Good Judgment in International Politics:
Three Psychological Perspectives

Philip E. Tetlock

This article explores three distinct psychological perspectives on good judgment in international politics. Advocates of one perspective (skeptics) argue that good judgment is extremely rare. The international political scene is so unpredictable that most people are inevitably wrong most of the time. Moreover, the skeptics argue that when prognosticators are right, their “hits” are largely the product of either luck or prudent use of probability theory. Advocates of the second perspective (complexifiers) argue that good judgment is closely linked to the ability to transcend common cognitive biases and errors that vitiate most intuitive predictions. Prognosticators who think in self-critical and integratively complex ways are less likely to accept facile historical analogies, to jump to conclusions from fragmentary evidence, to persevere with first impressions despite contradictory evidence, and to fall prey to overconfidence and certainty-of-hindsight effects. Advocates of the third perspective (simplifiers or fundamentalists) argue that good judgment is closely linked to the ability to focus on a few basic processes (rational responses to economic, technological, and geopolitical trends) and to ignore short-term distractions. Drawing on the results of a preliminary study of experts’ forecasts in three policy domains (the future of American-Soviet relations and of the Soviet Union itself, South Africa, and the Gulf crisis), I argue that each perspective has components of validity and that the research challenge is to identify the conditions under which each is most useful.

KEY WORDS: forecasting; good judgment; indeterminacy; cognitive bias; expertise; integrative complexity; Soviet Union; South Africa; Gulf crisis.

INTRODUCTION

It is hard to achieve consensus on what constitutes good judgment in international politics. The concept means radically different things to different observ-

1Institute of Personality and Social Research, 2150 Kittridge Street, Room 26, Oxford Court, University of California, Berkeley, California 94720.

517
0162-895X/92/0003-051/750}$ 50.01 © 1992 International Society of Political Psychology
ers. Protectionists versus free traders, realists versus idealists, deterrenrs versus reassurers—each camp has its own distinctive perspective on the forces that drive events and the policies that will most effectively promote peace and prosperity.Attributing good judgment to others is often just another way of saying that one agrees with them.

There would be little point in writing about good judgment if we were merely to rediscover the same partisan and theoretical cleavages that already dominate the intellectual landscape. I approach the problem therefore from a different angle—that of cognitive and social psychology. In this article, I am interested less in the validity of overarching theoretical perspectives on international politics that I am in the validity of the forecasts (conditional or unconditional) that individual observers of the political scene advance. I divide the article into three parts. In the first section, I explain why current theory sheds disappointingly little light on the nature of good judgment. The problem lies, in part, with the multiplicity of theories on which one can draw. Each suggests a quite different conception of good judgment. The problem also lies with the vagueness and openendedness of the major theoretical positions. Each is riddled with escape clauses that protect it from empirical refutation. Unlike more scientifically developed or paradigmatic disciplines (medical diagnosis, meteorology, econometrics), there are no clearly specified rules for generating forecasts from theories. Our theories give us either conflicting or ambiguous advice.

In the second section, I consider three psychological perspectives on good judgment: the skeptics, the complexifiers and the fundamentalists. Skeptics doubt the very existence of the phenomenon. In their view, good judgment is largely, perhaps entirely, the product of good luck. Complexifiers believe not only that good judgment exists but that it has systematic cognitive stylistic correlates. Observers who think in more multidimensional self-critical ways should be less prone to common cognitive biases and, as a result, are more accurate in their forecasts. Finally, the fundamentalists adopt what is in key respects the mirror image position. Good judgment is the ability to identify the small number of truly diagnostic cues in international politics and to ignore the plethora of distractor variables in which these diagnostic cues are embedded.

In the third section, I describe a small ongoing study of good judgment. Over the last eight years, I have asked expert observers of the international scene to make predictions and then to explain them. The results to date offer something for everybody. Skeptics can note that, taken as a group, the experts performed no better than one would expect from chance; complexifiers can point to the positive (albeit modest) correlation between multidimensional patterns of thinking and accuracy of forecasts; and fundamentalists can point to the long-term accuracy of certain key predictions that rested on either macro-economic necessity or Realpolitik logic. I close with a plea for more systematic study of good judgment.

LIMITS OF CURRENT THEORY

It is easy, but not particularly helpful, to poke fun at the predictive power of current theoretical approaches to international relations and security issues. All one needs to do is to measure the current crop of theories against a sufficiently high epistemological standard. For example, our theories fell far short of the standards of the covering-law model of explanation (Hempel, 1965). According to this model, it should be possible, in principle, to plug in one's mathematical model with its well-specified laws, to insert the necessary information on antecedent conditions (the data on which the laws will operate), and then to sit back as the hypothetico-deductive engine of inference cranks out forecasts that are as accurate as measurement error permits. Good judgment is the product of good theory in this view. The predictive engine stalls however when it comes to international affairs—and for a number of instructive reasons.

In part, the reasons lie outside the control of theorists. Even if we had elegant, comprehensive theories that specified the exact causal relationships among key variables, we still could not predict anything of consequence as long as we lacked reliable information on current world conditions (the "antecedent conditions" on which Hempelian laws must operate). This obstacle to prediction holds across the spectrum of levels of analysis. One cannot use personality theories to predict the behavior of individual leaders if one has no means of assessing the relevant personality variables; one cannot predict the stability of an international system if there is enormous uncertainty about the economic and military strength of key national actors within that system and the rate at which relative strengths are changing.

Prediction will also fail if the political environment is inherently indeterminate—again, hardly the fault of the theorists. The butterfly effect may be alive and well not only in meteorology but also in international affairs. Small causes do not always have small effects; sometimes, they are amplified through the system. If the flapping of a butterfly's wings in Brazil can cause a tornado in Iowa (at least in computer simulations of the weather—Gleick, 1987), perhaps the steadiness of an assassin's hand can determine the fate of a nation or a silly tactical oversight the outcome of a war.

The theorists cannot however completely escape blame. Current theories are riddled with escape clauses and are virtually impossible to pin down empirically. In a recent review, for example, Gaddis (1992) distinguishes three major approaches to forecasting international politics—the behavioralists, the structuralists, and the evolutionists—each of which has given birth to a host of theories and empirical generalizations. He notes that the vast majority of these theories and generalizations failed to yield specific (still less accurate) predictions concerning the end of the U.S.-Soviet Cold War. The behavioralists could not agree.
on the appropriate generalizations (e.g., sometimes arms races lead to war, sometimes not, and we cannot specify when—Downs, 1990); the structuralists could not agree on the consequences of bipolarity versus multipolarity (Levy, 1989); the evolutionists could not agree on the length of hegemonic cycles or on whether nuclear weapons had broken the historical pattern of recurring hegemon-challenger wars (Levy, 1989).

Although internal disagreements within theoretical camps are understandable, they make it well-nigh impossible to assess the predictive utility of competing approaches. Three examples illustrate the conceptual stand-offs that emerge even when the political world gives us seemingly decisive feedback.

Deterrence Versus Reassurance

Deterrence theorists and conflict spiral theorists held starkly different views of how to manage the U.S.-Soviet relationship. Although neither camp was well-prepared for the abrupt end of the Cold War, both camps were well-positioned to offer post hoc explanations (Tetlock, McGuire, & Mitchell, 1991). Advocates of deterrence argued that Western resolve and strength precipitated liberalization within the Soviet Union and Eastern Europe. (Some claimed that Kennan’s containment prophecy of 1947 had finally been vindicated; others argued that the “technological trump card” of the Strategic Defense Initiative convinced the Soviets of the futility of continued competition with the West.) But these theorists were equally well-prepared to predict a neo-Stalinist retreatment in the East bloc (what other options did the beleaguered Communist leaders have than to crack down, reinforce internal discipline, and pursue confrontational policies with the West?). In fact, some advocates of deterrence thought Gorbachev was a “closet hardliner” right up to the attempted military coup of August, 1991.

Advocates of the more dovish persuasion are not in a stronger epistemological position. They traced the radical transformations in the East bloc to internal developments within those polities. They argued that the new thinking of Gorbachev and his allies arose despite of, not because of, the recent American defense build-up. Conflict spiral theorists, however, hardly predicted such an outcome. Indeed, many warned us that the Reagan build-up would greatly increase the influence of Kremlin hardliners (White, 1984).

Rational Actor Versus Psychodynamics

Advocates of rational actor and psychodynamic perspectives tell radically different stories about political decision-making. It turns out, however, that they often offer difficult-to-distinguish advice. For instance, many rational actor theorists expected Saddam Hussein to withdraw from most of Kuwait because he

recognized that the balance of power had tipped hopelessly against him and that there was still a possibility of holding on to part of Kuwait if he could fracture the American-led alliance. By contrast, another (albeit smaller) camp of rational actor theorists expected him not to back down—either because he had reasonable grounds to believe the U.S. was bluffing or because he feared the domestic political consequences in Iraq of appearing weak or because he believed that even a military defeat would enhance his prestige and power (a Nasser effect). Psychodynamic theorists (who emphasized his “malignant narcissism,” his “rigid cognitive style,” or “groupthink” in his inner circle) advanced an equally wide range of expectations. Some predicted a cunning retreat in the form “aggressive withdrawal” (a skill acquired in Ba’ath Party maneuvering and the war with Iran); others predicted a reckless attempt to embroil the entire region in war (“better to be a rooster for a day than a chicken for a year,” Saddam reportedly said). Again the theorists in each camp were well-positioned to explain the full spectrum of outcomes.

Boomsters Versus Doomsters

Free market economists (“boomsters”) and ecologists who see growing human needs and dwindling global resources (“doomsters”) have starkly different perspectives on the prospects for long-term prosperity. When one theorist from each camp (Julian Simon versus Paul Ehrlich) agreed to a wager in 1980 on the price of five finite-resource commodities (e.g., copper, tin, tungsten) a decade hence, the confrontation between viewpoints was laid out with admirable clarity (Tierney, 1990). The boomster expected real prices to decline, the doomster expected them to rise. Real prices declined (in fact, even most nominal prices fell) and Julian Simon won the wager. But such a contest made little impression on the opposing camp. Ehrlich compared Simon to a man who jumps from the top of the Empire State Building and as he passes the 10th floor euphorically proclaims “All’s well so far!” It is unclear, moreover, on strictly logical grounds whether anyone should have changed his mind. Ehrlich can argue that the choice of resources or time frame was wrong for testing the pending positions. And Simon could have responded with his own set of theoretical defenses had he lost (cost-effective technological substitutes have yet to be found, but soon will be; unusual macro-economic conditions prevailed). Once again, we find ourselves at a conceptual impasse, this time between laissez-faire optimism and Malthusian pessimism.

What conclusions should we draw from these examples? I draw two major lessons. First, although current approaches to international security usually fail the falsifiability test that Popper (1972) upheld as the sine qua non of scientific theory, we should not conclude that the theories are useless. Most philosophers
of science accept the Duhem-Quine thesis that theory-testing is difficult even in controlled experimental settings where it invariably occurs in the context of background assumptions about materials, instrumentation, and procedures (Suppe, 1977). It follows that the theory-testing is all the more problematic in complex political and economic systems in which a multitude of uncontrolled variables are operating. One can be right for the wrong reasons (as the loser Ehrlich claimed of the winner Simon) or wrong for the right reasons (the theory is fundamentally correct but was misapplied). Many of the post hoc corrections that theorists advance may be quite justified, no matter how suspicious they look. Accordingly, one should be exceedingly cautious in accepting or rejecting a theory based on a few forecasting hits or misses. Theoretical arguments are rarely resolved by decisive empirical knock-outs (cf. Tetlock & Levi, 1982; Tetlock & Manstead, 1985).

A second lesson is that forecasting is currently more art than science. Forecasting is not a matter of deducing a consequence in hypothetico-deductive fashion from a major theoretical premise (principle x is valid) and a minor empirical premise (the antecedent conditions for activating principle x are present in the real world). International politics is simply too complex. There are too many potential processes at work and there is too much uncertainty about how these processes interact.

In short, current theory gives us little guidance into the nature of good judgment. The theories are too fragmented and open-ended to generate falsifiable forecasts. As a result, we know virtually nothing about the correlates of good forecasting judgment. How much weight do more or less successful forecasters place on different levels of analysis in making judgments about the future? Do they draw on psychological theories of leadership, careful monitoring of domestic political events and economic trends, macro-economic theory, neorealism? Do they rely on complex, configural judgment strategies or do they focus on a surprisingly small number of underlying cues? In the remainder of this paper, I advance a number of hypotheses about good judgment and report an ongoing study to test some of these hypotheses. It is important to bear in mind, however, the limits of this exercise. We should harbor no illusions that the best forecasters will necessarily be the best theorists or vice versa. Forecasters (like theorists) can be right for the wrong reasons or wrong for the right reasons. We should not enshrine the thought protocols that lie behind successful prognostications. I make only the following limited claim: careful analysis of the rationales behind more or less successful forecasts will advance our understanding of what we often loosely call good judgment in the domain of international politics.

Skeptics, Complexifiers, and Fundamentalists

The most fundamental question is, of course, whether good judgment even exists. To be sure, some people are occasionally right about some things. But the skeptics argue that such hits are best attributed to luck. If one concludes that the skeptics are wrong (or, more modestly, not completely right), it is reasonable to explore two other major schools of thought on good judgment—the complexifiers and the fundamentalists.

The Skeptics

Psychometricians remind us that some outcomes in life are just not very predictable and that no amount of qualitative or quantitative analysis will improve the accuracy of our forecasts beyond a certain point (Armstrong, 1977). For example, in predicting occupational choices of students or recidivism of prisoners, it has proven exceptionally difficult to account for more than 15% or 20% of the variance. Moreover, one can often attain this degree of prediction with relatively little effort. One doesn’t need highly trained clinical staff to conduct in-depth interviews. Clinicians don’t do any better and sometimes they even do worse than minimal baseline efforts (Einhorn & Hogarth, 1981).

Chaos theorists have also recently reminded us of the limits on predictability (Gleick, 1987). They note that a system such as an organism could be either muted or amplified through complex causal systems. The most widely cited example is the “butterfly effect” in which a tiny data entry error in a computer model of global climate led to major perturbations in the unfolding of simulated storms (an effect analogous in magnitude to the flapping of a butterfly’s wings in Beijing causing a tornado in Iowa). Moreover, there may be a pervasive tendency for large, causally interactive systems (plate tectonics, ecosystems, financial markets) to evolve toward “critical states” in which minor events are sufficient to trigger catastrophes (Bak & Chen, 1991). If so, this tendency bodes ill for efforts to predict the behavior of all such systems.

Applied to international politics, the skeptics’ position has several counterintuitive implications (all likely to be unpalatable to the communities of relevant experts). One implication is that the search for systematic correlates of good
judgment will yield little or nothing. The yield will be meager because only a small handful of participants will even come close to making consistently accurate forecasts and most of these “hits” can be most parsimoniously attributed to chance. The political environment is simply too indeterminate for good judgment to be a common or reliable phenomenon. Forecasting accuracy will often not exceed levels that one would expect from the operation of chance alone.

A second implication is that we should expect few differences in forecasting accuracy between seasoned experts and relative novices. Previous psychological studies of expert judgment have revealed the diminishing marginal predictive utility of knowledge in such diverse fields as criminology, meteorology, radiology, and technical analysis of stock market trends (Kahneman, Slovic & Tversky, 1982). Widely cited political experts may do little better than advanced undergraduates or attentive readers of the New York Times in anticipating whether crises will be resolved peacefully, command economies will successfully navigate the perilous transition to free markets, or multi-ethnic states will move toward dictatorship or devolution. Beyond a certain minimum level of knowledge, it is extraordinarily difficult to translate new insights into increments in forecasting accuracy.

A third implication of the skeptic’s position is that insofar as some experts do consistently outperform others in their prognostications, their superior performance should be attributed less to political insight and more to perceptive applications of probability theory. Consider, for instance, the problem of predicting which of 50 current world leaders will be in power on January 1, 1997. The skeptics would treat this problem as formally analogous to placing bets at a roulette table. The most effective strategy is to compute the relevant “base rates” (admittedly often difficult in international politics—Jervis, 1990): What is the average (or median) length of time in office for national leaders? In the absence of any other information (about correlates of political longevity or about the costs of making different predictive errors), the rational strategy is to predict the most common outcome across the board.

In sum, the skeptics’ position is that sophisticated political knowledge is neither a necessary nor a sufficient condition for accurate forecasting. Indeed, under some conditions, knowledge may be a handicap (leading experts to search for nonexistent ways of improving upon simple base rate or multiple regression predictions). The remaining two psychological perspectives on good judgment are not so dismissive, however, of expertise and domain-specific knowledge. Both positions lead us to expect systematic relationships between styles of reasoning and forecasting accuracy.

The Search for Wisdom: Advocates of Complexity

Experts may fail to make accurate forecasts not because they have been asked to do the impossible (predicting the unpredictable), but because they are thinking about the problem in the wrong way. A major theme of recent work in social-political cognition is that people are creatures of bounded rationality—“cognitive misers” whose penchant for simple, easy-to-execute heuristics renders them vulnerable to an assortment of errors and biases (Fiske & Taylor, 1991). People are too quick to rush to judgment and too slow to revise their beliefs in response to new evidence. One example is the use of history (Jervis, 1976; Neustadt & May, 1986). Khong (1991), for instance, makes a strong case that experts focused too exclusively on the similarities between the Vietnamese and Korean conflicts, ignoring key differences. Tetlock et al. (1991) note the variety of post-Vietnam conflicts that experts (writers on the New York Times Op-Ed page) have compared to Vietnam, including Afghanistan (the Soviet’s Vietnam), Chad (Libya’s Vietnam), Eritrea (Ethiopia’s Vietnam), Lebanon (Israel’s Vietnam), Angola (Cuba’s Vietnam), El Salvador (a new American Vietnam) and Kampuchea (Vietnam’s Vietnam). Each analogy contained implicit or explicit political predictions and policy prescriptions. It is unlikely that the Vietnam analogy fits each case equally well.

The tendency to rush to historical judgment would be less serious were it not accompanied by an equally powerful tendency to be slow to change one’s mind (Nisbett & Ross, 1980; Tetlock & McGuire, 1986). As a result, observers are continually at risk of falling out of phase with geopolitical reality. Many observers of the Soviet political scene who were successful in the stagnant Brezhnev era had considerable difficulty in adjusting to the dramatic changes of the Gorbachev period (Tetlock, 1991). In the next decade, a similar fate may await those who scored quite spectacular hits in the late 1980s. We should recall that the same Winston Churchill who, as early as 1933, was uncannily correct about Hitler’s aspirations attributed almost equally malign motives to Gandhi and his independence movement (Gilbert, 1990). In short, good judgment may sometimes be the product of fortuitous coincidences of slowly changing preconceptions in a rapidly changing world.

Complexifiers argue, however, that good judgment need not always be fortuitous. The literature on cognitive biases and errors captures only modal
tendencies, not all of human nature. Thoughtful human beings do exist, and their
thoughtfulness gives them some advantage in anticipating what will happen next.
Fortunately, research on social and political cognition not only highlights short-
comings in how people judge and anticipate events; it also points to correctives.
For example, consider the following frequently mentioned shortcomings in our
intuitive judgment strategies:

1. People seek "premature closure" in judging new events. They try to force
new events into preexisting cognitive templates even when these events do not fit.
For most new events, there is not just one historical precedent, there are
many, each of which captures some aspects of the phenomenon but fails to
capture others (Vertzberger, 1990).

2. People are "cognitive conservatives" who resist revising preconceptions
even in the face of mounting disconfirming evidence (Fiske & Taylor, 1991;

3. People are often overconfident in the correctness of their judgments. On
average, when intelligence analysts claim to be 90% confident, they are correct
roughly 75% of the time (Fischhoff, 1982; Tetlock & Kim, 1987).

4. People have much lower thresholds for "seeing" evidence that supports
as opposed to challenges their pet hypotheses (Nisbett & Ross, 1980).

Work on "debiasing" judgment has revealed a simple but surprisingly potent
corrective to these cognitive flaws. The key is "preemptive self-criticism"
(Tetlock, 1985, 1991). People who are explicitly encouraged to think of reasons
why they might be wrong are less likely to persist with initially incorrect impres-
sions of a problem and more likely to report realistic degrees of confidence in
their predictions (Fiske & Taylor, 1991; Tetlock & Kim, 1987). Applied to
international politics, this argument suggests that good judgment will be strongly
 correlated with "integratively complex" reasoning strategies in which people (a)
display tolerance of ambiguity and sharply opposed perspectives on issues (On
the one hand . . ., on the other hand . . .); (b) attempt to generate integrative
or bridging cognitions that specify how processes identified by opposing viewpoints
interact to produce outcomes that devotes of the original viewpoint might
never have expected. Integratively complex (or dialectical) thinkers will be less
swayed by simplistic historical analogies and more inclined to note both simi-
larities and differences between current and past problems (cf. Neustadt & May,
1986; Tetlock, 1991). Integratively complex thinkers will be less likely to over-
assimilate contradictory evidence into their preconceptions and more open to the
possibility that they might have been wrong. And integratively complex thinkers
will be less prone to Manichaean worldviews that attribute only good charac-
teristics to their side and only bad characteristics to the other. If international
politics is indeed governed by a dissonant cacophony of causal processes, there is
good reason to believe that integratively complex thinkers will be especially
well-positioned to pick up on the subtle and contingent predictive cues that do
exist.

The Search for Wisdom Continued: The Fundamentalists

Whereas the complexifiers emphasize tolerance of ambiguity, contradiction,
and change as critical ingredients for good judgment, the fundamentalists take a
starkly different approach. First, they discount the possibility of short-term pre-
diction and focus on the long term. A useful comparison is to the stock market.
Historically, it has been easier to predict whether the market would be higher or
lower a decade hence than it has been a month hence. Fundamental economic
forces produce a long-term upward trend in market prices. Short-term forces
produce a great deal of oscillation (the random walk) around this trend (Malkiel,
1990). This combination of causal processes produces the pseudo-paradox of
short-term unpredictability and long-term predictability.

Second, the fundamentalists stress the importance of a "field-independent"
cognitive style that allows the perceiver to see simple patterns in complex events.
The key to good judgment is to focus on the underlying geopolitical, economic,
and technological forces that determine the relative power of nations and to
ignore irrelevant short-term distractor variables that dominate the headlines and
shape public opinion of the moment (cf. Waltz, 1979). There are many examples
of this approach. Mearsheimer (1990), for example, draws on neorealist theory
to argue (quite against the optimistic spirit of the times) that the break-up of the
Soviet bloc has actually increased the likelihood of wars by creating a whole new
set of security dilemmas in Eastern Europe. He can even claim an early "hit"
with the Yugoslav secession crises of 1991. Bueno de Mesquita and colleagues
(1985) draw on rational choice theory to develop a checklist of key questions
that need to be answered to reproduce the expected utility calculus of key decision-
makers and apply their method to predict the fate of Hong Kong. Mancur Olson
(1982) points to a small number of reliable political-economic signs that a nation
has lost its capacity to respond resiliently to internal and external challenges as a
result of the growing influence of special interest groups. Paul Kennedy (1987)
suggests that imperial powers have a life cycle and tend to mismanage the
inevitable trade-offs among consumption, investment, and defense in ways that
lead to their ultimate demise.

In short, fundamentalists may disagree among themselves as to exactly what
is fundamental; they agree, however, that good judgment requires an eye for the
abstract forces underlying the superficial flux of events that define the present.

Complementarity of Styles of Thinking

There are obvious tensions among the three approaches to good judgment.
The skeptics warn against the perils of one type of inferential error: the danger of
looking for order when randomness and chaos reign supreme. At best, we will waste
our time; at worst, we will delude ourselves (at least temporarily) into thinking we
can predict the unpredictable. Both the complexifiers and fundamentalists disagree and warn us of the opposite error: the danger of failing to look for patterns when they do exist and when knowledge of patterns could forestall disaster. This alliance between the complexifiers and fundamentalists dissolves quickly because they disagree over the form that order takes and over how to find it. The "complexifiers" focus on the danger of premature closure, belief perseverence, and overconfidence (the risk of ignoring relevant information); the fundamentalists focus on the danger of being distracted by trivia in a world with a very unfavorable signal-to-noise ratio (the risk of heeding irrelevant information). Each side can point to examples of these different types of error. There are occasions when observers should probably have tried to be more integratively complex by noting differences as well as similarities between the current problem and preferred historical precedents (Vietnam was not Korea—Khong, 1991; Neustadt & May, 1986). And one can cite occasions when observers may be well-advised to focus on one or two fundamental indicators and ignore everything else (Do we really need to understand the intricacies of North Korean political culture to predict the eventual collapse of the regime? Is it not sufficient to note the drying up of its foreign currency reserves and of support from its former allies and the inevitable collapse of its military-industrial complex?).

This argument suggests that good judgment should be identified with neither the pro-complexity nor the pro-simplicity camp, but rather with the "metacognition" camp that places great importance on the ability of observers to shift styles of thinking in response to changing situational demands. The wise observer can move flexibly from simple to complex to simple strategies of thought as the situation warrants. The key question here becomes how one determines what the situation warrants. Is it possible to identify specific cues that should serve as indicators of the usefulness of different styles of thought? Or do such metacognitive judgments call lie outside the range of social science? Is good judgment ultimately an art form?"
tions about the likely intentions of the upcoming generation of Soviet leaders, the stability of both the political and economic system, intermediate-range and strategic nuclear competition with the United States, geopolitical competition with the United States, frequency of high-level diplomatic contact (including summits), conventional forces in Europe, human rights and emigration, trade, and continuation of current alliance structures. We also posed a summary question: “On balance, do you believe American-Soviet relations will be better, worse, or essentially the same (1) (5) (10) years from now?”

In the case of South Africa, the interviews explored whether respondents expected a right-wing military coup, continuation of the status quo, liberalization within the ruling political party, a return or renewal of apartheid, legalization of the African National Congress (ANC), free multiparty elections, severity of violence within the society, degree of international isolation, and a set of other issues. We also posed a summary question: “On balance do you believe the South African government (1) (5) (10) years into the future will be more repressive, less repressive, or essentially the same as the government currently in power?”

In the case of the Gulf crisis, the interviews explored whether respondents believed Iraq would be forced through either military or economic pressure to withdraw its forces from Kuwait, whether the same governments would remain in power in Iraq, Saudi Arabia, Jordan, and Syria, the future of Saddam Hussein in particular, and whether there would be any significant changes in borders between countries in the region. The summary question was: “Which of the following outcomes do you see as most likely within the next (1) (5) (10) years: peaceful negotiated settlement, continuing military stand-off, or a major war? If you anticipate a major war, what outcome do you expect?”

In the case of the Korean peninsula, we asked respondents whether they believed North Korea was a politically and economically viable state, whether Kim II Sung or his son would still be in power, what the likely economic and geopolitical consequences of reunification were, and whether reunification could be achieved in a reasonably bloodless manner. The summary question was: “Which of the following outcomes do you see as most likely in the next (1) (5) (10) years: the same regime in power, a reformist regime, or the process of merging with the South will be well-advanced?”

The most obvious aspect of the data is their consistency with the skeptics’ position on good judgment. The results testify to the variety of ways in which well-informed and intelligent people can get things wrong. Of the eight specialists on the Soviet Union and American-Soviet relations, four believed that American-Soviet relations in 1988 would be very similar to American-Soviet relations in 1983 (a continuation of the “new Cold War” status quo), two believed that relations would be significantly worse as a result of the American defense build-up and significant shift toward a more hard-line, nationalist, and even neo-Stalinist leadership in Moscow, and two believed there would be a significant improvement in American-Soviet relations. Interestingly, there was little connection between political ideology and optimism-pessimism of the forecasts. Liberals and conservatives who held sharply discrepant political preferences held surprisingly similar images of the future. Advocates of the conflict spiral school of thought expected the legacy of the Reagan defense build-up to be a bitter and protracted nuclear arms race and an intensified struggle for influence in the Third World (Afghanistan, Angola, Nicaragua). Defenders of the Reagan administration were not noticeably more optimistic. They tended to subscribe to a rigidly “essentialist” image of the Soviet political system that left little room for internally initiated reform of the regime (the “evil empire”). The two observers—one liberal and one conservative—who were optimistic (and relatively accurate in the five-year range) were also the most difficult to classify ideologically and the most willing to offend existing orthodoxies (a point to which we shall return in discussing correlates of accuracy).

Of the five specialists on South Africa, three expected a continuation of the status quo, one expected a significant shift toward the right in the white-minority government, and one expected significant movement toward liberalization (similar in spirit to the policy initiatives of the de Klerk government). Again, there was not a strong connection between political point of view and forecasts. Liberals (who endorsed tough sanctions) and conservatives (who opposed them) agreed that little was likely to change and that there was a significant risk of provoking extremists in the Pretoria government and security establishment. The sole optimist straddled the conventional political divisions and drew on arguments from both camps to formulate forecasts.

Of the seven experts interviewed on the Gulf crisis, two expected peaceful negotiated compromise. In this scenario, Saddam Hussein would attempt (successfully) to fracture the U.S.-led coalition by engaging in aggressive withdrawal from Kuwait. The immediate result would be de facto annexation of a small part of Kuwait, the strengthening of Saddam’s political position in Iraq and the Arab world at large, and growing doubts about the capacity of the U.S. to protect its friends in the region. The long-term results would be extension of Iraqi influence into Syria and Jordan (both predicted the end of the current Syrian and Jordanian governments), the resurgence of Arab militancy toward Israel, and formal Arab recognition of the partial annexation. Two other experts expected continuing military standoff. In this scenario, Saddam Hussein would call the U.S.-led alliance’s “bluff,” hold on to all of Kuwait, wait out economic sanctions (which would prove “porous”), and continue as leader of an Iraq that, with Kuwaiti oil wealth, would emerge as the dominant regional power. Finally, three experts correctly predicted war between the U.S. alliance and Iraq. One of these experts predicted a long (six months plus) war that would provoke widespread sympathy for Iraq in the Arab world and destabilize pro-Western regimes; two others
predicted a brief, decisive campaign against Iraq that would lead to Saddam Hussein's rapid fall from power and greatly bolster U.S. influence in the region (one expert going so far as to see a successful multilateral peace conference on the Arab-Israeli conflict as a causally linked result by 1995).

Neither time nor events allow us to reach any conclusions about the accuracy of predictions about the Korean peninsula. For the record, however, the five specialists were unanimous that the days of the North Korean regime were numbered. Beyond that point, the experts understandably differed among each other over how imminent the demise of the regime was. Four were quite or very confident that North Korea would not exist as a separate state and would be absorbed into the South by 1996. The other expert cautioned against underestimating the staying power of the Pyongyang government and saw the most likely five-year outcome as “continuation of the status quo but worse—greater misery, deprivation, and repression.” Even this expert was unwilling, however, to predict that the regime would last a full decade (2001). The experts also agreed that the political survival of the Pyongyang regime hinges on the physical survival of Kim Il Sung and his son. After these political personalities disappear from the scene, the experts expected rapid disintegration of the regime—either along the lines of the German Democratic Republic (a reformist party leader tries unsuccessfully to contain popular resentment and desire for reunification) or, more likely, along the lines of Romania (a brief but violent civil war between elite loyalist forces and large sections of the army).

In those cases where we can now gauge the accuracy of experts’ forecasts (American-Soviet relations in 1988, South Africa in 1991–1992, and the immediate one-year outcomes of the Gulf crisis), accuracy does not appreciably diverge from chance. For example, if we had randomly assigned experts to the positions that the U.S.-Soviet relationship would either improve or deteriorate or remain the same, the experts so assigned would have had hit rates comparable to, or better than, the experts interviewed for this study. The same argument holds for experts who ventured forecasts on South Africa and the Gulf crisis. They also fared no better than if we had assigned them randomly to the logically exhaustive set of possible outcomes within their domains. This aspect of our results is clearly consistent with the skeptics' hypothesis that international politics is, by and large, too complex and indeterminate to permit successful forecasting.

From the standpoint of the skeptics, if one asks enough people enough questions, some answers are probabilistically fated to come out correct. I cannot rule out this hypothesis with the current data. My sample size is too small and the performance of the experts too close to chance accuracy to dismiss the skeptics with any confidence. We should not however be inhibited by the skeptics from carefully examining alternative hypotheses concerning the nature of good judgment in international politics. Although forecasting accuracy does not exceed chance, this observation does not require us to conclude that luck and luck alone is at work. Some ways of thinking about international politics may translate into more accurate forecasts than other ways, and it would be a serious mistake not to take this possibility seriously.

Complexifiers argue that experts who display greater tolerance for ambiguity, change, and contradictions in their thought protocols will tend, on average, to have superior track records in anticipating major international trends and events. We can test this hypothesis by looking for correlations between accuracy on the one hand and the styles of reasoning by which experts generate forecasts on the other. For instance, we subjected the interviews with experts to content analysis to assess the frequency with which experts invoked (a) conflicting or mutually reinforcing causal arguments in their analyses of events (do experts distinguish contradictory causal forces at work or do experts view the causal forces at work as pushing events in essentially the same direction?); (b) integrative rules, principles, or criteria for specifying how much weight one should attach to different causal processes.

The following examples illustrate the acknowledgement of causal contradictions:

The Reagan defense buildup is probably (and remember I am guessing) having two major effects on the Soviet leadership. Some of them are as mad as hatters and inclined to up the ante. Others see the [geopolitical and military] competition with the West as an exercise in futility at best and self-defeating at worst.

Economic sanctions aimed at South Africa polarize the ruling party, driving hard-core racists toward a fascist bunker mentality and persuading the pragmatists in the business community that apartheid must be dismantled if South Africa is to be reintegrated into the world community.

Saddam Hussein is performing a balancing act. The longer he stands up to the West, the longer he can keep Kuwait, the stronger his claim to leadership of the anti-Zionist, anti-Western Arab world. But eventually the U.N. sanctions will start to hurt. And eventually the international coalition may strike hard against Iraq itself.

Passages like this leave the listener hanging. In the interviews, I prodded respondents to be more specific and some accommodated by generating integrative cognitions that specified guidelines or rules for assigning priority to one or another line of causal argumentation.

Examples of integrative-cognitions include the following:

One effect will probably be larger than the other. My guess is that the new generation of Soviet leaders we’ve been talking about for so long is really quite fed up with the mismanagement of the old generation. The desire to open up to the world, modernize, will trump the desire to confront the West.

It is hard for an outsider to gauge the relative influence of factions in the secretive world of Afrikaner politics. Perhaps I’m underestimating the xenophobia and influence of the hardliners. But my view is that most of them will ultimately vote their pocketbooks. The pragmatists will win out, although the hardliners will be able to slow down the process of dismantling the system. They may even resort to selective assassinations to stir up trouble.

Some people think of Saddam as a good poker player. How could he have survived the treacherous world of Ba’ath Party politics otherwise? Be that as it may, we shouldn’t overestimate this guy’s rationality. Remember the Iranian invasion. And remember he is
surrounded by sycophants and sees the world through a glass darkly. All in all, I see a serious risk that he will overplay his hand and the West will have no choice but to attack. War is a 70/30 proposition I'd say.

On balance, the data suggest that experts who both acknowledged causal contradictions and made self-conscious attempts to reach integrative conclusions made somewhat more accurate forecasts than other experts. As noted earlier, only two of the eight American-Soviet experts, one of the South Africans, and three of the experts on the Gulf crisis made reasonably accurate forecasts on the central questions of interest here. Of these six experts, five received high scores on the causal-contradiction and integrative-cognition indices and one received a low score. Of the 14 experts whose forecasts were patently inaccurate, only four received high scores on the causal-contradiction and integrative-cognition indices and 10 received low or moderate scores. The three experts who made the least accurate forecasts regarding Soviet-American relations and South African politics received among the lowest scores on the causal-contradictions and integrative-cognitions indices. These results are of borderline statistical significance by the chi-square test and are consistent with the hypothesis that a cognitive capacity and willingness to confront causal contradictions is, if not a necessary condition, then at least a facilitative condition for reasonably accurate forecasting of international trends.

The data also offer some support for the fundamentalist perspective on good judgment. Both experts who accurately foresaw improvements in the U.S.-Soviet relationship mentioned the enormous economic pressures on the Soviet Union to scale back its military spending and geopolitical commitments. The single expert who foresaw major reform in South Africa also anchored his analysis in compelling trade and financial interests that, it was argued, would encourage “new thinking” from the business community and pragmatists within the National Party. It is also worth noting that the two experts who were most inaccurate in their forecasts of trends in American-Soviet relations (predicting deterioration) and the single expert on South Africa who predicted movement toward a more oppressive apartheid regime in Pretoria almost completely ignored economic trends in their forecasts. These data are quite consistent with the hypothesis that good judgment in international politics requires an appreciation of both economic and balance-of-power arguments.

It is clearly possible, however, to overestimate the rationality of political leaders and their responsiveness to the calculus of power. A vivid example comes from the two experts on the Gulf crisis who expected Saddam Hussein to engage in “aggressive withdrawal” from Kuwait. Both experts (one quite self-consciously) applied Hans Morgenthau's method of rational reenactment in which they tried to place themselves in Saddam's position and to identify the most effective way of fracturing the increasingly formidable coalition arrayed against Iraq. The three experts who accurately foresaw outbreak of war all invoked psychological or cultural-political reasons why Saddam Hussein would not make the “rational” balance-of-power response and withdraw in the face of superior force (reasons that included Saddam's narcissistic personality, the emergence of extreme groupthink among Saddam's sycophantic advisers, and the systematic attempt by President Bush to cut off all face-saving avenues of retreat from Kuwait). In this political domain at least, a sensitivity to psychological arguments appears to have translated into forecasting accuracy. But it translated into accuracy only up to a point. One of the three experts incorrectly foresaw a long war and, although the jury is still out on whether Saddam's influence was enhanced by the conflict, that proposition looks dubious at the moment; the other two correctly foresaw a short war, but the jury is still out on whether Saddam will fall from power as a result of his massive defeat.

CONCLUDING REMARKS

The study reported here is work in progress. The sampling of experts, political domains, and time frames is far too sparse to justify drawing strong conclusions about the validity of the three theoretical perspectives on good judgment. This methodological disclaimer is not, moreover, merely ritualistic. There are good reasons for suspecting that each of the three perspectives—skeptic, the complexifier, and the fundamentalist—contains important insights into the nature of good judgment at particular junctures in history. Few scholars would deny that accurate forecasting is sometimes the product of chance. Who could have predicted that the 1944 assassination attempt on Hitler's life would fail because someone inadvertently pushed the briefcase bomb behind a heavy oak leg of the conference table? Or that the Doctors' Plot purges would be brought to a sudden halt by Stalin's death as a result of a cerebral hemorrhage? In a similar vein, few scholars would categorically deny that some failures of forecasting can be traced to the failure either to recognize causal contradictions or to integrate such contradictions into projections of future trends. Who would deny that skillful political leaders must often strike compromises between conflicting values and constituencies (market efficiency versus social equality, crime control versus due process, economic growth versus environmental protection)? How can one hope to anticipate what such leaders will do if one does not understand the integratively complex decision strategies that guide them? And few scholars would challenge the fundamentalist claim that some failures of forecasting stem directly from obliviousness to long-term economic, technological, and military trends? Who would deny that, Hitler alive or dead, the
fate of Nazi Germany was sealed well before the assassination attempt of August 1944? Or, on a broader time scale, that the global economic dominance of the United States has eroded quite dramatically in the post-World War II period? Each perspective on good judgment surely has some validity; it is important, however, to advance beyond open-mindedness. A proper theory of good judgment should specify—at minimum in rough outline—the conditions under which each perspective is most likely to prove useful. A host of such boundary-condition questions need to be addressed: Does chance play an especially pivotal role in fluid, rapidly changing political situations (e.g., revolutions, civil wars?) in extremely rigid, totalitarian systems (Hitlers, Stalins, and Kim II Sung?) Does chance recede in importance in politics with widely accepted institutional procedures for transferring and maintaining power? Is chance more important in the short term and less important the further one looks into the future? Are fundamental causal processes more important in the long term and less important in the short term? Are experts with complex cognitive styles better equipped to anticipate the policy directions of leaders with similar cognitive styles? In closing, we should be careful not to underestimate the empirical and conceptual obstacles to achieving a more precise and differentiated understanding of good judgment. There is a serious risk in this line of research of attributing substantive significance to chance conjunctures of reasoning styles and forecasting hits and misses. The strong form of the skeptic’s hypothesis is that it makes as little sense to talk of good judgment in international politics as it does in games of pure chance, such as roulette. There will frequently be patterns in randomly generated data. An alternating series of red-black or black-red may arise or a series of increasing prime numbers or decreasing even numbers may emerge. For any given spin or sequence of spins of the roulette wheel, there will usually be at least one expert who can claim vindication. A comprehensive test of the three major perspectives on good judgment requires a much larger-scale research program than the study reported here.

There is also a serious risk in unequivocally categorizing forecasts as hits or misses so soon after the specified deadlines elapsed. Some observers in the 1970s foresaw the disintegration of the Soviet Union by 1984 (Amalrik, 1970). In 1985 we would have concluded they were wrong; in 1991, such forecasts look prophetic. In a similar vein, observers who expected a neo-Stalinist retrenchment by 1989 were wrong but they could claim at least weak vindication from the attempt at military coup in August 1991 (as Richard Pipes was quick to claim on the second day of the coup—New York Times, August 21, 1991). There was a core group of traditional Marxist-Leninists who refused to go quietly into the night of history, and perhaps a more cohesive conspiracy could have successfully seized power. It is appropriate therefore to end this article on a note of humility. We may not know for many years what hidden nuggets of wisdom lie within...


