suggests that there is considerable room for improving performance through training analysts in probabilistic reason and recognizing cognitive bias.³⁵ Analysts need to be good Bayesians too, updating their assessments in response to new or updated information.³⁶ There is also evidence that forcing analysts to change perspective—for example, by red teaming,³⁷ or by role-playing a conflict scenario (rather than simply reading about it)³⁸—improves predictive accuracy.

Second, team judgments appear to outperform judgments by single analysts. Several methods have been developed over the years to harness this effect, from Delphi polls of subject matter experts through to prediction markets wherein participants essentially bet on the outcome of specified events, thereby generating a price signal that can be used as an indicator of collective predictive judgment. Prediction markets have been shown to outperform individual predictors. However, simple team discussions may be as, or more, effective when certain aggregation techniques are used,³⁹ and when groups are established and run in ways that are most likely to elicit a productive interplay of ideas.⁴⁰ This later point—that effective group dynamics involve more than simply

assigning multiple people to a task—is something that some managers know intuitively, but is rarely formally taught as a managerial skill.

Group forecasts are also improved by having greater diversity within the group.⁴¹ While the extent of formal professional qualifications is usually a poor indicator of predictive accuracy, all expert judgment seems to improve with group discussion.⁴² Even more impressive results can be obtained by selecting participants based on past forecasting performance. Indeed, civilian “superforecasters” in the Good Judgment Project seem to even out-predict professional intelligence analysts.⁴³ Such superforecasters tend to score higher on measures on fluid intelligence, are more likely to update beliefs, have a greater appetite for challenge, energetically collect information, and have superior domain knowledge (although not necessary superior formal qualifications).⁴⁴

Accountability and feedback mechanisms have been shown to improve the performance of analysts. For the most part, however, the intelligence community “has not systematically evaluated the accuracy of its estimates”⁴⁵ and has resisted using numerical indicators of estimated probability and other indicators and processes that would facilitate such evaluation both within and across agencies.⁴⁶

Communications

The manner in which early warning or other forecasting is communicated can have substantial effect on the way it is received by policymakers, and hence on its subsequent impact on policy. Care is needed in the communication of assessments so that a reader will understand a statement to mean exactly what an analyst intends. Many organizations thus provide explicit guidance on what words of estimative probability such as “likely” or “very likely” are meant to convey. In the case of Canada’s IAS, some divisions took this a step further and required analysts to include numerical probability statements (“9/10”) alongside predictions in draft assessments. While these faced some criticism for implying a false quantitative precision, in general they were found to have helped clarify intent, facilitate discussion, and reduce some of the most common errors involved in expressing future probability.⁴⁷

Analysts need to also take care to avoid what Sherman Kent called the “lurking weasel”⁴⁸—that is, the tendency to hedge bets by using ambiguous statements, uncertain timeframes, or offering non-falsifiable predictions or other conclusions. Best practice in the intelligence community has been to discourage this. In the academic and foreign policy community, however, it is commonplace. A quick review of both scholarly journals⁴⁹ and diplomatic reporting (for example, US State Department cables published by Wikileaks)⁵⁰ reveals that identification of trends is usually so hedged in vagueness or qualifiers as to offer no unambiguous signal as to what the analyst means.

The physical format of assessments and briefings may also influence both predictive accuracy and the communication of strategic assessments to policymakers. In Canada, IAS shifted to shorter reports that could be more rapidly produced and which were more easily digested by clients. However, fears were expressed that this format—which some analysts derisively referred to as “tweeting”—could prove less effective in providing strategic warning than longer and more substantive assessments.⁵¹ The frequent use of PowerPoint in briefings for senior officials can have its problems too. Unless prepared well, slides can limit or miscommunicate information. Information retention may be adversely affected too. Indeed, the failure of a PowerPoint presentation to effectively communicate key risks was identified as a contributing factor in the 1986 Challenger disaster.⁵²

Christoph Meyer et al, writing about the onset and escalation of armed violence within states, argue that the process of early warning is best understood through the lens of persuasion. Specifically, they warn about “the tendency to concentrate only on the scientific accuracy of forecasting, while neglecting communicative, cognitive, and political dynamics that connect producers and consumers of warning.”⁵³ While the notion of strategic forecasting as “persuasion” resonates better in the atrocity-prevention (NGO and academic) community than it does within intelligence institutions committed to some degree of policy neutrality, their approach does highlight the complex interplay between analysis, communication of that analysis, and the institutional and political context within which warning (or other types of policy-relevant prediction) takes place. In a similar vein, Fingar argues that analytical support to decision-makers is most useful if there is ongoing dialogue between analysts and clients. He emphasizes, “Analysts can, should,

and do regard reminding customers of long-term trends and strategic implications of current decisions as an important part of their job, but they must do so within the parameters of trust, temporal pressures, and the agendas of those they support. The alternative is to be regarded as unhelpful or irrelevant.”⁵⁴

As the Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security has noted, several factors complicate this dialogue. One of these impediments is differences in status, which make it difficult for junior analysts to ask senior policymakers questions about what information they need and why they need it. A second problem can be that of presumed “common knowledge,” whereby analysts assume clients see the world in more similar ways than is actually the case.⁵⁵

The contemporary media environment adds a further complication. Policymakers today have easier access to more competing information sources than ever before. On the positive side, this makes them less vulnerable to bureaucratic capture due to dependence on official information. On the other hand, it means that strategic analyses must compete for attention with a welter of emails, websites, social media campaigns, and online (often flawed or partisan) media of varying quality. It thus becomes increasingly important that analysts not only communicate their best assessment of emerging trends and future possibilities, but also proactively respond to some of the other less reliable analyses to which clients may have been exposed.

Institutions

Whether attention is paid to forecasts depends not on their content (that is, whether it is timely, relevant, and communicates new informational and analysis), but also on that elusive creature known as “political will.” Will is a complex and nebulous mix of how national interests are weighed and perceived, the nature of threats, domestic political calculations, the available means and resources, and perceived prospects for successful policy engagement. For the most part it lies outside the scope of this chapter.

However, institutions also affect the quality and impact of strategic forecasting in other fundamental ways too. They can serve to empower analysts to forecast better.

Conversely, they can encourage practices, organizational cultures, and sustain institutional pathologies that weaken strategic assessment and its effects. In his important work on the UN response to early warnings of potential genocide in Rwanda, for example, Michael Barnett places considerable weight on these sorts of factors.⁵⁶ Institutional factors will determine who is charged with strategic forecasting, whom they are expected to serve, and how that task is defined. In order to better serve primary clients and so as to develop comparative advantage in key areas, different institutions may do different things in different kinds of way. Institutional factors also shape the ways in which individual analysts approach their task, whether through the impact of hiring, training, human resources management, or the structure of the assessment production process. Institutional factors will shape how information is filtered, prioritized, and communicated.

Here, discussion will focus on four organizational pathologies which tend to inhibit the effective production and dissemination of policy-relevant prediction: politicization of intelligence assessment; stove-piping and the “not written here” syndrome; the difficulties of organization learning; and the question of accountability.

Relatively little needs to be said about the problems that arise when analysts reshape or self-censor their assessment to fit with the policy preferences of their political masters—it has long been recognized as a problem. In the recent past the Bush and Blair administration in the US and UK were accused of having done so in the run up to the Iraq war, and in Canada the Harper government’s allergy to contradictory information may have created something of a chilly climate for unbiased assessment on some issues (climate change and Middle East politics come to mind) in some government agencies. Yet, as Richard Betts has noted, however, some politicization is inevitable. Analytic agencies thus need to show some political acumen if they want their analyses to be given due consideration.⁵⁷

While it is simple enough for analysts to complain about pressures from politicians and policymakers, internal institutional distortions can be just as serious. The “not written here” problem is one that arises from a combination of bureaucratic turf wars and the propensity of analytical teams to value (even over-value) in-house production over that produced by others. Agencies with implementation responsibilities are often

especially reluctant to accept outside analysis that runs counter to current or proposed policy. In the Canadian case, it is not unusual to hear officials at the Department of Foreign Affairs, Trade, and Development (DFATD) downplay the value of trend assessment by IAS, the Department of National Defence (DND), or the Canadian Security Intelligence Service (CSIS)—although Canadian diplomatic reporting, like those of many countries, is curiously short of unambiguous forecasting. Similarly, IAS, DND and CSIS analysts have been known to be jealous of their respective turf when one agency strays into predictive terrain that the other considers in its own backyard. For similar reasons, scanning and foresight work by Policy Horizons Canada appears to have relatively little impact on the national security community.⁵⁸ In this case, difference in work style and communicative language—including the latter’s rather more trendy and edgy style—probably contributes to the gap. The US National Intelligence Council’s *Global Trends* reports and the UK Ministry of Defence’s *Strategic Trends* assessments seem to be slightly better received within their respective countries, perhaps because they are produced within national security/foreign policy institutions.⁵⁹

In the Canada, the small size of the analytic community means that some of barriers to collaboration can be easily overcome through personal relationships. In the US, however, the sheer size of the Department of Defense, State Department, and CIA magnifies the problem of institutional stove-piping. This has been offset, however, by the impact of 9/11 and evident need for closer collaboration generated by war-fighting in Afghanistan, Iraq, and elsewhere. The 9/11 attacks in particular led to a series of reforms of the US intelligence community intended to enhance communication and cooperation. While it is doubtful that all of these organizational changes had the desired effect, there can be little doubt that war provides a powerful impetus for learning.

Not all learning is necessarily positive, however. The failure to detect some Iraqi WMD (weapons of mass destruction) programmes in the early 1990s contributed to inflated assessment of Baghdad’s capabilities in the run-up to the 2003 war.⁶⁰ In turn, the embarrassing failure to find any active Iraqi WMD program then led the CIA to “learn” to be more cautious again, for in its assessment of Iran’s nuclear programme.

Particularly problematic for positive learning by strategic assessment organizations is the absence of clear accountability mechanisms and metrics. The

political process often holds the intelligence community to unreasonable expectations, conveniently uses it as a scapegoat, or wields it as a tool in political and policy fights. Within agencies, there is little retrospective review of past assessments and usually no way of systematically tracking the accuracy or utility of forecast.⁶¹ Perhaps indicative of this problem, when DRDC issued an unclassified qualitative report on IAS predictive accuracy, senior officials—possibly embarrassed by criticism or suggestions for change—quickly had the report withdrawn from public circulation.⁶²

Final Thoughts

This chapter has briefly surveyed the challenges presented by policy-relevant prediction. It has argued that the evidence both from assessment agencies (such as IAS) and from growing scholarship on cognition and forecasting indicates that there are a number of ways in which performance can be improved.

If forecasting is to be improved, however, the very first step is to identify and acknowledge shortcomings when and where they exist. Doing so will require metrics whereby performance can be measured. It also requires accountability mechanisms that recognize and reward success on the one hand, and diagnose and address deficiencies on the other. To date, these still remain conspicuous by their absence in many organizations entrusted with strategic analysis.

Enhancing forecasting capabilities also requires understanding and applying the insights of recent research on the topic. Within analytical agencies, a surprising number of mid- and senior-level management personnel have only limited exposure to this. There is also a tendency to emphasize the “art” of assessment over the “science.” While there is indeed much art to the process, it is foolish to not fully engage the latter too in an effort to improve strategic forecasting. Similarly, analyst training needs to be informed by the latest evidence as to how more accurate prediction can best be fostered.

Finally, many of the most serious challenges are institutional ones. Does the format of analytical products, the assessment process, management style and institutional culture within an analytic unit provide an optimum setting for effective forecasting? How can client preferences and analytical rigor best be balanced when the two pull in

somewhat different directions? Does the institution even recognize that prediction is part of what it does? These questions need to be explicitly addressed, and not simply settled by bureaucratic inertia or the path of least resistance. Here, managers have a key role to play in thinking through the challenge of fostering appropriate incentives, structures, and behaviors.

References

¹ IAS is tasked with providing the prime minister, cabinet, and senior government clients with “original, policy-neutral assessments of foreign developments and trends that may affect Canadian interests,” as well as “coordinat[ing] assessment work that involves more than one federal department or agency” and “foster[ing] and strength[ing] strengthen relationships with allied international assessment organizations.” Privy Council Office, “Privy Council Office Secretariats,” accessed 4 August 2015, <http://www.pco-bcp.gc.ca/index.asp?lang=eng&page=secretariats>. The present author was the primary coder for the project, which was led by David Mandel (Defense Research and Development Canada) and Alan Barnes (IAS). The study reviewed more than 1,500 individual predictive judgments in IAS assessments. See David Mandel and Alan Barnes, “Accuracy of Forecasts in Strategic Intelligence,” *Proceedings of the National Academy of Science* 111 (2014), accessed August 4, 2015, doi: 10.1073/pnas.1406138111.

² Some of the literature draws a distinction between “forecasting” and “prediction,” although there is far from any consensus on what that difference might be. Herein the terms are treated as largely interchangeable.

³ Jack Davis argues “warning analysis should give primacy of place to avoidance or limitation of damage—and not to the unrealistic standard of avoidance of surprise. In other words, the ultimate goal of effective warning is to maximize damage limitation not predictive accuracy.” While he is right in highlighting the inevitability of some—perhaps many—surprises, his conception of damage limitation still hinges heavily on a predictive function, namely assessing the future plausibility or likelihood of particular strategic

developments. Jack Davis, *Strategic Warning: If Surprise is Inevitable, What Role for Analysis?*, Sherman Kent Center for Intelligence Analysis Occasional Papers 2 (2003), accessed 1 August 2015, <https://www.cia.gov/library/kent-center-occasional-papers/vol2no1.htm>. In the 1970s Cynthia Grabo noted that “Strategic warning is not a forecast of imminent attack, but rather a forecast of probable attack and it is this above all which the policy official and commander need to appreciate,”—again highlighting that even if forecasts are not predictions of specific events at a specific time, they are nonetheless probabilistic statements of possible or likely future developments. Cynthia Grabo, *Anticipating Surprise: Analysis for Strategic Warning*, Jan Goldman ed., (Washington DC: Joint Military Intelligence College, 2002), p. 118.

⁴ See, for example, Klaus Knorr, “Failures in National Intelligence Estimates: The Case of the Cuban Missiles,” *World Politics* 16 (1964); Avi Schlaim, “Failures in National Intelligence Estimates: The Case of the Yom Kippur War,” *World Politics* 28 (1976); US House of Representatives, Subcommittee on Evaluation, Permanent Select Committee on Intelligence, Staff Report, *Iran: Evaluation of U.S. Intelligence Performance Prior to November 1978* (Washington DC: US Government Printing Office, 1979); Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (New York: Cornell University Press, 2010); National Commission on Terrorist Attacks Upon the United States (9/11 Commission), *The 9/11 Commission Report* (Washington DC: US Government Printing Office, 2004), accessed 2 August 2015, <http://www.9-11commission.gov/report>.

⁵ F. Gregory Gause III, “Why Middle East Studies Missed the Arab Spring: The Myth of Authoritarian Stability,” *Foreign Affairs* 90 (2011); US Department of Defense, “DIA

Five-Year Plan Updates Strategic Warning Mission,” *DoD News*, 18 July 2012, accessed 8 August 2015, <http://www.defense.gov/news/newsarticle.aspx?id=117160> ; “Arab spring took British intelligence by surprise, report says,” *The Guardian*, 12 July 2012, accessed 8 August 2015, <http://www.theguardian.com/uk/2012/jul/12/arab-spring-british-intelligence-report> .

⁶ Stefanie Levitz, “Arab Spring Caught Canada By Surprise: Government Report,” *Huffington Post*, 6 May 2013 (from Canadian Press), August 4, 2015, http://www.huffingtonpost.ca/2013/05/06/arab-spring-canada-government-report_n_3224719.html

⁷ David Mandel notes that at one workshop he attended, an intelligence professional objected to efforts to assess predictive accuracy on the grounds that “intelligence is not in the business of making predictions.” David Mandel, Alan Barnes, and Karen Richards, *A Quantitative Assessment of the Quality of Strategic Forecasts*, Technical Report TR2013-036 (Toronto: Defence Research and Development Canada, March 2013), p. 3.

⁸ Of course, much strategic surprise arises precisely because something was thought to be unlikely or impossible. Given that, there is also considerable value in alternative futures exercises and similar techniques that call upon planners to strategize against the most harmful, and not necessarily the most likely, threat.

⁹ Richard Betts, “Analysis, War, and Decision: Why Intelligence Failures are Inevitable,” *World Politics* 31 (1978).

¹⁰ For a discussion, see James Bruce, “Making Analysis More Reliable: Why Epistemology Matters to Intelligence,” in Roger George and James Bruce, eds.,

Analyzing Intelligence: Origins, Obstacles, and Innovations (Washington DC: Georgetown University Press, 2008).

¹¹ Timur Kuran, “The East European Revolution of 1989: Is It Surprising that We Were Surprised?” *American Economic Review* 81(1991). Preference falsification further amplifies the effects of informational cascades, creating the potential for what Kuran describes as “revolutionary bandwagoning.”

¹² For information on deaths and human rights abuses, see the Syrian Observatory for Human Rights website at <http://www.syriaahr.com>; for open-source information and analysis of Syrian CW usage see *Bellingcat* at <https://www.bellingcat.com/category/news/mena> and the earlier *Brown Moses* blog at <http://brown-moses.blogspot.ca>.

¹³ See, for example, J.M. Berger and Jonathon Morgan, *The ISIS Twitter Census: Defining and Describing the population of ISIS supporters on Twitter* (Washington DC: Brookings Institution, 2015), as well as the work of the International Centre for the Study of radicalization at King’s College London, <http://icsr.info>.

¹⁴ See, for example, *The Arkenstone* (a blog devoted to opensource IMINT and technical analysis of the Iranian military) at <http://thearkenstone.blogspot.com>; or the Sentinel Project (which uses crowdsourcing and OS imagery for the purposes of mass atrocity prevention) at <https://thesentinelproject.org>.

¹⁵ Thomas Fingar, “Analysis in the U.S.Intelligence Community: Missions, Masters, and Methods,” in Baruch Fischhoff and Cherie Chauvin, eds, *Intelligence Analysis: Behavioral and Social Scientific Foundations*, Committee on Behavioural and Social

Science Research to Improve Intelligence Analysis (Washington DC: National Academies Press, 2011), p. 11.

¹⁶ Alain Auger, “Identification of Emerging Scientific and Technical Trends Using Scientometric Analysis,” in NATO Science and Technology Organization, *Horizon Scanning and Strategic Futures Analysis*, STO Educational Notes EN-IST-135, June 2015.

¹⁷ On the challenges of more effective use of OSINT, see comments by Doug Naquin (Director of the Open Source Center), *Central Intelligence Agency Retirees Association Newsletter* 4 (2007), accessed 10 August 2015, <http://fas.org/irp/eprint/naquin.pdf>

¹⁸ Bess Puvathingal and Donald Hantula, “Revisiting the Psychology of Intelligence Analysis: From Rational Actors to Adaptive Thinkers,” *American Psychologist* 67 (2012).

¹⁹ Nassim Nicholas Taleb, *The Black Swan: the Impact of the Highly Improbable* (New York: Random House, 2010), p. xxii.

²⁰ Charles Kurzman, *The Unthinkable Revolution in Iran* (Cambridge: Harvard University Press, 2005).

²¹ Merouan Mekouar, “No Political Agents, No Diffusion: Evidence from North Africa,” *International Studies Review* 16 (2014). See also Jeff Goodwin, “Why We Were Surprised (Again) by the Arab Spring,” *Swiss Political Science Review* 17 (2011); and Rex Brynen, Pete Moore, Bassel Salloukh, and Marie-Joëlle Zahar, *Beyond the Arab Spring: Authoritarianism and Democratization in the Arab World* (Boulder: Lynne Rienner Publishers, 2012), pp. 3, 109-112.

²² Philip Tetlock, *Expert Political Judgment* (Princeton: Princeton University Press, 2005), pp. 49-54. For more general discussion of the analytical challenges of prediction see Dan Gardner, *Future Babble: Why Expert Predictions Fail—and Why We Believe Them Anyway* (Toronto: McClelland & Stewart, 2010). For a popular summary of the findings of the Good Judgment Project, see Philip Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction* (New York: Crown, 2015).

²³ “Accurate” here indicates that the predicted even occurred, even if different in significant (but not fundamental) ways from what was originally predicted. One major methodological difference between the DRDC/IAS study on the one hand and the Tetlock and Good Judgment Projects on the other arises from the source of predictions. In the former, many predictions were self-identified and selected by analysts. In the latter, predictions were externally provided as questions (“Will country X have a revolution in the next year?”) to participants in the experiment.

²⁴ Mandel and Barnes, “Accuracy of Forecasts in Strategic Intelligence,” p. 10985. See also Mandel, Barnes, and Richards, *A Quantitative Assessment of the Quality of Strategic Forecasts*.

²⁵ See, for example, the work of the Predictive Services Program of the (US) National Interagency Fire Center at <http://www.predictiveservices.nifc.gov/predictive.htm>. I am grateful to Kedra Hildebrand for suggesting the analogy.

²⁶ Levitz, “Arab Spring Caught Canada By Surprise: Government Report.”

²⁷ See comments by academic-turned-CIA-analyst-turned academic Paul Pillar, “The Battle of the Bridge,” *The National Interest*, 28 December 2010, accessed August 12, 2015, <http://nationalinterest.org/blog/paul-pillar/the-battle-the-bridge-4641>. See also the

various contributions on academics and the national security community in *Perspectives on Politics* 8 (December 2010).

²⁸ Richards Heuer, Jr., *The Psychology of Intelligence Analysis* (Washington DC: Central Intelligence Agency, 1999), accessed August 9, 2015, <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/PsychofIntelNew.pdf>

²⁹ Heuer, *The Psychology of Intelligence Analysis*, p. 56

³⁰ See, for example, the analytic techniques discussed in Grabo, *Anticipating Surprise: Analysis for Strategic Warning*; Central Intelligence Agency, *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington DC: Central Intelligence Agency, 2009), accessed August 10, 2015, <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Tradecraft%20Primer-apr09.pdf>

³¹ CIA, *A Tradecraft Primer*, p. 5.

³² In particular, see Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security, *Intelligence Analysis for Tomorrow: Advances from Behavioural and Social Sciences* (Washington DC: National Academics Press, 2011); and Fischhoff and Chauvin, eds, *Intelligence Analysis: Behavioral and Social Scientific Foundations*. See also Mandeep Dhimi et al, “Improving Intelligence Analysis with Decision Science,” *Perspectives on Psychological Science* (forthcoming 2015), accessed August 16, 2015, http://www.researchgate.net/publication/278966670_Improving_Intelligence_Analysis_with_Decision_Science

³³ Robert Sinclair, *Thinking and Writing: Cognitive Science and Intelligence Analysis*, revised ed. (Washington DC: Center for the Study of Intelligence, 2010), pp. 26-27.

³⁴ Tetlock, *Expert Political Judgment*, p. 2. Daniel Drezner has suggested that most academic (international relations) scholars are “hedgehogs,” which may hamper their ability to generate accurate predictions. Daniel Drezner, “Hedgehogs and Foxes,” *The National Interest*, April 22, 2009, accessed August 15, 2015, <http://www.nationalinterest.org/article/hedgehogs-and-foxes-3104>. For a dissenting view, see Henry Farrell, “Hedgehogs and Foxes,” *The Monkey Cage* blog, April 22, 2009, accessed 15 August 2015, http://themonkeycage.org/2009/04/hedgehogs_and_foxes/.

³⁵ *Good Judgment Project*, accessed August 8, 2015, <http://goodjudgment.com>. See also Barbara Mellers et al, “Psychological Strategies for Winning a Geopolitical Forecasting Tournament,” *Psychological Science* 25 (2014); Barbara Mellers et al, “The Psychology of Intelligence Analysis: Drivers of Prediction Accuracy in World Politics,” *Journal of Experimental Psychology: Applied* 21 (2015).

³⁶ So named after British statistician Thomas Bayes (1701-1761). Bayes theorem describes how new data should update beliefs. Experiments with intelligence analysts suggest that training in Bayesian methods improves assessment skills. David Mandel, “Instruction in information structuring improves Bayesian judgment in intelligence analysts,” *Frontiers in Psychology* (6) 2015. For a broader and more popular discussion, see Nate Silver, *The Signal and the Noise: Why So Many Predictions Fail—but Some Don’t* (New York: Penguin Press, 2012).

³⁷ Read-teaming typically involves having an alternative analytical group challenge an the assumptions, priorities, and plans of an organization or analyst. For an overview, see UK Ministry of Defence, *Red Teaming Guide* 2nd ed. (Shrivenham: Development, Concepts and Doctrine Centre , 2012), and *Red Team Journal*, at <http://redteamjournal.com>.

³⁸ Kesten Green, “Forecasting Decisions in Conflict Situations: A Comparison of Game Theory, Role-Playing, and Unaided Judgement,” *International Journal of Forecasting* 18, (2002); Kesten Green, “Game Theory, Simulated Interaction, and Unaided Judgement for Forecasting Decisions in Conflicts: Further Evidence,” *International Journal of Forecasting* 21 (2005). The interactive aspect of role “play” seems to be particularly important—experiments in role “thinking” (that is, putting yourself in someone else’s shoes) suggests it is markedly less effective. Kesten Green and J. Scott Armstrong, “Role thinking: Standing in other people’s shoes to forecast decisions in conflicts,” *International Journal of Forecasting* 27 (2011).

³⁹ Lyle Ungar et al, “The Good Judgment Project: A large-scale test of different methods of combining expert predictions,” Association for the Advancement of Artificial Intelligence Technical Report FS-12-06 (2012), accessed August 15, 2015, <https://www.aaai.org/ocs/index.php/FSS/FSS12/paper/viewFile/5570/5871>.

⁴⁰ Reid Hastie, “Group Processes in Intelligence Analysis,” in Fischhoff and Chauvin, eds, *Intelligence Analysis: Behavioral and Social Scientific Foundations*.

⁴¹ Ilan Yaniv, “Group diversity and decision quality: Amplification and attenuation of the framing effect,” *International Journal of Forecasting* 27 (2011). See also Scott Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and*

Societies (Princeton, NJ: Princeton University Press, 2007), whose “diversity prediction theorem” suggests that diverse groups will outperform single experts or homogenous groups.

⁴² Mark Brugman et al, “Expert Status and Performance,” *PLOS One*, July 29, 2011, accessed 15 August 2015, doi: 10.1371/journal.pone.0022998

⁴³ “So You Think You're Smarter Than A CIA Agent,” *NPR Parallels*, April 2, 2014, accessed August 15, 2015, <http://www.npr.org/sections/parallels/2014/04/02/297839429/-so-you-think-youre-smarter-than-a-cia-agent>.

⁴⁴ Barbara Mellers et al, “Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions,” *Perspectives on Psychological Science* 10 (2015).

⁴⁵ Jeffrey Friedman and Richard Zeckhauser, “Why Assessing Estimative Accuracy is Feasible and Desirable,” *Intelligence and National Security*, published online November 24, 2014, accessed August 15, 2015, doi: 10.1080/02684527.2014.980534.

⁴⁶ Philip Tetlock and Barbara Mellers, “Intelligent Management of Intelligence Agencies,” *American Psychologist* 66 (2011).

⁴⁷ Alan Barnes, “Making Intelligence Analysis More Intelligent: Using Numeric Probabilities,” *Intelligence and National Security*, published online January 13, 2015, accessed 15 August 2015, doi: 10.1080/02684527.2014.994955. See also Joab Rosenberg, “The Interpretation of Probability in Intelligence Estimation and Strategic Assessment,” *Intelligence and National Security* 23 (2008).

⁴⁸ Sherman Kent, “Words of Estimative Probability,” *Studies in Intelligence* (1964), Central Intelligence Agency website, accessed August 15, 2015,

<https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/sherman-kent-and-the-board-of-national-estimates-collected-essays/6words.html>

⁴⁹ Gerald Schneider, Nils Petter Gleditsch, and Sabine Carry note the general doubts many international relations scholars have about forecasting events, although they are more optimistic about the capacity of scholars to do so. “Forecasting in International Relations: One Quest, Three Approaches,” *Conflict Management and Peace Science* 28 (2011). Bruce Bueno De Mesquita is perhaps the contemporary political scientist most optimistic about forecasting events, in his case using a game theory/rational choice framework that incorporates expert assessments of the preferences and policy influence of key decision-makers and institutions. *The Predictioneers Game: Using the Logic of Brazen Self-Interest to See and Shape the Future* (New York: Random House, 2010).

⁵⁰ US diplomatic cables from Tunis before the 2011 overthrow of President Ali, for example, highlighted the growth of pervasive corruption at length. Other than a single observation that “those at the top...[are] likely to remain in power,” however, the report offered no assessment whatsoever of the future impact or consequences of the problem. Indeed, even the term “likely” (which usually conveys around a 25% chance of something else happening instead) seems to have been used in order to hedge bets rather than offer any sort of political forecast. US Embassy Tunis, *Corruption in Tunisia: What's Yours is Mine*, Tunis 000679 (Top Secret), 23 June 2008, accessed August 15, 2015, <https://wikileaks.ch/cable/2008/06/08TUNIS679.html>. Similarly, a report on the future of Tunisia and Tunisian-US relations offered vague assessments such as “in a region in turmoil, Tunisia has better prospects than most even though it is troubled.” US

Embassy Tunis, *Troubled Tunisia: What Should We Do?* Tunis 00492, 17 July 2009, accessed August 15, 2015, <https://wikileaks.ch/cable/2009/07/09TUNIS492.html>.

⁵¹ Levitz, “Arab Spring Caught Canada By Surprise: Government Report.”

⁵² “PowerPoint Makes You Dumb,” *New York Times*, 14 December 2003; Edward Tufte, *The Cognitive Style of Powerpoint 2nd* ed. (Cheshire CT: Graphics Press, 2006).

⁵³ Christoph Meyer et al, Recasting the Warning-Response Problem: Persuasion and Preventive Policy,” *International Studies Review* 12 (2010), p. 573.

⁵⁴ Fingar, “Analysis in the US Intelligence Community,” p. 13.

⁵⁵ Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security, *Intelligence Analysis for Tomorrow*, p. 76.

⁵⁶ Michael Barnett, *Eyewitness to Genocide: The United Nations and Rwanda* (Ithaca, NY: Cornell University Press, 2002).

⁵⁷ Richard K. Betts, “The Politicization of Intelligence: Costs and Benefits,” in Richard K. Betts and Thomas Mahnken, eds., *Paradoxes of Strategic Intelligence: Essays in Honor of Michael I. Handel* (London: Frank Cass, 2003). In its most recent intelligence reform, Israel has actually encouraged its Directorate of Military Intelligence (AMAN) to make policy recommendations. This, however, likely reflects the peculiar nature of civil-military relations within Israel, which are unlike those in most Western countries. Yosef Kupperwasser, *Lessons From Israel’s Intelligence Reforms* (Washington DC: Brookings Institution, 2007), p. 16.

⁵⁸ Policy Horizons Canada was established “an organization within the federal public service that conducts strategic foresight on cross-cutting issues that informs public servants today about the possible public policy implications over the next 10-15 years.”

“Who We Are,” *Policy Horizons Canada*, accessed August 17 2015,

<http://www.horizons.gc.ca/eng/content/who-we-are>.

⁵⁹ Office of the Director of National Intelligence, “National Intelligence Council: Global Trends,” accessed August 20, 2015,

<http://www.dni.gov/index.php/about/organization/national-intelligence-council-global-trends>; and UK Ministry of Defence, “DCDC Strategic Trends Programme,” accessed August 20, 2015, <https://www.gov.uk/government/collections/strategic-trends-programme>

⁶⁰ Michael Morrell and Bill Harlow, *The Great War of Our Time: The CIA's Fight Against Terrorism from al Qa'ida to ISIS* (New York: Twelve, 2015), p. 101. Morell is a former Deputy and Acting Director of the CIA.

⁶¹ Friedman and Zeckhauser, “Why Assessing Estimative Accuracy is Feasible and Desirable;” Tetlock and Mellers, “Intelligent Management of Intelligence Agencies.”

⁶² Levitz, “Arab Spring Caught Canada By Surprise: Government Report.”