H-Diplo | ISSF

Roundtable, Volume VII, No. 19 (2015)

Keren Yarhi-Milo. _Knowing the Adversary: Leaders Intelligence and Assessment of Intentions in International Relations._ Princeton, NJ: Princeton University Press, 2014. ISBN: 9780691159164 (paperback, $32.95/ £22.95); 9780691159157 (hardcover, $95.00/£65.00).

Contents

Introduction by Jack S. Levy, Rutgers University

How political leaders and their intelligence agencies assess the long-term intentions of their adversaries in international politics, how their assessments change in response to changes in the adversary's capabilities or behavior, and the extent to which political leaders rely on their intelligence agencies are old questions in the study of international relations. The assessment of long-term intentions is an extraordinarily difficult task, and the development of generalizable theory about the process is equally difficult. Keren Yarhi-Milo's recent book, *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*, is an enormously valuable contribution to our understanding of these questions. Unlike many studies of intelligence, it is well-grounded in international relations theory, and it effectively builds upon theories of social psychology, cognitive science, and organizational theory. Yarhi-Milo distinguishes herself from many other theorists by emphasizing that the assessment processes of political leaders may differ from those of state intelligence organizations, but at the same time she integrates both within a single overarching theoretical framework. Yarhi-Milo tests her theoretical arguments against leading alternative interpretations in three sets of important and revealing historical cases: British assessments of Germany's intentions from 1934-1939; and U.S. assessments of Soviet intentions during the years leading to the collapse of détente (1976-1980) and during the end of the Cold War (1985-1988). Yarhi-Milo's in-depth comparative studies utilize historical archives, published documents, and, for the U.S.-Soviet cases, interviews with key participants.

The core of Yarhi-Milo's argument is the 'selective attention thesis,' which consists of three primary hypotheses: the 'vividness hypothesis,' which is grounded in affective models of decision-making; the 'subjective credibility hypothesis,' which is grounded in strictly cognitive models of decision-making; and the 'organizational expertise hypothesis,' which is grounded in models of organizations and information processing.[1] The first two focus on how political leaders form inferences about the long-term intentions of the adversary, while the third focuses on how intelligence organizations process information about adversary intentions. Together, these hypotheses explain both assessments of intentions and variations in those assessments across different governmental actors.

The vividness hypothesis posits that political leaders tend to rely on information that has an emotional impact, that is spatially and temporally proximate, and that is easily imaginable. Personal impressions from private interactions with adversary leaders are particularly important, as are those leaders' responses to 'litmus tests' devised by

---

[1] On affective, cognitive, and organizational models, respectively, see Irving L. Janis and Leon Mann, *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment* (New York: Free Press, 1977); Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976); and Graham T. Allison and Philip Zelikow, *Essence of Decision: Explaining the Cuban Missile Crisis*, 2nd ed. (New York: Longman, 1999). For an analysis of threat perception that integrates affective (motivated) and cognitive (non-motivated) psychological biases, see Janice Gross Stein, "Threat Perception in International Relations," in Leonie Huddy, David O. Sears, and Jack S. Levy, eds., *Handbook of Political Psychology*, 2nd ed. (New York: Oxford University Press, 2013), Chap. 12.

individual decision-makers. The subjective credibility hypothesis posits that decision-makers' preexisting world views and theoretical expectations exert a disproportionate impact on their perceptions of adversary intentions. Leaders tend to see what they expect to see. Consequently, they are relatively unreceptive to incoming information that does not fit their world views; they are slow to update their beliefs in response to new information; and decision-makers with different pre-existing beliefs or world views generally make different inferences about intentions based on new information. The organizational expertise hypothesis posits that intelligence organizations rely on different indicators to assess the adversary's long-term intentions than do political leaders. Given organizational missions, goals, and expertise, intelligence organizations prioritize the assessment of adversary capabilities, and they tend to assess adversary intentions through the prism of those capabilities.

Yarhi-Milo makes a serious effort to engage leading alternative theories of how states assess the long-term intentions of the adversary. From the literature she identifies the 'capabilities thesis,' the 'strategic military doctrine thesis,' and the 'behavior thesis.' The capabilities thesis is drawn from offensive realist theory and posits that uncertainty about the current and future intentions of the adversary leads state actors to make worst-case assumptions and base their assessment of adversary intentions on adversary capabilities. Increases in adversary capabilities are assumed to reflect an increase in the hostility of intentions, while slow-downs in military build-ups are assumed to reflect a turn toward more benign intentions. The strategic military doctrine thesis suggests that the adversary's military doctrine (offensive, defensive, or deterrent, for example), military deployments, and public declaratory statements reveal information about its intentions. The behavior thesis posits that state actors focus on non-capability-related adversary actions as diagnostic indicators of intentions. These actions might include military conquests, interventions, and withdrawals, and decisions to join or exit international organizations or arms control agreements, among other things. Yarhi-Milo distinguishes between current actions (undertaken by the current political leadership) or past actions (undertaken by a prior leadership), in an attempt to evaluate whether inferences about intentions are specific to a particular adversary leader or more generally valid for the adversary state.

Each of the reviewers offers substantial praise for *Knowing the Adversary*. David Edelstein argues that the book is likely to be a "go-to work" in the study of adversary intentions. Jennifer Sims praises Yarhi-Milo for her "bold leap" over the longstanding, wide, and seemingly unbridgeable gap between theories of intelligence and theories of international politics. Jeffrey Taliaferro argues that this is a "superb book." I agree. *Knowing the Adversary* is an important contribution both to the theoretical literature on the assessment of intentions and to the historiography of three important cases of threat assessment. Each of the reviewers raises some issues, of course, and to these I now turn.

David Edelstein identifies five major issues. The first three have to do with time, capabilities, and uncertainty. Edelstein notes that Yarhi-Milo focuses on the adversary's "long-term" intentions, which she defines as "longer than a week or two into the future" (263, fn. 2). Edelstein questions the analytic utility of this dichotomous conception of short-term and long-term, and argues that a more differentiated categorization of the long term

would be preferable.[2] He hypothesizes that the longer the time horizons, the greater the reliance on more general adversary characteristics like ideology and the less reliance on specific behavioral indicators. As for capabilities, Edelstein accepts the argument that capabilities do not automatically determine intentions. He argues, however, that Yarhi-Milo underestimates the extent to which material capabilities nonetheless constrain intentions, a point raised by other reviewers. That is, capabilities are necessary to implement intentions but not sufficient to determine those intentions. Edelstein also emphasizes the inadequate attention Yarhi-Milo gives to the role of uncertainty. He notes that although Yarhi-Milo's historical analyses occasionally mention the high degree of uncertainty that analysts attached to their estimates, she fails to integrate uncertainty into her theoretical analysis.[3] Edelstein applies the themes of time, capabilities, and uncertainty to Yarhi-Milo's efforts to establish congruence between beliefs about intentions and the policies that follow, while recognizing that this is not Yarhi-Milo's primary aim in the book. Edelstein ends with a set of methodological issues, including the question of case selection, noting that all three of Yarhi-Milo's cases focus on democracies assessing the intentions of non-democracies. Jeffrey Taliaferro also raises the case selection issue, and I return to it later.

Jennifer Sims focuses on a number of definitional and conceptual issues. She argues, first of all, that Yarhi-Milo fails to provide a clear definition of her central concept of 'intentions.' She points out that governments often do not know what they intend to do, especially over an extended period of time, and that this fact needs to be incorporated into the definition of intention.[4] Another definitional issue concerns the concept of 'perceiver' or 'decision-maker,' and how they differ from intelligence organizations. As Sims points out, Yarhi-Milo restricts her organizational expertise hypothesis to joint intelligence community products, not to individual or agency-level perceivers. This generates some additional theoretical issues. It suggests the possibility of a gap between the assessments of joint intelligence products and the assessments of individual heads of agencies – with the former being shaped primarily by estimates of capabilities (reflecting the expertise of intelligence organizations) and the latter being subject to some of the same affective and cognitive biases affecting decision-makers.

---

[2] For Edelstein's own work on time horizons, see David M. Edelstein, "Conclusions: Rethinking Interwar Grand Strategies," in Jeffrey W. Taliaferro, Norrin M. Ripsman, and Steven E. Lobell, eds., *The Challenge of Grand Strategy: The Great Powers and the Broken Balance Between the Wars,* ed. (New York: Cambridge University Press, 2012), 318-25.

[3] On the importance of acknowledging uncertainty in estimates, and of the need for sensitivity to and a proactive search for information that might invalidate one's estimates, see Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, NY: Cornell University Press, 2010).

[4] Sims's argument that governments do not always know what they intend to do can be linked to Edelstein's point about uncertainty. One of the things to be estimated is not just the single best estimate of adversary intentions, but the degree of uncertainty attached to that judgment. If there is any reason to believe that the adversary is uncertain about its own intentions, or has a leading decision-maker who is known to be volatile or unpredictable, that information should be incorporated into the degree of confidence associated with one's assessments of adversary intentions. For a similar argument see Jack S. Levy, "Misperception and the Causes of War," *World Politics* 36, 1 (October 1983): 97.

A related issue emerges from the fact that some political leaders previously served as the head of an intelligence organization, and some of them were deeply socialized and trained in those organizations. This raises the question of whether their earlier training has any influence on their threat assessments when they become political decision-makers, or whether they make a clean break. The implication of *Knowing the Enemy* would seem to be that Vladimir Putin's sixteen years in the KGB had no impact on the analytic framework he later employed during his role as Russian president and prime minister. Turning to democratic leaders, did Ehud Barak's many years in Israeli military intelligence affect his assessments of adversary intentions while later serving as prime minister? Similarly, did George H. W. Bush's role of director of the Central Intelligence Agency influence his assessments of adversary intentions while he was president? Did Robert Gates many years of experience in the CIA, including his role as director, affect his assessments of adversary intentions while he was secretary of defense?

Sims also raises the question of how political decision-makers interact with intelligence agencies. She argues that this interaction goes beyond Yarhi-Milo's emphasis on decision-makers' reading of formal and joint intelligence products to include their contact with individuals at different levels of the intelligence organization. That interaction varies with the nature of the issue, the pace of decision-making, the availability of sources on adversary intentions, and (presumably) on personal contacts. Sims also emphasizes the point raised by Edelstein about the importance of the time horizons associated with a particular issue, arguing that the most useful indicators of adversary intentions vary depending on the time horizons of interest.[5] She also argues that if the most useful indicators for a particular analysis – human intelligence, for example – are not be available, intelligence analysts tend to resort to capabilities. Sims ends by briefly articulating her own views, based on many years of experience in the intelligence community, of the keys to success in intelligence analysis.

Like Edelstein and Sims, Jeffrey Taliaferro also raises the question of whether Yarhi-Milo underestimates the role of capabilities in assessments of long-term adversary intentions. Taliaferro is sympathetic to the role of psychological factors, but he emphasizes the difficulty of empirically disentangling decision-makers' subjective assessments of the intentions of the adversary from their perceptions about the current balance of power and about trends in relative power. He notes that there is a high correlation between shifts in relative power and assessments of adversary intentions in all of Yarhi-Milo's cases.[6] In both

---

[5] Time horizons involve not only how far actors look in to the future, but the shape of the discounting function they apply to the future. Most experimental and field studies suggest that individuals discount the short-term future more, and the more distant future less, than is posited by the standard exponential discounting model that is utilized in nearly all economic analyses. Philip Streich and Jack S. Levy, "Time Horizons, Discounting, and Intertemporal Choice," *Journal of Conflict Resolution*, 51, 2 (April 2007): 199-226.

[6] Theoretically, this relationship might depend on the existing balance of capabilities. If the adversary is currently stronger but declining in relative capabilities, under some conditions we might predict more

the British-German case of the 1930s and the U.S.-Soviet case of the late 1970s, he argues, individuals and organizations in the democratic state perceived an adverse shift in relative power and moved toward a more "malign" interpretation of the adversary's intentions.[7] In contrast, the late 1980s witnessed both an objective shift in power in favor of the United States and a shift toward more benign U.S. perceptions of the Soviet Union. Taliaferro also notes, in a point that echoes Edelstein's concerns about case selection, the limited variation in the nature of the entity whose intentions are being assessed: each is a longstanding adversarial state of the democratic state. Taliaferro suggests that it would be useful for future research to examine perceptions of the intentions of actors posing novel types of security threats, such as transnational terrorist networks.

Having summarized the reviewers' critiques of *Knowing the Adversary*, I would now like to add some comments of my own. These comments, like those of the reviewers, are made in the context of great admiration for the book and appreciation of its contribution to our understanding of how states assess the long-term intentions of the adversary.

It is important to emphasize, first of all, that each of the three competing theories of assessments of intentions that Yarhi-Milo identifies – the adversary capabilities, strategic military doctrine, and behavior hypotheses – is a rational unitary actor theory. Their unitary actor assumption leads to the prediction of the absence of variation in assessments of adversary intentions across political leaders or between political leaders and intelligence organizations. Moreover, each assumes a rationalist costly signaling model of assessment of intentions. Given incentives for deception in international politics, the only words and actions of the adversary that reveal information are those that incur some costs for the adversary – costs that either constrain its future actions or involve an expenditure of resources that cannot be recovered. 'Cheap talk' is not credible.[8] This rationalist signaling modeling is now the conventional wisdom in most international relations research communities on international conflict. *Knowing the Adversary* should be read as a critique

---

hostile adversary intentions, driven by preventive logic to maintain the adversary's favorable position while the opportunity is still available.

[7] The late 1930s are complicated. By early 1939 British military intelligence had recognized that the pace of German rearmament was beginning to exceed the capacity of the German economy and that the current pace of the German arms build-up could not be sustained. That, along with British rearmament efforts, led the British to conclude that the relative balance of power would soon shift back in their favor. Assessments of Germany's intentions, however, were moving in the opposite direction, certainly after the Prague coup in March 1939, toward a more malign interpretation of German intentions. See Norrin M. Ripsman and Jack S. Levy, "Playing It Straight or Politicized Process? British Military Intelligence and the Nazi Threat, 1933-39," paper presented at a conference on Signaling and Perceptions: Theoretical Debates and Historical Perspectives," Princeton University, April 28-29, 2011.

[8] James D. Fearon, "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs," *Journal of Conflict Resolution* 41, 1 (1997): 68–90; Andrew H. Kydd, *Trust and Mistrust in International Relations* (Princeton: Princeton University Press, 2007).

of this conventional wisdom. It constructs an alternative theory of the assessment of long-term intentions, one that is non-unitary and, for the most part, non-rational.[9]

Others have advanced theoretical arguments about the limitations of costly signaling model, but the extensive and systematic evidence provided in Yarhi-Milo's historical case studies sets *Knowing the Adversary* apart from previous work. Yarhi-Milo also makes a distinctive and important new theoretical contribution to the literature by emphasizing the importance of personal impressions from face-to-face meetings and by specifying the psychological mechanisms that help explain the impact of those impressions on the perceiver. Although historians have occasionally emphasized the importance of personal impressions in individual cases,[10] this factor has been generally neglected in the theoretical literature. The causal importance of personal impressions relates to my point in the previous paragraph about costly signals: because political leaders can do a great deal to manipulate the impressions they convey in face-to-face meetings, information and interpretations derived from those impressions should not be particularly credible. Nonetheless, Yarhi-Milo demonstrates that decision-makers often give this this information considerably more weight than it warrants in terms of its evidentiary value.

Yarhi-Milo makes another important theoretical contribution by analytically separating political decision-makers and intelligence organizations. Analyses that fail to make this distinction – which include any analysis based on the capabilities, strategic military doctrine, or behavior theses described by Yarhi-Milo – cannot explain situations in which political leaders and intelligence organizations come up with different estimates of adversary intentions, or in which political leaders ignore intelligence estimates and substitute their own judgments. These situations are sufficiently common – most readers can probably think of examples beyond Yarhi-Milo's three cases – to suggest the limitations of any unitary actor model of threat assessment and to highlight the utility of analytically separating political leaders from intelligence organizations.

The utility of distinguishing between political decision-makers and intelligence organizations is evident in any attempt to explore the question – often asked by students of intelligence – of whether intelligence shapes policy or whether policy shapes intelligence. The failure to distinguish between political decision-makers and intelligence organizations leads to the failure to recognize two quite distinct causal mechanisms underlying the policy-drives-intelligence hypothesis. One involves a political mechanism – the politicization of intelligence, or top-down pressure by political leaders on intelligence agencies and analysts to shape intelligence to support the leaders' policy preferences.[11] The

---

[9] One can debate whether the organizational expertise model is rationalist or non-rationalist.

[10] A good example is Frank Costigliola, *Roosevelt's Lost Alliances: How Personal Politics Helped Start the Cold War* (Princeton: Princeton University Press, 2012).

[11] For a good analysis of the politicization of intelligence, see Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, NY: Cornell University Press, 2011).

other involves a psychological mechanism – the motivated biases (or motivated reasoning) of political leaders, where a political leader's preferences for a particular policy may lead them to see what they want to see, to rationalize their pre-existing policy preferences. These are analytically distinct processes. Among other things, the first can be explained by a rational model; the second cannot.

Despite the utility of Yarhi-Milo's theoretical framework for facilitating the differentiation between these two different paths to an outcome in which policy influences intelligence, Yarhi-Milo gives little emphasis to either the politicization of intelligence or to the role of motivated biases in her theoretical analysis of the assessment of intentions. Let me begin with the politicization of intelligence hypothesis. This hypothesis clearly does not fit Yarhi-Milo's vividness hypothesis, subjective credibility hypothesis, or the capabilities focus of the organizational expertise hypothesis, nor any of the three alternative explanations, each of which posits a unitary rational actor. Yet we can surely think of cases for which the politicization of intelligence hypothesis is a leading candidate for explanation. As Robert Jervis argues with respect to the political mechanism, "Policy often drives intelligence as much as intelligence drives policy."[12] Many argue, for example, that pressure by the George W. Bush administration on the Central Intelligence Agency contributed significantly to flawed U.S. estimates of the Iraqi nuclear weapons program.[13]

Yarhi-Milo briefly mentions the politicization interpretation in her discussion of British intelligence agencies' assessments of German intentions. She notes that Prime Minister Neville Chamberlain strategically used estimates from the Chiefs of Staff that supported his policy preferences, but argues that "there is no evidence, however, that Chamberlain intentionally manipulated British intelligence to reach these conclusions" (108).[14] This is an important analytic distinction, and her conclusion that Chamberlain did not attempt to influence the intelligence product is an important piece of the story of British intelligence in the late 1930s. Still, I would have preferred that the politicization hypothesis be explicitly mentioned in the theoretical section as an alternative explanation that needs to be considered. It probably does not play a role in the majority of cases, but it may play an important role in some cases, and it needs to be put on the radar screen of the empirical investigator.

---

[12] Robert Jervis, "Intelligence and Foreign Policy," *International Security* 11, 3 (Winter 1986–87): 154.

[13] Rovner, *Fixing the Facts*, chap. 7. For an interpretation that gives more emphasis to cognitive biases than to politicization in U.S. intelligence on Iraq, see Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, NY: Cornell University Press, 2010), chap. 3. Note that politicization includes not just top-down pressure, but also the establishment of a new agency, or a major modification of an existing agency, in order to bypass normal intelligence channels and provide a path for getting the intelligence that supports one's preferred policy. See Uri Bar-Joseph and Jack S. Levy, "Conscious Action and Intelligence Failure," *Political Science Quarterly*, 124, 3 (Fall): 475. 61-88.

[14] Ripsman and Levy reach the same conclusion in "Playing It Straight or Politicized Process?"

I now turn to the motivated reasoning argument. This hypothesis does not fit the subjective credibility hypothesis, which is grounded in strictly cognitive rather than motivated biases, or the organizational expertise hypothesis. Although the vividness hypothesis, like the motivated reasoning argument, does incorporate affective factors, those affective factors relate to face-to-face meetings and personal litmus tests, which are different than motivated reasoning.

One can dig one step deeper, however, and ask how political decision-makers form their personal impressions of adversary leaders in their face-to-face meetings. One possibility is that leaders' personal impressions are significantly influenced by motivated reasoning. Leaders go into meetings with adversary leaders with certain policy preferences. In those meetings they may subconsciously form impressions that help to reinforce those policy preferences. Such impressions may also help minimize any cognitive dissonance that might result from preferring one policy but viewing the adversary leader in a way that would contradict that policy.[15] That is, in their face-to-face meetings there may be some tendency for leaders to see what they want to see. If so, we could posit a causal chain leading from policy preferences to motivated reasoning to personal impressions. This is not to say that this is the only mechanism shaping the formation of personal impressions, but that it is one of several alternatives.

Although Yarhi-Milo does not address this possibility of a more generalized motivated reasoning hypothesis in her theoretical analysis, she mentions this alternative interpretation in her analysis of Chamberlain's assessments of German intentions. She notes the possibility that the Prime Minister's views may have been influenced by motivated biases, but she concludes that he was instead influenced by his personal impressions. She does not consider the possibility that motivated biases might have shaped Chamberlain's personal impressions (88). Later, in what looks like a possible contradiction, she includes Chamberlain's "motivated defensive avoidance" (a form of motivated bias) among several factors reinforcing his selective attention to information about German intentions (92).

It is not surprising that Yarhi-Milo gives only brief attention to motivated reasoning (or to politicization, for that matter), given that selective attention thesis points us in other directions and provides little reason to focus on this factor in empirical analysis. My point is that motivated reasoning can be integrated into the vividness hypothesis, and that this theoretical refinement would have led to a sharper empirical analysis of the possible role of motivated reasoning in the assessment of adversary intentions. This is not to say that motivated biases necessarily played an important role. In fact, Yarhi-Milo's discussion of how Neville Chamberlain, Foreign Secretary Anthony Eden, and Lord Halifax, who succeeded Eden as Foreign Secretary, reacted to their face-to-face meetings with German Chancellor Adolph Hitler provides some evidence that works against any motivated

---

[15] Cognitive dissonance is sometimes treated as a cognitive bias rather than a motivated one, but the emotional discomfort generated by cognitive dissonance provides a psychological motivation to reduce dissonance by aligning perceptions with preferences. See Leon Festinger, *A Theory of Cognitive Dissonance* (Stanford, CA: Stanford University Press, 1957).

reasoning interpretation. Two comparisons are particularly revealing. First, although Chamberlain had fairly strong policy preferences going into the meetings with Hitler at Munich, Eden did not, yet they each came away with quite positive impressions of Hitler and of his willingness to reach a negotiated settlement (81). Halifax's policy preferences were initially closer to Chamberlain's, but after the meeting with Hitler at Godesberg in September 1938 Halifax's perceptions of Hitler' intentions shifted much more significantly than did Chamberlain's. Halifax concluded that Chamberlain had been excessively influenced by Hitler's assurances (73).

Yarhi-Milo's affect-based vividness hypothesis and cognitive-based subjective credibility hypothesis, along with the motivated reasoning hypothesis, each emphasize individual belief systems and reflect an individual psychological model of judgment and decision-making. One might ask about the role of shared beliefs in the processes of assessment of adversary intentions. These shared beliefs might be the product of a societal-level ideology or cultural frames, or an organizational-level strategic culture.[16] Theoretically, these shared beliefs and cultural frames are often associated with constructivism.[17] In a footnote (269, fn 56) Yarhi-Milo refers to constructivism and explains why she does not incorporate changes in ideology or domestic practices into her analysis. She mentions the example of a convergence in two countries' ideologies and domestic practices, which should lead to more positive assessments by each state of the other's intentions, and argues that the predictions of this hypothesis should be indistinguishable from those of the behavior thesis. Fair enough. But ideologies, identities, domestic practices, and norms might affect assessments in another way. By shaping political leaders' beliefs about the adversary or about causal relationships in international relations, they can influence how leaders process information about adversary intentions. More specifically, ideological and cultural frames can lead to selective attention to information, anchoring on pre-existing beliefs shaped by those frames, and to inefficient updating in response to new information.

In one sense, this shared beliefs hypothesis fits perfectly into Yarhi-Milo's subjective credibility hypothesis of the selective attention thesis. One difference, however, is that shared beliefs would predict minimal variation in assessments across political leaders. One of Yarhi-Milo's important findings is that there are often substantial variations in assessments of adversary intentions across decision-makers. One cannot be too confident in generalizing from these three cases, however, and the investigator must be open to the possibility that selective attention to information might be driven by shared beliefs and cultural frames.

---

[16] Small-group dynamics, in the form of pressures for conformity, might also lead to shared beliefs. See Irving L. Janis, *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*, 2nd ed. (Boston, MA: Houghton Mifflin, 1982).

[17] Ted Hopf, *Social Construction of International Politics: Identities and Foreign Policies, Moscow, 1955 and 1999* (Ithaca, New York: Cornell University Press, 2002).

One such case might be the Israeli intelligence failure of 1973. One of the leading interpretations of that failure traces it to a collective mindset that resulted in selective attention to information that was consistent with that mindset and that filtered out contrary information. In this view, Israeli intelligence officers and political leaders alike shared the beliefs that Egypt would not go to war unless it was able to mount air strikes deep into Israel to neutralize Israel's air force, and that Syria would not go to war without Egypt. Because the first condition was not satisfied, these shared beliefs, later labeled 'the conception,' blinded Israeli intelligence to information that should have alerted them to the likelihood that Egypt and Syria were preparing for war.[18] Israel's Agranat Commission traced the intelligence failure to the "persistent adherence to 'the conception.'"[19] Regardless of the validity of this interpretation, one can certainly imagine shared beliefs playing the leading causal role in some other cases. It would be useful to think about how shared beliefs and cultural frames, grounded in a constructivist world view, might be incorporated into Yarhi-Milo's analysis, either as a supplement to the subjective credibility hypothesis, or perhaps as an alternative hypothesis.

Finally, I would like to comment on the argument that *Knowing the Adversary* is based on a somewhat narrow selection of cases. Edelstein, for example, while correctly applauding Yarhi-Milo for her sensitivity to the dangers of over-generalizing from these cases, suggests that "the book would have been strengthened by a more serious consideration of how non-democracies might perceive intentions differently or how democracies might view differently the intentions of other democracies." I agree with Edelstein about the value of additional comparative studies to explore how the assessment of adversary intentions might vary across regime type (and, I might add, across different types of democratic regimes and across different types of authoritarian regimes). I believe, however, that his advice for a broader comparative analysis is better read as a guide for future research than as a suggestion as to how *Knowing the Adversary* might have been strengthened. Given the realities of publishing, supplementing Yarhi-Milo's 350-page book with additional cases, as well as with additional theoretical discussion of impact of regime type, could only have been implemented at the expense of diluting the quality of Yarhi-Milo's existing empirical research. Validating the key hypotheses of the selective attention thesis, and testing them against alternative interpretations, is a highly data-intensive task. Yarhi-Milo's extensive research into the archives and published documents, along with her interviews, are crucial in demonstrating the validity of many of her claims in these particular cases. Gains in external validity (generalizability) from adding more cases would have come at the expense of a loss of internal validity – the validity of causal inferences within a particular case or set of data. My strong hunch is that the expansion in the number and variety of

---

[18] For example, Israeli military intelligence interpreted the unprecedented magnitude of Syrian and Egyptian deployments at the front lines as evidence of routine Egyptian military exercises and Syrian defensive moves, not as preparation for war.

[19] Agranat Commission, The Agranat Report [Hebrew] (Tel Aviv: Am Oved, 1974); Avi Shlaim, "Failures in National Intelligence Estimates: The Case of the Yom Kippur War," *World Politics* 28, 2 (April 1976): 348–380. For an interpretation that places more weight on individual actions within the Israeli intelligence community, see Bar-Joseph and Levy, "Conscious Action and Intelligence Failure."

cases in *Knowing the Adversary* would have involved a net loss. As a suggestion for future research, however, Edelstein's suggestions are right on target, as is Taliaferro's suggestion of examining assessments of the intentions of actors posing new types of security threats. Our understanding of how states assess the long term intentions of other actors would be well served by such studies, especially if future studies were guided by the same high standards of research set by Yarhi-Milo in *Knowing the Adversary*.

Participants:

Keren Yarhi-Milo is an Assistant Professor of Politics and International Affairs at Princeton University's Department of Politics and the Woodrow Wilson School of International and Public Affairs. Her work has been published in numerous academic journals including *International Security*, *International Organization*, *International Studies Quarterly* and *Security Studies*. Her book, *Knowing The Adversary: Leaders, Intelligence Organizations and Assessment of Intentions in International Relations* (Princeton University Press, 2014) is based on her Ph.D. dissertation, which won the 2010 Kenneth Waltz Award from the American Political Studies Association. She is currently working on a new book project that examines variation in leaders' concerns for international reputation in international politics. Yarhi-Milo holds a Ph.D. and a Master's degree from the University of Pennsylvania, and a B.A., *summa cum laude*, in Political Science from Columbia University.

Jack S. Levy is Board of Governors' Professor of Political Science at Rutgers University, and an Affiliate of the Saltzman Institute of War and Peace Studies at Columbia University. He is past-president of the International Studies Association and of the Peace Science Society. Levy's primary research interests focus on the causes of interstate war, foreign policy decision-making, and qualitative methodology. He is author of *War in the Modern Great Power System, 1495-1975* (1983); co-author (with William R. Thompson) of *Causes of War* (2010) and of *The Arc of War: Origins, Escalation, and Transformation* (2011); and co-editor of *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals* (with Gary Goertz, 2007), *The Oxford Handbook of Political Psychology*, 2nd ed. (with Leonie Huddy and David O. Sears, 2013), and *The Outbreak of the First World War: Structure, Politics, and Decision-Making* (with John A. Vasquez, 2014).

David M. Edelstein is Associate Professor in the Edmund A. Walsh School of Foreign Service and the Department of Government at Georgetown University. He is also a core faculty member in Georgetown's Security Studies Program. He received his Ph.D. and M.A. from the University of Chicago and his B.A. from Colgate University, and he has held fellowhsips at Stanford University and Harvard University. He is the author of *Occupational Hazards: Success and Failure in Military Occupation* (Cornell University Press), and he has been published in *International Security*, *Security Studies*, and *Survival*. He is currently completing a book manuscript on the role of time horizons in great power politics.

Jennifer Sims is a Senior Fellow with the Chicago Council on Global Affairs and a member of the Senior Advisory Group to the Director of National Intelligence (DNI), James Clapper. She is also a consultant on intelligence and homeland security for private corporations and the U.S. government. In 1998, Sims received the intelligence community's highest civilian

award, the National Distinguished Service Medal. Sims is currently writing a history of intelligence in international politics with support from the Smith Richardson Foundation. Her publications on defense technology and arms control include *Icarus Restrained: an Intellectual History of Nuclear Arms Control in the United States from 1945 to 1960* (Westview Press 1985), and "The American Approach to Nuclear Arms Control: A Retrospective," *Daedalus* (Winter, 1991). She has co-edited with Burton Gerber, *Vaults, Mirrors and Masks: Rediscovering US Counterintelligence* (Georgetown University Press 2009) and *Transforming U.S. Intelligence* (Georgetown University Press, 2005) and has published numerous articles and book chapters on intelligence theory, intelligence reform, counterintelligence and intelligence support to counterterrorism.

Jeffrey W. Taliaferro is an Associate Professor of Political Science at Tufts University, where he has taught since 1997. His research and teaching focus on security studies, international relations theory, international history and politics, and United States foreign policy. Taliaferro is the author of *Balancing Risks: Great Power Intervention in the Periphery* (Cornell University Press, 2004) which won the American Political Science Association's Robert L. Jervis and Paul W. Schroeder Award for the Best Book in International History and Politics. His articles have appeared in *International Security, Security Studies*, and *Political Psychology* and several edited volumes. He is co-editor, along with Steven E. Lobell and Norrin P. Ripsman, of *Neoclassical Realism, the State, and Foreign Policy* (2009) and of *The Challenge of Grand Strategy: The Great Powers and the Broken Balance between the World Wars* (Cambridge University Press, 2012).

Review by David M. Edelstein, Georgetown University

I n June 2001, President George W. Bush held his first meeting with Russian President Vlaidimir Putin at Brdo Castle in Slovenia. Asked following their meeting whether he could trust Putin, Bush famously replied, "I looked the man in the eye. I found him to be very straightforward and trustworthy. . .I was able to get a sense of his soul."[1] Many were unimpressed with the new American President's method for discerning the trustworthiness of a potential adversary with thousands of nuclear weapons, but as Keren-Yarhi Milo writes in her important new book, Bush's approach was not unusual for politicians (17). Yarhi-Milo seeks to answer the question of how states judge the intentions of other states. She argues that leaders tend to rely on particularly "vivid" encounters with the leaders of other states in order to judge their intentions. Intelligence organizations, on the other hand, tend to rely on indicators of intentions that fit with their particular organizational expertise, which often means simply measuring the material military capabilities of others.

While theorists of international relations have long debated the relative importance of beliefs about intentions in how states perceive threats, Yarhi-Milo focuses on the question of how states come to form beliefs about each other's intentions.[2] Aside from the large literature on the security dilemma and the related debates over the offense-defense balance, more general theoretical discussions of how actors in international relations discern each other's intentions have not been as common as one might expect.[3] Yarhi-Milo's answer to this important question begins by disaggregating the state and recognizing that different actors within a state may have different methods for judging other's intentions. In what she calls her "selective attention thesis," she suggests three different hypotheses that account for divergent beliefs about an adversary's intentions. First, as suggested above, individual leaders tend to be impressed by their vivid, personal encounters with other leaders. Second, the information that leaders view as credible indicators of intentions is likely to vary, based on the pre-existing beliefs of those leaders. Third, intelligence organizations rely on indicators of intentions that accord with their organizational expertise. She distinguishes her theoretical account from three existing

---

[1] "Press Conference by President Bush and Russian Federation President Putin," Brdo Castle, Brdo Pri Kranju, Slovenia, 16 June 2001. Accessed at:
http://georgewbushwhitehouse.archives.gov/news/releases/2001/06/20010618.html

[2] For the argument that beliefs about intentions are a significant element of threat perception, see Stephen M. Walt, *The Origins of Alliances* (Ithaca, N.Y.: Cornell University Press, 1987).

[3] On the security dilemma, see Robert Jervis, "Cooperation Under the Security Dilemma," *World Politics* 30, no. 2 (1978): 167–214; on the offense-defense balance, Keir A. Lieber, "Grasping the Technological Peace: The Offense-Defense Balance and International Security," *International Security* 25, no. 1 (2000): 71–104; Charles L. Glaser and Chaim Kaufmann, "What Is the Offense-Defense Balance and Can We Measure It?," *International Security* 22, no. 4 (1998): 44–82; for one discussion of how to assess intentions and the effect of beliefs about intentions on strategy, see David M. Edelstein, "Managing Uncertainty: Beliefs about Intentions and the Rise of Great Powers," *Security Studies* 12, no. 1 (2002): 1–40.

arguments in the literature: a capabilities thesis in which states simply assume the worst about intentions and equate capabilities with intentions, a military doctrine thesis in which states judge intentions based on how offensive the doctrine of another country is, and a rationalist behavior thesis in which costly signals—either in the past or present—are relied upon to judge future intentions.

After delineating the logic of these different theories, Yarhi-Milo empirically adjudicates between them using an impressive study of how beliefs about intentions were formed in three different cases: British beliefs about German intentions in the 1930's, the beliefs of the Jimmy Carter administration about Soviet intentions in the 1970's, and the beliefs of the Ronald Reagan administration about Soviet intentions in the 1980's after Mikhail Gorbachev had come to power. The empirical material is presented systematically and draws persuasively on evidence from a variety of primary sources, including archival materials, published documents, and personal interviews with key decision-makers. In the end, Yarhi-Milo offers a significant contribution to the literature on threat perception that is likely to become a go-to work on the subject of beliefs about intentions in international relations. That said, in this review I offer five different critiques. Four of them focus on theoretical and conceptual issues while the fifth has to do with the empirical research design.

The first critique involves Yarhi-Milo's conception of intentions. She repeatedly states that she is interested in how leaders and intelligence organizations judge an adversary's "long-term intentions" (14), which she defines as "strategic-political in nature with a time horizon longer than a week or two into the future." (fn. 2, 263). Such a definition, however, underappreciates the extent to which temporal dynamics might affect how intentions are perceived. Intentions are, by definition, about how an actor intends to behave in the future, but how one judges those intentions may depend on how far into the future one is attempting to judge. Beyond "a week or two into the future" simply is not discrete enough to capture this variation. If one is interested in a state's intentions a month from now, one could hypothesize that leaders and intelligence organizations alike are more likely to rely on specific behavioral signals related to a particular interest than more general indicators of intentions such as, for example, ideology. The longer one goes into the future, one might expect the converse to be true: behavioral signals may be viewed as ephemeral and less reliable while underlying characteristics may be seen as more credible. Why this is consequential is evident empirically in Yarhi-Milo's case study of British beliefs about German intentions in the 1930's. As the 1930's progressed—and, in particular, after the remilitarization of the Rhineland, the *Anschluss* in Austria, and the takeover of the Sudetenland—British time horizons were compressed, much more focused on immediate intentions than what could fairly be considered 'long-term' intentions, yet Yarhi-Milo treats this case the same as all her others. If one allows that actors must judge intentions over time horizons that vary more than just by whether they are more or less than two weeks

out, then one might reach different conclusions about how leaders and intelligence organizations attempt to discern intentions.[4]

Second, Yarhi-Milo insufficiently appreciates the relationship between intentions and capabilities. Even though she finds that actors within states do not simply assume the worst about intentions and, therefore, judge intentions based on capabilities alone, Yarhi-Milo too easily throws out the capabilities baby with the offensive-realist bathwater.[5] Yarhi-Milo finds consistent evidence that intelligence organizations judge intentions based on capabilities, but for reasons of organizational expertise rather than the uncertainty rationale suggested by offensive realists. Even still, her analysis of how political leaders judge intentions fails to recognize that an actor's intentions are necessarily judged conditional on an actor's capabilities. I might express my intention to run a four-minute mile next year, but if others know that I simply do not have that capability (or at least that it is highly unlikely that I have that capability) then they will discount that stated intention. A fuller, more satisfying account of how actors judge each other's intentions would theorize the relationship between beliefs about intentions based on personal interactions, pre-existing beliefs, and organizational expertise and an actor's likely capabilities to realize particular intentions.

My third critique revolves around the nearly unavoidable element of uncertainty involved in judging a state's future intentions. Yarhi-Milo is nearly silent, at least theoretically, on the issue of uncertainty, which generates the impression that actors come to far more confident conclusions about others' intentions than is likely, especially given the possibility that intentions can always change. Any prediction about future behavior is unlikely to be made with complete confidence, in particular as those predictions become more and more forward-oriented. Actors and organizations certainly must make some estimates about an actor's future intentions, but one would expect different levels of confidence about those estimates based on varying amounts of uncertainty. Even if an actor is seventy-five percent certain of another's intentions, that residual uncertainty is consequential and is likely to shape the policies for which it advocates as a result. Empirically, Yarhi-Milo recognizes the challenge that uncertainty poses to discerning intentions. For example, in discussing why U.S. intelligence agencies discounted reassuring signals from Gorbachev, Yarhi-Milo acknowledges, "One explanation for this is that leading intelligence figures understood the task of assessing Gorbachev's motivations and intentions as subject to inescapable uncertainty" (239). Yet the role of such uncertainty is scarcely discussed in the theoretical section of the book even though it might offer an alternative explanation for the behavior of intelligence organizations. Such agencies might rely on capabilities to judge intentions not simply because of their organizational expertise, but rather because they can be relatively

---

[4] On these temporal dynamics, see David M. Edelstein, "Rethinking Interwar Grand Strategies," in *The Challenge of Grand Strategy: The Great Powers and the Broken Balance Between the Wars*, ed. Jeffrey W. Taliaferro, Norrin M. Ripsman, Steven E. Lobell (New York: Cambridge University Press, 2012), 308-36.

[5] For the classic statement of offensive realism, see John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York: W. W. Norton, 2001).

certain about the capabilities of another while other potential indicators of intentions are rife with uncertainty, especially as one goes further into the future.

Fourth, Yarhi-Milo attempts to demonstrate that beliefs about intentions affect the policies that leaders advocate, but the connections offered are not always persuasive. The main purpose of the book is not to present a theory connecting beliefs about intentions to particular policies, and she readily acknowledges that other factors are likely contribute to certain policy positions (41). Nonetheless, in each of her empirical cases, Yarhi-Milo recognizes that there is little purpose in establishing how leaders and agencies judge the intentions of other states if those beliefs about intentions were inconsequential for the policies that leaders advocated. So, she seeks to establish congruence between variation in the beliefs that actors held about intentions and variation in the hawkishness of the policies they supported. While appreciating that this analysis is not the core of the book, its inclusion only prompts further questions related to the issues of time, capabilities, and uncertainty that I have raised above. The translation of beliefs about intentions into particular policies is almost certainly subject to (a) the time frame of the intention and the proposed policy, (b) the capabilities of the actor to carry out its apparent intention, and (c) the degree of uncertainty about the suspected intention. A fuller appreciation of these dynamics would enable a more persuasive account of how beliefs that another state has certain intentions are reflected in policies that either try to encourage more favorable intentions or resign themselves to preparing for aggressive behavior.

Fifth, while the breadth and depth of the empirical material presented in the book is truly impressive, there are at least three discomfiting elements in how the research is designed and executed. Yarhi-Milo's argument about how leaders form beliefs about intentions relies heavily on the notion of "litmus tests" through which others reveal their intentions by how they perform on the test (17). But Yarhi-Milo provides no way of recognizing such litmus tests ex ante. While she suggests that leaders may explicitly and consciously construct such litmus tests, it may also be the case that litmus tests emerge on their own without conscious prompting. When is an actor's reaction to a certain stimulus likely to be taken as an important litmus test of its intentions and when is it not? Additionally, Yarhi-Milo's "selective credibility" argument relies on the familiar claim that leaders tend to fit new information into pre-existing beliefs.[6] Only strong and credible disconfirming evidence is likely to alter a leaders beliefs, yet Yarhi-Milo offers little insight into when such evidence is likely to be sufficient to change beliefs (21). In her empirical chapters, Yarhi-Milo points to examples of changing beliefs about intentions, but one might reasonably wonder why certain pieces of information, but not others, were sufficient to alter beliefs about intentions. Finally, all three of the case studies that Yarhi-Milo examines involve democracies judging the intentions of non-democracies. She is admirably candid about the dangers of over-generalizing from the results of her study, but the book would have been strengthened by a more serious consideration of how non-democracies might perceive intentions differently or how democracies might view differently the intentions of other

---

[6] Robert Jervis, *Perception and Misperception in International Politics* (Princeton, N.J.: Princeton University Press, 1976).

democracies (250). Certainly, advocates of the liberal democratic peace might contend that the presence of liberal democracy trumps all else when democracies are considering the intentions of other democracies.[7]

While I view all of these critiques as significant, none of them should be taken as undermining the fundamental importance of this book. If international relations scholars are going to insist that beliefs about intentions shape state behavior, then it is critical to understand how those beliefs form, and Yarhi-Milo has made substantial progress in unpacking this difficult question. At the end of the day, though, and as Yarhi-Milo herself acknowledges, her question is really only important if one is convinced that beliefs about intentions do, in fact, influence the policies that particular leaders advocate. For this claim to be persuasive, subsequent work ought to consider more how time, capabilities, and uncertainty affect not only the beliefs states form about intentions but also the policies and strategies that they adopt as a result.

---

[7] Of course, gathering sufficient empirical material for how non-democracies perceive intentions would have been a substantial challenge. Moreover, Yarhi-Milo might suggest that looking at how democracies view the intentions of other democracies is less relevant since she has limited her universe of cases to states forming beliefs about the intentions of other states already acknowledge as adversaries, and contemporary democracies tend not to be each other's adversaries. That said, the choice to limit her analysis to a universe conditioned on the existence of an adversarial relationship might itself be problematic.

Review by Jennifer E. Sims, Chicago Council on Global Affairs

Structural Realism, perhaps the dominant theory of international politics, hides within its vast literature a deep contradiction: on the one hand, the assumption that states live in constant fear for their security because they cannot know that they are safe; and, on the other hand, the assumption that states know enough to appreciate the distribution of power within the overall system, fear for their interests, and act accordingly. How is it that states are presumptively uncertain, yet knowledgeable enough to balance against threats? And if we presume states can learn, must they inevitably be caught in the security dilemma wherein one state's efforts to shore up its own security always induces others to fear for their own?

Karen Yarhi-Milo's book is both bold and encouraging because it tackles these kinds of theoretical problems in the context of decision-making and system- level dynamics without neglecting the role of the intelligence process. In this sense, the author's Selective Attention thesis fits well with the pioneering work of such International Relation (IR) theorists as Robert Jervis, James Fearon and Charles Glaser while compensating for the neglect of many intelligence theorists of the decision-maker's role and system-level outcomes.[1] After 9-11, for example, intelligence scholars got busy reviewing intelligence agencies' practices in order to reform them, but left to battling memoirs the essential question of why decision-makers failed to listen to (as opposed to hear of) intelligence on Al Qaeda's intentions until it was too late.[2]

I therefore find the author's opening thesis – that decision makers selectively attend to intelligence products – refreshing, provocative and a timely corrective for those busy digging and constructing in the sub-field of intelligence theory.

Yarhi-Milo hypothesizes that intelligence analysts, shaped by their organizations, tend to rely on military capabilities to estimate the intentions of adversaries while decision makers estimate intentions less analytically and more affectively: by vivid experiences and

---

[1] See James D. Fearon, "Signaling Foreign Policy Interests: Tying Hands Versus Sinking Costs," *Journal of Conflict Resolution*; 41 (1), 2002; 68-90; Robert Jervis, *Perception and Misperception in International Politics, Princeton University Press* (1976); and Charles Glazer, *Rational Theory of International Politics: The Logic of Competition and Cooperation* (Princeton University Press) 2010. Since my view is that intelligence succeeds or fails only in the context of winning international competitions through conflict or cooperation (i.e. decision-making), much intelligence theorizing seems similar to designing the perfect sailboat absent wind and water. For a good review of intelligence theory, see Gregory Treverton and Wilhelm Agrell, eds., *National Intelligence Systems: Current Research and Future Prospects,* (Cambridge: Cambridge University Press) 2009. Also see Peter Gill, Stephen Marrin, and Mark Pythian, eds., Intelligence Theory: Key Questions and Debates (New York: Routledge) 2009.

[2] One book, Amy Zegart's *Flawed by Design: The Evolution of the CIA, NSC, and JCS* (Stanford: Stanford University Press, 2000) was published before 9-11 but had considerable impact on the movement for reform after the event. See also Zegart's *Spying Blind: The CIA, The FBI, and the Origins of 9-11* (Princeton: Princeton University Press, 2009. The memoirs are endless, but a good place to begin is George Tenet, *At the Center of the Storm*, (New York: Harper, 2007).

subjective beliefs derived from personal experiences. If she is right, then defensive realists' assumptions that states can signal their intentions may be true, but their assumption that states can do so through military means may be wrong. In a series of richly documented case studies, she finds validation of her theory.

Although there is much good to say about Yarhi-Milo's book, I will focus on three problems with her underlying theory in the hopes of advancing further work: definitions, conceptual issues related to intelligence, and complexity. In the first category, the definition of 'intentions' may be the most significant. Since few people know what they intend to do next year – or for that matter in the morning before they have had coffee – it seems safe to assume that governments often do not know what they intend either, especially over the long-haul. To overcome skepticism perhaps, Yarhi-Milo categorizes intentions as expansionist, opportunistic, and status quo – such broad terms that 'intention' becomes divorced from 'decision' and thus appears in her case studies as a relic of systemic theorizing dragged down to the level of government decision making and individual agency. I would thus suggest that the author operationalize the indicators for each of her categories. In doing so she may, however, run into problems. For example, 'opportunistic' would seem to be a good way to describe *the absence of* intention or, in other words, a reactive foreign policy. One suspects few governments actually *intend* to have no strategy, so are we expecting to look for signals of the equivalent of a pre-coffee morning? Distinguishing 'opportunistic' from 'limited-expansionist' intent would make this book's findings seem less vague and seemingly subjective.

In any case, from a theoretical and research standpoint, it is probably simplest and most productive to think of a state's 'intention' as the phase of activity from decision to execution, whether that intention relates to tactics or strategy. After the moment of decision, governments know their intent but have not fully acted in ways others can discern. Because decisions are often documented and can be discovered, intentions are, in this sense, certainly knowable. Absent good intelligence on the knowable, analysts must estimate intentions on some other basis – estimating that varies in strength based on what is known, how it is known, and the method of projecting the implications. The point is, by failing to clearly define what she regards as intentions in policy-making terms, Yarhi Milo generates confusion regarding the differences between divination (non-evidentiary projections – i.e. faith or unproven theory), assessment (treaty violations mean the adversary intends to cheat), and formal estimation (the new reprocessing facility indicates the intent to acquire nuclear weapons). Indeed, short-handing the intelligence process more generally – a point I will return to later – adds to these difficulties.

The author could also, for the purposes of her research, define the "perceivers" (9) more convincingly. For example, there are seemingly inconsistent claims about what sets intelligence analysts apart from decision makers: first, that because of their organizational affiliations intelligence analysts use military capabilities to estimate adversarial intentions; second, that they will not necessarily draw conclusions based on military capabilities; and third, in a seeming reversal, that the organizational expertise hypothesis, which holds that intelligence analysts' perceptions are shaped by their organizations, actually only applies to

joint intelligence community *products,* not individual or agency-level perceivers.[3] Clarification here would help other scholars replicate her work. Of course, clarification might limit the theory's application to particular types of states at particular historical times, making her theory more descriptive than parsimonious – a point I return to below.

The definition of "decision-maker" is equally unclear (9). The author concentrates in her case studies on "the primary decision maker (i.e. the prime minister or president) and their closest advisors on issues of foreign policy" (9) and characterizes them, unlike intelligence analysts, as uninhibited by bureaucracy (and egocentric, biased toward vivid rather than objective information, impressionable and prejudiced.) Given that many decision makers, even very senior ones, are often also time-pressed bureaucratic managers and former scholars, lawyers, military officers, or intelligence professionals, it is difficult to buy into the premise that their liberties increase but their analytical training fails as they rise.[4] But even if a researcher were to accept these characterizations, replicating the author's work would be troublesome. After all, the author does not make clear when we might predict in the course of an individual's career, which may involve bureaucratic promotions or a shift from intelligence to policy, that analytical skills at assessment or estimation will give way to the influence of experiential vividness and prior mindsets.[5] Indeed, the author makes her claim about decision-makers' proclivities and then backs off, claiming, for example, in the case of the vividness hypothesis, that "we should not anticipate that personal interactions will always cause leaders to infer intentions." This is leaky theorizing that leads to selectivity in applying her hypotheses to individuals.

A second general concern with Yarhi-Milo's theoretical approach has to do with her premises about intelligence. She identifies national intelligence as typically defense dominated, organizationally driven, and branded by joint estimative analysis. This provocative characterization is flawed for several reasons. First, it renders her theory

---

[3] See pages 23-25.

[4] It is worth pointing out that British Prime Minister Neville Chamberlain, featured in the author's first case study, has been authoritatively described as very much an organization man: "…a lucid expositor, and a competent defender, of Departmental policy." (Sir Arthur Salter as quoted in Ernest R. May, *Strange Victory* (New York: Hill and Wang, 2000), 171. Indeed, T.G. Otte in *The Foreign Office Mind: The Making of British Foreign Policy, 1865–1914* (Cambridge and New York: Cambridge University Press, 2013) makes a good case that one cannot understand senior level decision making in foreign policy without understanding the bureaucracies that shape the attitudes of senior decision makers (1-22). At the very least, Yarhi-Milo needs to warrant her proposition that decision makers are unencumbered by either the culture or the processes of their host bureaucracies.

[5] This methodological problem is recognized by the author but is brought up in the context of what might be done to extend her research further. My critique instead makes the point that the methodological issue is integral to her approach itself, and so dissuades anyone from extending or even applying her theory. For example, would the senior American diplomat Richard Holbrook, who successfully psyched-out Croatian, Serbian, and Bosnian intentions *while* serving in the Department of State's Bureau of European Affairs and as a special be considered a "decision maker" regarding the Balkans but not regarding the rest of Europe? If so, would the author expect to find him less analytical in his latter role than in his former one or, for that matter, when he was assistant secretary of state for the Far East?

irrelevant for state systems before 1900. Few great powers had national intelligence organizations prior to this date, let alone joint estimative products. Second, the author's emphasis on joint written estimates means that evidence of a decision-maker's rejection of these products becomes the main indicator of his rejection of analytical thought regarding the adversary more generally. Such logic reflects too narrow an understanding of the intelligence process, and offers only one (poor) test for the analytical quality of leaders' decision-making. Decision-makers engage intelligence professionals about sensitive source collection as well as oral all-source briefings and written analysis. All-source intelligence analysts do not necessarily know the full range of compartmented information available to decision-makers (such as through secret diplomacy, covert action or offensive counterintelligence operations). Senior decision-makers therefore regard even "all-source" products as supplemental to what they know through other means. Moreover, if decision-makers determine that collection against a target would be risky, they may purposefully limit it, dismissing the resulting analysis less because of their own subjective biases or non-analytical preferences than because of its poor factual underpinnings. The author's research cannot reveal such counterevidence to decision-makers' subjectivity because she has defined the alternative argument away. Finally, the author's supposition of a defense bias in intelligence institutions, if correct, would seem to validate decision-makers' discounting of their estimative products anyway and for analytical rather than subjective reasons. Even setting aside Yarhi-Milo's questionable characterization of intelligence, the larger point is that how decision makers interact with intelligence agencies will vary depending on the nature of the conflict at issue, the pace of decision-making, and the availability of sources on the adversary's intentions.

Variances in how intelligence works would therefore seem to be important to the selection of cases. The Cold War, which frames two of Yarhi-Milo's three cases, entailed a bipolar system, delicate strategic force balances, and a dearth of good human sources on Soviet decision-making because of the extreme secrecy of the Soviet state. Given that the arms race also implicated treaty regimes such as the Strategic Arms Limitation Talks (SALT), intelligence production was bound to focus on litmus tests based on those regimes and collection architectures constructed for the purpose of assessing and estimating military capabilities. Decision makers likely influenced those very choices through the requirements process. Estimating Soviet intentions included the degree to which weapons treaties were violated and the testing and deployment of forces because that was where the Superpowers had achieved some level of agreement and cooperation. In contrast, Queen Elizabeth I's intelligence advisors estimated Spanish King Philip II's strategic intentions in the mid-sixteenth century by the number of times his agents tried to infiltrate England and kill her. Requirements, collection, analysis, and estimation are all linked through dynamic interaction in case-specific kinds of ways. Although this point is significant for the larger issue of case selection in this book, the more nuanced point is this: without knowing the requirements or sourcing behind intelligence production in a given competition, it is difficult to ascribe biases in information processing or analysis to either intelligence analysts or decision makers.

Strategic intentions also have varied time horizons in international politics – a fact the Cold War seemed to mask. Suppose the strategic issue is whether the North Koreans will collude

with terrorists (ally for power), whether Iran intends to develop a nuclear weapons program (build power in a bid for regional supremacy), whether Saudi Arabia intends to stabilize oil prices (implicate U.S. power), or whether the proto-Islamic State intends to bring war to the continental United States (conduct war)? In these cases, estimating intentions using offensive or defensive military capabilities will work less well than considering past behavior (e.g. airborne particulates after testing, or imagery of disturbed earth suggesting a hidden reprocessing facility), or declaratory policies (either public or private "chatter" on cell phones). In some cases, the best indicator would come from having a double agent with access (Human Intelligence - HUMINT) in the room at the moment strategic decisions are made. In sum, it is doubtful that any analyst or decision maker in the above cases would *prefer* to derive intentions from military capabilities if other indicators were available.

Thus, while I applaud Yarhi-Milo's attempt to operationalize Offense-Defense theory's implications for perceivers, care must be taken when abbreviating intelligence processes for the purpose of testing systemic theory. All sorts of methodological issues arise. As already explained, it is easy to confuse decision-makers' unmet informational needs, lack of intercepts, or lack agent access with cognitive bias.[6] In asserting causation of the latter kind, one has to control for the former or explain convincingly the reason for not doing so. In other words, if decision-makers' receptiveness to intelligence products has more to do with the sourcing, timing, and relevancy of the intelligence than with its analytical methodology, then theories built on the latter will be unconvincing. This obstacle does not mean that joining intelligence theory to systemic theory is impossible; indeed it is sorely needed. The author, who clearly is aware of the underlying issue, might be planning to test her thesis in more demanding cases than she does in this book. If so, it is necessary to keep her terms clear and import as little high-level theorizing as possible into her case analyses of perception and intent. Finding generalizable trends out of casework is, after all, how behavioralists, such as Daniel Kahneman and Ammos Tversky, have influenced system-level theories of how economies work.[7]

I would suggest that findable documents should not become the basis for intelligence theorizing. This is always a danger when working in the declassified domain. For example, declassified joint intelligence estimates seem tempting indicators of decision-makers' attentiveness to intelligence in part because governments have released them relatively

---

[6] The author notes that U.S. analysts during the Carter administration relied on military capabilities disproportionately in gauging Soviet intentions. But her research on the Team A / Team B exercise demonstrates only that Team B wanted estimates of intentions to rest not only on capabilities but on such ephemeral drivers as Marxist ideology, ideas, and the like. This reveals the point: how does the author know that reliance on military capabilities in National Intelligence Estimates [NIEs], rather than reflecting organizational bias, simply reflected the lack of spies or double agents who might have been able to provide hard evidence on plans. In fact the simplest explanation for all the NIEs during this period was that the US intelligence community had a HUMINT collection problem, not an organizationally bias.

[7] Daniel Kahneman and Amos Tversky, eds., *Choices, Values and Frames* (Cambridge University Press, 2000).

consistently, making their use as indicators feasible. But this very provenance makes their use suspect. Moreover, as the author is aware, Intelligence Community (IC) products are the least likely of all intelligence products to be read by senior decision makers for substantive reasons[8] This suggests that the author could usefully find supplemental intelligence indicators for her theory, keeping in mind that decision-makers prefer sensitively sourced, customized intelligence for their specific needs and frequently demand, therefore, restricted distribution. Unfortunately, such restrictions and sensitivity also makes their finding and declassification troublesome. It may therefore make sense to seek more general indicators. I have theorized, for example, that the willingness of decision-makers to share policy objectives with intelligence officials (a robust requirements process) and the latter's sharing of sources and methods with decision-makers (risk sharing) constitute the best indicators of decision-makers' attentiveness to intelligence products. In any event, joint estimates that are highly likely to leak or be declassified will be seen by decision makers as archival, academic, or potential time-bombs. The suggestion that such products are ignored at high levels – except, perhaps, for what political impact they might have if leaked – is equivalent to calling out gambling in a casino.

All this brings us to the matter of complexity. The author's manipulation of so many variables in her own thesis seems difficult enough, but her effort to contrast its supporting evidence with that applicable to three other theories (the Capabilities, Behavioral, and Strategic Doctrine theses) leads to confusion for the reader. Although I found the repetitive explanations of what we were learning from each case helpful and even necessary, they also seemed to cost us case content in the write-up. Absent evidentiary richness, including the exploration of alternative explanations, the analysis cannot convincingly tease out cause and effect.

Not surprisingly, problems in definitions, delimitations, and intelligence theory lead to errors in evidence gathering and findings. For instance, in the section on British decision-making before World War II, the author finds that joint assessments relied more on military capabilities than on relational or political indicators of intentions and that decision-makers did not seem influenced by them. Describing Prime Minister Neville Chamberlain as being impressed by his encounters with Adolf Hitler and, later, discussing his erratic behavior, she concludes that he found such non-analytical sources the most influential to his decisions. A case could equally well be made, however, that Chamberlain was attentive to the intelligence he was receiving through liaison with the French services and his own spy rings run by his senior intelligence analyst, Sir Robert Vansittart, Permanent Undersecretary in the Foreign Office. These espionage reports suggested that, despite Nazi Germany's capabilities, its leadership was fractured, some of Adolf Hitler's most senior military and intelligence officers were disaffected, and a few were even

---

[8] The author acknowledges that "decision makers have repeatedly complained about 'the ponderousness obviousness, and mushiness of typical (National Intelligence) estimates, and consider reading them a waste of time.'" [253, quoting in the original from Richard K. Betts's Article Review of Sulmaan Wasif Khan's "The Aesthetic of Analysis," H-Diplo Article Review 202 (November 28, 2008), http://h-diplo.org/reviews/PDF/Betts-Khan.pdf]

plotting a coup.[9] Chamberlain may have thought that if he could buy some time, more sober heads in Germany might prevail, lessening the need for continuing costly military expenditures *despite* Germany's growing capabilities. Such an assessment by Chamberlain would have had little to do with vividness or non-analytic biases. It would have been helpful if the author had more fully explored such counter-arguments.

After all, a credible alternative hypothesis is that the more competitive and complete the sourcing on intentions (behavioral, military, declaratory), the more influential the resulting analysis will be and the less instinctual or prejudicial decision-makers will be in estimating intentions and avoiding surprise. By extension, no matter what type of conflict is involved, intelligence estimates that tease out the hinge assumptions and press sources to affirm or counter them will tend to be more convincing than estimates based on skimpy personal experiences or military capabilities alone.

The foregoing criticisms notwithstanding, I am heartened by this book. The divide between intelligence theory and theories of international politics has long been wide and seemingly unbridgeable. Karen Yarhi-Milo has contributed a bold leap across it. While quibbling with the author's underlying theory and methodology, I conclude that the author is not proposing so much a theory as an analytical framework for scholars who are trying to understand past successes and failures in perceiving intentions. Yarhi-Milo's work should also stand as a warning to intelligence theorists that they have not yet made an impact on IR theory, partly because many of them ignore systemic constraints to decision making. Decision-making in conditions of uncertainty is risky and dangerous. Common sense tells us that, given these risks, decision-makers are likely to rely on their own knowledge supplemented by what they view to be good sources. Intelligence works to provide the latter, recognizing that leaders will trust their own first-hand knowledge over more distant sources of uncertain credibility. Intelligence theorists can contribute to IR theory by more rigorous investigation into how sourcing influences decision-making and the estimative process.

As an intelligence theorist, I remain convinced that intelligence is best understood as competitive learning. When employed by governments, it works to amplify what decision-makers need to know to win, either improving or degrading their power to affect the international system. It is my firm belief that the most productive theories do not turn on how individuals or organizations think, but on what makes intelligence succeed in clarifying the distribution of power, interest, and related actions among states. My own view is that there are four measurable keys to success: rich collection; connectedness to decision makers; flexible, semi-independent management; and a capacity for selective secrecy. Analysis is the vehicle for mind-merging with the decision-maker and is important, but not definitive. Strategic nuclear war was avoided during the Cold War because *both sides* created formidable, counterbalancing collection machines that decision makers trusted to get the principal question right: whether the other side was planning a surprise attack. The U.S. may have had advantages in technological collection and the Soviets in

---

[9] Ernest R. May, *Strange Victory*, (New York: Hill and Wang, 2000), 162.

secrecy and agent penetration, but both became good enough in collection to reduce the estimative problem to manageable levels and bring the Cold War to a close without strategic war. Unfortunately, these finely-tuned intelligence machines were less suited to the international system that rapidly evolved in the 1990s. They lost decision makers' trust as a result. Appreciating how states estimate others and learn to adapt requires, in my view, more study of how intelligence works.

Review by Jeffrey W. Taliaferro, Tufts University

H ow do great powers, or more properly, the senior officials charged with the formulation of foreign and defense policies and the intelligence agencies that serve them, actually make assessments about the long-term intentions of rival great powers? What prompts changes in how policymakers and intelligence analysts perceive the intentions of their adversaries? How often do the assessments of policymakers about the intentions of foreign leaders differ from those of their intelligence organizations?

Keren Yarhi-Milo's *Knowing the Adversary* is an admirable effort to answer these difficult questions. It joins a long and distinguished line of scholarly books that draw insights from social psychology, cognitive psychology, information processing, and organization theory to develop hypotheses challenging largely rationalist international relations theories and that use detailed historical case studies based on extensive archival research to test those hypotheses.[1] It also joins a growing number of recent scholarly books that advance theories about dilemmas in intelligence analysis and the interactions between intelligence analysts and policymakers, especially in the United States.[2]

Knowing the Adversary is a superb book. It is based upon Yarhi-Milo's dissertation, which won the American Political Science Association's (APSA) Kenneth N. Waltz award in the field of international security.[3] She asks big and important questions and develops an interesting theoretical framework to answer those questions. The book's qualitative research design is well executed. Drawing extensively upon archival research, Yarhi-Milo presents detailed comparative case studies of the assessments of Germany by the British government and military intelligence in the 1930s and assessments of the Soviet Union by the Carter administration, the Reagan administration, and the U.S. Intelligence Community in the late 1970s and the late 1980s, respectively. This book will be widely read by scholars and advanced graduate students interested in security studies, international history, and intelligence studies.

---

[1] See for example, Deborah Welch Larson, *Origins of Containment : A Psychological Explanation*, (Princeton, N.J.: Princeton University Press, 1985); Deborah Welch Larson, *Anatomy of Mistrust : U.S.-Soviet Relations During the Cold War*, (Ithaca, NY: Cornell University Press, 1997); Yuen Foong Khong, *Analogies at War : Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*, (Princeton, N.J.: Princeton University Press, 1992); Jonathan Mercer, *Reputation and International Politics*, (Ithaca, N.Y.: Cornell University Press, 1996); Barbara Farnham, *Roosevelt and the Munich Crisis : A Study of Political Decision-Making*, (Princeton, N.J.: Princeton University Press, 1997); James W. Davis, *Threats and Promises : The Pursuit of International Influence*, (Baltimore: Johns Hopkins University Press, 2000).

[2] See for example, Robert Jervis, *Why Intelligence Fails : Lessons from the Iranian Revolution and the Iraq War*, (Ithaca: Cornell University Press, 2010); Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence*, (Ithaca N.Y.: Cornell University Press, 2011); Thomas Fingar, *Reducing Uncertainty : Intelligence Analysis and National Security*, (Stanford, California: Stanford Security Studies, 2011).

[3] Keren Yarhi-Milo, "Knowing Thy Adversary: Assessments of Intentions in International Relations" (Ph.D., University of Pennsylvania, 2009).

I am quite sympathetic to incorporating insights from social and cognitive psychological theories into the study of international relations. However, I am not persuaded that British and U.S. leaders' assessments of German and Soviet intentions were not inextricably linked to their assessments of the current balance-of-power and anticipated power trends. This is not discount Yarhi-Milo's central claims about the importance of psychological biases of presidents, prime ministers, and their close advisors and the importance of organizational expertise in intelligence agencies in making assessments of great power rivals. Rather, I contend that is often difficult to disentangle leaders' subjective assessments of an adversary state's intentions from their perceptions and prior expectations about the balance-of-power and power trends.

Yarhi-Milo's selective attention thesis holds that senior policymakers rely far more on information that is vivid, salient, and personal (for example, personal interactions with foreign leaders) than on abstract or statistical information about an adversary's capabilities, even when latter has greater evidentiary value. Policymakers make highly subjective evaluations about the credibility of information about an adversary's intentions based upon their preexisting beliefs and expectations of that adversary's behavior. The selective attention thesis suggests three testable hypotheses: the first two seek to explain the inference process of policymakers regarding other states' intentions, and the third seeks to explain how intelligence agencies process information about intentions.

The vividness hypothesis holds that, "All else being equal, personalized information that is positive should lead observers toward a more benign assessment of the adversary's intentions, whereas personalized information that is negative should lead observers toward a more hostile assessment of the adversary's intentions" (254). Furthermore, policymakers tend to exhibit an egocentric bias, often believing that certain negative or positive actions by an adversary were a direct response to the policymaker's own actions (19).

The selective credibility hypothesis suggests that decision makers will likely vary in their prior degree of distrust toward a particular adversary and the extent to which they believe the adversary's intentions are hostile. "Specifically, those decision makers who already hold hawkish views about the adversary's intentions are unlikely to perceive its costly reassuring actions as credible signals of benign intentions, and consequently such signals will not necessarily change their perception of their adversary's intentions" (21). Decision makers who hold more hawkish views of the adversary's intentions are more likely to categorize hostile behavior as a credible signal of malign intentions.

The organizational expertise hypothesis purports to explain why policymakers and intelligence analysts favor different indicators of an adversary's intentions. Intelligence agencies have expertise in producing quantifiable analyses of other states' military arsenals and are often reluctant to adjust estimates in response to new information. In short, whereas intelligence agencies collect, analyze, and disseminate intelligence on an adversary's military capabilities and changes in such capabilities over time and use them to draw inferences about adversary's intentions, policymakers are unlikely to be receptive unless intelligence product confirms their preexisting beliefs about the adversary.

Yarhi-Milo pits her selective attention thesis against rival hypothesis from the capabilities, strategic military doctrine, and behavior theses, all of which draw upon various strands of neorealism (specifically, John Mearsheimer's offensive realism, offense-defense theories of defensive realism, and rationalist approaches to the security dilemma) and the deterrence literature.[4] The three alternative theses all posit an important role for costly signals of an adversary's intentions, whether in the form of military forces and extant weapons systems or under development (the capabilities thesis), declaratory statements of grand strategy, military doctrine, or operational war plans (the doctrine thesis), or observable behavioral patterns (the behavior theses).

The capabilities thesis holds that observers will infer intentions based on calculations of military capabilities. If the adversary already enjoys military superiority, the observer will perceive any further increase in military capabilities as evidence of hostile intentions. However, a decrease in the adversary's military capabilities will be seen as a reassuring signal, and thus evidence of benign intentions.

The military-doctrine thesis holds that observers infer intentions based on the perceived orientation of their adversary's military doctrines. Observers will see an adversary's offensive, or war-fighting nuclear doctrines will be seen as evidence of hostile intentions, whereas they will see defensive and deterrence-based doctrines as evidence of benign intentions.

Last, the behavior thesis holds that observers infer intentions based upon costly behavioral actions by the adversary. Foreign military intentions and territorial conquests, withdrawal from arms control treaties and international institutions, and the dismantling of democratic institution (where the observer is a democracy) will be seen as evidence of hostile intent, whereas withdrawal from foreign military interventions, entry into binding arms control treaties and international institutions, or the adoption of democratic institutions will be seen as evidence of benign intent. The current-actions hypothesis holds that observers infer intentions from costly actions taken by the adversary's current leadership, whereas the past-actions hypothesis holds that the observer infers intentions from costly actions taken by the adversary's past leadership.

Yarhi-Milo finds strong support for her selective attention thesis in three comparative case studies: the assessments of Prime Minister Neville Chamberlain, Foreign Secretary Antony Eden and his successor Lord Halifax, and British military intelligence about Germany's intentions between 1934 and 1939, the years between Adolf Hitler's initiation of large-

---

[4] John J. Mearsheimer, *The Tragedy of Great Power Politics*, Updated edition. ed., (New York: W.W. Norton & Company, 2014); Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30, no. 2 (1978): 167-214; Stephen Van Evera, *Causes of War: Power and the Roots of Conflict*, (Ithaca: Cornell University Press, 1999); Charles L. Glaser, *Analyzing Strategic Nuclear Policy*, (Princeton, N.J.: Princeton University Press, 1990); Charles L. Glaser, *Rational Theory of International Politics : The Logic of Competition and Cooperation*, (Princeton, N.J.: Princeton University Press, 2010).

scale rearmament and the outbreak of World War II; [5] the assessment of the Soviet Union's intentions by President Jimmy Carter, Secretary of State Cyrus Vance, National Security Advisor Zbigniew Brzezinski and the U.S. Intelligence Community between 1977 and 1980, the years leading to the Soviet invasion of Afghanistan and the final collapse of superpower *détente*; and the assessments of President Ronald Reagan, Secretary of State George Schultz, Secretary of Defense Caspar Weinberger, and the U.S. Intelligence Community between 1985 and 1988, the years between Mikhail Gorbachev's selection as General Secretary of Communist Party and the end of the Cold War. She contends the rival capabilities, strategic military doctrine, and behavior theses all fail to account for the empirical patterns observed.

Yarhi-Milo's book is a commendable attempt to address one of the more important and yet intractable phenomena in the study of international politics, namely the problem of how states—or more properly, the senior officials and the supporting bureaucracies charged with strategic intelligence, foreign policy, and defense—actually perceive external threats and make assessments about other states' intentions and likely future behavior.

In trying to discern how states make such assessments, Yarhi-Milo subjects thousands of declassified documents to three types of tests: (1) a covariance test that examines the fit between the four theses' predictions and observed changes in perceived intentions; (2) a reasoning test that considers the evidence that senior policymakers and intelligence analysts cite in support of their assessments; and (3) a policy consistency test exploring the fit between the policymaker's statements about adversary intentions and the specific diplomatic and military policies they support, as well as the evolution of policies pursued by the Chamberlain government and by the Carter and the Reagan administrations over time (39-43).

Yarhi-Milo's case selection seems appropriate and defensible. As she notes, the universe of potential cases includes all pairs of state adversaries (35). The three cases selected have differing values on the dependent variable—perceptions of the adversary's intentions— and whether that adversary becomes more or less hostile over time. She writes, "Both the British-German interaction in the 1930s and the US-Soviet relationship in the late 1970s tell a story of an evolution toward a malign interpretation of the adversary's intentions, while US-Soviet interaction in the late 1980s tracks the process by which Soviet intentions were perceived as becoming more benign" (38).

---

[5] Yarhi-Milo notes (58) that British Prime Ministers James Ramsay MacDonald (June 1929-June 1935) and Stanley Baldwin (June 1935-May 1937) were not significantly engaged in foreign policy; hence her analysis focuses largely on Chamberlain (who served as the Chancellor of the Exchequer in Baldwin's government before succeeding him as Prime Minister), Eden, and Halifax, as well as Robert Vansittart and Alexander Cadogan, who served successively as Permanent Undersecretary of State in the Foreign Office, and Eric Phillips and Neville Henderson, who served successively as the British Ambassador to Germany. One could quibble about the inclusion of the last four individuals as "policymakers." Vansittart and Cadogan were the senior civil servants in the Foreign Office, while Ambassadors Philips and Henderson were presumably in Berlin.

The cases have clearly identifiable chief policymakers—Chamberlain, Carter, and Reagan—and a number of secondary policymakers—Eden, Halifax, Vance, Brzezinski, Shultz, and Weinberger. The cases have variation on the independent variables posited by the capabilities, military doctrine, and behavior theses within and between those cases. The final case selection criterion was the accessibility of archival material. Yarhi-Milo chose three historical cases for which a broad range of archival holdings, document compilations, private papers, and oral histories available, although she acknowledges that a significant number of relevant documents from the Reagan administration are still classified (193).

In the empirical chapters, Yarhi-Milo goes to great lengths to distinguish between policymakers' and intelligence organizations' respective assessments of German and Soviet intentions, on the one hand, and the resultant policies that Great Britain and the United States pursued vis-à-vis Germany and the Soviet Union, on the other. The problems begin in trying to discern the extent to which those assessments of adversary's intentions are also influenced by: (1) the relative positions that Britain and the U.S. occupied in the international system during the 1930s, the late 1970s, and the late 1980s; and (2) the extent to which the target great power (and especially the regime) had a history of strategic rivalry (or even enmity) with Britain or the United States. I discuss each in turn.

Britain was one of several great powers in the 1934-1939 period and the only one with extensive strategic commitments in multiple geographic theaters—Western Europe (including the home islands), the Mediterranean, East Asia, and the empire. Furthermore, its material resources were quite limited given its strategic commitments and the number of potential adversaries it faced in different theaters, namely Germany (in Europe) and Japan (in East Asia), and to a lesser extent Italy (in the Mediterranean), in addition to growing nationalist movements in various parts of the British Empire, such as India, Egypt, and Palestine. In the mid-1930s, therefore, cabinet ministers, diplomats, military officers, and intelligence analysts in Whitehall faced the daunting task of assessing which components of potential adversaries' capabilities were more (or less) threatening to British interests, the time horizons in which those threats would become manifest, and medium to long-term intentions of potential adversaries and potential allies.[6]

The United States during both periods under examination (1977-1980 and 1985-1988) was one of two superpowers. Like Britain in the 1930s, the U.S. had extensive strategic commitments in multiple theaters, including a network of overseas military bases, and various bilateral and multilateral alliances. However, unlike Britain in the 1930s, the U.S. had one primary adversary, the Soviet Union. It also had a sufficient resource base to meet its strategic commitments, even in the wake of seemingly adverse international trends, such as the Soviet acquisition of strategic parity in nuclear warheads and intercontinental

---

[6] Steven E. Lobell, "Bringing Balancing Back In: Britain's Targeted Balancing, 1936–1939," *Journal of Strategic Studies* 35, no. 6 (2012): 747-773. See also Steven E. Lobell, "Britain's Grand Strategy During the 1930s : From Balance of Power to Components of Power," in *The Challenge of Grand Strategy : The Great Powers and the Broken Balance between the World Wars*, ed. Jeffrey W. Taliaferro, Norrin M. Ripsman, and Steven E. Lobell (New York: Cambridge University Press, 2012), 147-170.

ballistic missiles (ICBMs) in the late 1960s and defeat in the Vietnam War, as well as domestic economic constraints, namely the high inflation and the recessions of the 1970s. By the mid-1980s, the United States' relative position in the international system had improved, thanks to the rebound on the U.S. economy and the massive buildup in conventional and nuclear military capabilities (which actually began under the Carter administration in FY 1980 and accelerated under the Reagan administration) and the Soviet Union's concurrent slide into deep and inevitable (and ultimately terminal) relative decline, especially in economic capabilities.[7]

It is notable that in both the British-German interaction of the 1930s and the U.S.-Soviet interaction of the late 1970s, "an evolution toward a malign interpretation of the adversary's intentions" (38) appears to track adverse shifts in the balance-of-power and resulting "windows of vulnerability" for Britain and the U.S. vis-à-vis their respective great power adversaries, Nazi Germany and the USSR. Conversely, the "the process by which Soviet intentions were perceived as becoming more benign" by the Reagan administration in the late 1980s tracks an objective power shift in favor of the U.S. and against the Soviet Union (38). This is not to dispute Yarhi-Milo's argument that British and U.S. policymakers relied heavily on information they considered vivid, salient, and personal in assessing the intentions of counterparts in Berlin and Moscow. Rather, I merely suggest that is more difficult to disentangle elite assessments of enemy intentions from elite assessments of the underlying relative power trends between adversaries.

The second observation is that all three of Yarhi-Milo's cases arguably involve situations in which policymakers and intelligence analysts in Washington and London confronted "known unknowns," specifically, the task of assessing the likely future intentions of long-term adversaries.[8] By 1934, Germany and Britain had a history of rivalry dating back to Kaiser Wilhelm II and his *Weltpolitik* in the 1890s. World War I was one only a dozen years past. The extent of Hitler's territorial ambitions and willingness to risk (another) major war in pursuit of them were unknowns for British ministers and intelligence analysts; but Germany's dissatisfaction with the Versailles Treaty and willingness to challenge its provisions were known. Likewise, by 1977, the Carter administration's first year in office, the Cold War was more than thirty-years old. Moreover, unlike Hitler in 1933, who had risen from being the leader of an obscure radical party in Bavaria to being the newly

---

[7] For assessments of relative power trends in the last two decades of the Cold War see Stephen G. Brooks and William C. Wohlforth, "Power, Globalization, and the End of the Cold War: Re-Evaluating a Landmark Case for Ideas," *International Security* 25, no. 3 (2000): 5-53; Stephen G. Brooks and William C. Wohlforth, *World out of Balance: International Relations and the Challenge of American Primacy*, (Princeton: Princeton University Press, 2008).

[8] This is a reference to former Secretary of Defense Donald Rumsfeld's answer to reporter's question about Iraq's suspected weapons of mass destruction (WMD) at press conference in February 2002. The full quotation is ""There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know." DoD News Briefing—Secretary Rumsfeld and Gen. Meyers, 12 February 2002 http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636 retrieved 20 November 2014.

appointed Reich chancellor in a decade, by 1977 Leonid Brezhnev had been CPSU general secretary (and as such, the paramount leader of the USSR) for thirteen years. Brezhnev was hardly a *tabula rasa*. In 1985, Gorbachev was more an enigma for foreign observers. On the one hand, he was an *apparatchik*, who had been socialized and groomed for leadership within an authoritarian one-party state. On the other, Gorbachev (then aged 56) was a full generation younger that his three immediate predecessors—Brezhnev, Konstantin Chernenko, and Yuri Andropov—and the first CPSU general secretary to have born after the 1917 Bolshevik Revolution. In future research, it would be interesting to test Yarhi-Milo's selective attention thesis and the alternative behavior theses against cases where policymakers and intelligence analysts make assessments about the intentions of actors posing novel or unprecedented types of security threats, such as the leaders of transnational terrorist networks.

One of Yarhi-Milo's findings is that British and U.S. policymakers, with the notable exception of Brzezinski (a widely published international relations scholar both before and after his government service), did not explicitly identify the indicators that shaped and reshaped their perceptions of German and Soviet intentions and relative capabilities. Furthermore, even within the same administration, different policymakers attached different weight to particular indicators of an adversary's capabilities and intentions. While this finding might seem to call into question the capabilities and the behavior theses, upon further reflection, it is not that surprising, in part because the types of politicians who become presidents, secretaries of state and defense, and even the national security adviser to the president (in the U.S.) or prime ministers, chancellors of the Exchequer, secretaries of state, and other cabinet ministers (in the UK) tend to be generalists.

In his study of British officials' assessments of relative decline at the dawn of the twentieth century, Aaron Friedberg observes:

> Power cannot be tested; different elements of power possess different utilities at different times; the relationship of perceived power to material resources can be capricious; the mechanics of power are surrounded by uncertainty; states possess different conversion ratios and comparative advantages; the perceived prestige hierarchy and the military distribution may not coincide for prolonged periods; states adopt asymmetrical strategies to maximize their positions and undercut rivals; signals get confused among allies, rivals, and domestic audiences.[9]

Friedberg's observations about the practical difficulties that officials in Whitehall in the 1895-1905 period encountered in assessing the balance-of-power and anticipating power trends is equally applicable to difficulties their successors encountered in assessing the long-term intentions of Hitler in the 1930s and that their (future) counterparts in the Carter and Reagan administrations and U.S. intelligence analysts encountered in assessing

---

[9] Aaron L. Friedberg, *The Weary Titan: Britain and the Experience of Relative Decline, 1895-1905*, (Princeton: Princeton University Press, 1988), 308.

the long-term intentions of Soviet general secretaries Brezhnev and Gorbachev in the late 1970s and late 1980s, respectively.

Despite the above critiques, I reiterate that *Knowing the Adversary* is a superb book. Yarhi-Milo advances the field in its theoretical understanding of how policymakers and intelligence agencies assess the intentions of adversary states.

Author's Response by Keren Yarhi-Milo, Princeton University

I am immensely grateful to Jack Levy, David Edelstein, Jennifer Sims, and Jeffrey Taliaferro for their deeply insightful and thoughtful assessments of my book, *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. In their wide-ranging and expert critiques, they have raised issues that I grappled with throughout the process of writing the book, and that have provided much intellectual fodder for a future research agenda on the themes therein.

As they have each noted, the book struggles with a handful of big questions pertaining to international relations theory, psychological explanations, and intelligence analysis that have overhung the study of international politics since Robert Jervis's seminal works on the logic of images and perception and misperception in international politics.[1] The theoretical scale of the project called for a number of difficult decisions when it came to the empirical strategy, in terms of case selection, the operationalization of variables, and the analysis of alternative explanations. In the book, I attempted to elucidate the logic behind each of these decisions, so that researchers who seek to replicate or extend my analysis might fruitfully explore alternative strategies in order to address the overarching research questions and arrive at similar or different conclusions.

The specific critiques offered by the reviewers are tremendously valuable as signposts for further research. Below, however, I respond to four broad categories of issues that I think deserve some clarifications and more thought: definitions; time, capabilities, and uncertainty; intentions and policies; and case selection.

The first set of issues pertains to definitions. Particularly important in this category is my definition of the dependent variable, which is the intentions of the adversary state. I define the adversary's intentions as "foreign policy goals (political intentions) with regard to changing or maintaining the status quo, and to a lesser extent, about the adversary's inclination to use military force to achieve these objectives (military intentions)" (14). Jennifer Sims raises the valid point that states—like individuals—often themselves may not know what they intend to do in the long term, which is the focus of my analysis. However, this fact alone does not preclude observers in other states from drawing conclusions about the intentions of a state. The theory I offer is one of how state A's beliefs about state B's intentions are formed and change over time; it is agnostic in terms of whether A's beliefs are accurate, or whether B is even aware of its own intentions. Moreover, the theory deals with intentions to do a particular thing, which is to revise or maintain the status quo in the long term (say, a period of five to ten years); and the means that states are likely to use in order to achieve such goals. While such intentions may undoubtedly change over time, my theory attempts to track how in the presence of uncertainty the *assessments* of these intentions—correct or incorrect—are formed, and why they change over time.

---

[1] Robert Jervis. *The Logic of Images in International Relations* (Columbia University Press, 1970); and *Perception and Misperception in International Politics* (Princeton University Press, 1976).

Sims's other definitional points are also spot on. While it is theoretically possible to distinguish clearly between intelligence analysts and decision makers, the empirical waters are considerably murkier. Particularly in the case of decision makers who may have been trained and socialized in the culture of intelligence agencies—a point also raised by Jack Levy—there may exist cognitive and organizational cultural factors that offset the expectation that decision-makers act in accordance with the vividness and subjective credibility hypotheses. Empirically, decision-makers will not always act according to these hypotheses, and there may be certain decision-makers such as the current president of Russia, Vladimir Putin, former prime minister of Israel, Ehud Barak, and former U.S. president George H. W. Bush who are more in tune with the assessment style of intelligence agencies due to their backgrounds. This insight offers an important avenue for further research into more fine-grained distinctions between decision-makers in relation to their assessments of the adversary. One might envision, for example, a study that seeks to survey decision-makers in order to unearth the individual-level correlates of different types of assessments.

David Edelstein offers three important dimensions along which one might further qualify the theory I present in the book. The first is the time horizon on which intentions are being judged. The length of the time horizon is likely to impact the nature of the assessments a state's leaders make of another state. With a focus on a specific set of intentions over a specific timeframe, i.e. intentions regarding the status quo over a period of five to ten years, I find that decision-makers tend to use one set of metrics and intelligence analysts another. It is entirely possible that by varying the timeframe under consideration we might find other metrics (such as behavioral indices) in play, and this again would be a way to extend the theory presented in the book. Still, the empirical analysis in the book attempts to clearly delineate between assessments in the short run (e.g. what Soviet leader Mikhail Gorbachev was looking to do in the next round of negotiations), which are not the focus of this book, and how his long term political plans are perceived (e.g. was Gorbachev seeking to 'buy time' in order to get stronger and pursue more aggressive policies down the line). At times, assessments about the former will shape one's views about the latter (and vice versa), but the empirical chapters focus exclusively on teasing out the indicators used to infer *long-term* political intentions.

The second dimension is capabilities, and this is an important issue that I wrestled with in writing the book. Edelstein and Jeffrey Taliaferro argue separately that while changing capabilities may not have a direct impact on the assessments of decision-makers, they are likely to indirectly influence their assessments of intentions. Capabilities are certainly an important aspect of any assessment of intentions; however, they are not the only aspect and often not even the most important aspect, depending on who is conducting the assessment. The theory I offer therefore makes two points with regard to capabilities. First, while assessments about the adversary's capabilities are viewed as important by decision-makers in forming decisions about defense spending and force structures, the empirical studies in this book clearly reveal that there are a number of non-costly signals (or 'cheap talk') such as face-to-face meetings, salient events, etc., that decision-makers consider

when assessing the long term political intentions of the adversary. It may be the case that in assessment of intentions or resolve during crises, as Daryl Press's work has shown, decision-makers reason by referring to the adversary's capabilities and interests.[2] The absence of such reasoning evidence among decision-makers even in cases where the adversary's capabilities were significantly increasing (such as in the case of interwar Germany) to support or explain their assessments of political intentions is something I did not expect to find when I started this research. To the extent that decision-makers were using such reasoning explicitly and repeatedly in their crisis assessments (as Press documents), the argument that "capabilities are something they think about but do not use explicitly" when judging long-term political intentions becomes less convincing.

At a minimum, the absence of strong evidence supporting the 'capabilities thesis' for decision-makers should indicate that even if they believed in the strong relationship between capabilities and intentions, they did not think that others would be convinced by such reasoning. As I note in the book, National Security Adviser Zbigniew Brzezinski repeatedly tried to convince President Jimmy Carter that the Soviets had expansionist designs. The fact that Brzezinski very rarely referred to the significant Soviet buildup in his reasoning (and in fact explicitly argued that capabilities were not an adequate indicator to discern political intentions) indicates at the very least that he did not think Carter would find this reasoning compelling. The second point about capabilities is that the importance accorded to capabilities in assessments depends on who is doing the assessing. In terms of the theory, decision-makers are less likely to focus on capabilities than are intelligence analysts for all the reasons I explain in the book. I therefore fully agree with Edelstein and Taliaferro that capabilities matter, and would qualify their argument by echoing Levy when he writes that "capabilities are necessary to support intentions but not sufficient to determine them." Although the empirical cases in my book contain significant variation over time in the perceived military capabilities of the adversary, future work can look at a slightly different universe of cases – rising and declining power – in order to maximize variation in this independent variable. Such cases could serve as 'easy' tests for this thesis, and we should expect to observe, if this thesis is correct, important reasoning evidence linking such trends to changes in political assessments about long term intentions.

The impact of uncertainty on assessments is another important qualifier and potential extension of my theory. Indeed, the book's very premise is that the task of gauging intentions, especially long-term political intentions, is so difficult because of uncertainty about things like the other side's determination to pursue its goals, its incentives to misrepresent, the ability to distinguish between signals and noise, etc. Another aspect of uncertainty relates to differences in the level of uncertainty attached to intelligence estimates, which is a vital facet of what Sims calls "variances in how intelligence works." Thus decision-makers may be less willing to accept a capabilities-based intelligence estimate for reasons more to do with the reliability of its source than any cognitive or

---

[2] Daryl G. Press, *Calculating Credibility: How Leaders Assess Military Threats* (Cornell University Press, 2007).

organizational factors. This point also raises questions about the diverse methods by which decision-makers obtain intelligence estimates of an adversary, with joint estimates being one among many sources. No doubt this is the case, and working in the domain of declassified documents is, as Sims notes, an intellectually fraught enterprise precisely for this reason. Therefore the best one can do in such circumstances is to offer the analysis in the book with an important caveat, which is that it is based on available sources that may be subject to selection bias. I attempted to mitigate such bias by conducting interviews with intelligence officers and decision-makers, triangulating primary and secondary evidence, and reviewing available assessments by particular intelligence agencies. Future work could probably use the methods suggested by Sims that measure a decision-maker's attentiveness toward a particular intelligence product, method, or sourcing. Such an approach would help address important questions that are beyond the scope of this book, such as the conditions under which intelligence methods and sourcing will be more or less likely reduce international tensions, and when they are more or less likely to be trusted by decision-makers. Nonetheless, even our best methods are likely to only partially address selection bias due to the non-random scarcity of data that characterizes this field of inquiry.

Edelstein raises an important question regarding the translation of assessments into policy. This is an area that is ripe for future research. Given the existing parameters of the book, policy outcomes are in a sense over-determined. It might be that assessments of intentions alone can account for certain policy choices – for instance, in the empirical analysis I show that major changes in assessments do result in policy changes (and that some major shifts in policies took place and that decision makers who supported these also referred to their assessments of intentions). In other cases, it might be that capabilities, time horizons, and degrees of uncertainty all interact with assessments to produce policy outcomes. And of course, there is a host of domestic and external factors that might constrain the translation of assessments into policy. These myriad factors would stretch the scope of inquiry well beyond the limits of a single volume. Future work in the area of linking assessments to policy outcomes in the future would constitute a major step forward for the research agenda on perception and signaling in international politics.

Each of the reviewers have highlighted the issue of case selection, and rightly so. Given that the universe of potential cases is vast, in order to make my project manageable I chose cases in a way that would control for certain key factors while providing variation on the independent and dependent variables. More important criteria for selecting the cases are documented at length in the book (38-40).

While these choices necessarily entailed a trade-off with generalizability, even a strict constraint on generalizing my findings to the set of democracies assessing non-democratic adversaries covers a large number of cases. Scholars who are interested in replicating my argument can therefore test it against other similar cases. There is also a great deal that scholars wishing to extend and qualify the arguments made in the book can do by extending its empirical focus to non-democracies assessing democracies, democracies assessing democracies, non-democracies assessing non-democracies, and each such type of state assessing new threats such as terrorist networks. We might theorize that autocrats

are likely to make assessments based on vivid indicators and trust their 'gut feeling' more so than democratic decision-makers because they view themselves as more capable and experienced in sizing up their opponents, and more powerful than other actors (such as intelligence agencies) in their regimes. Iraqi leader Saddam Hussein, for example, thought of himself as more capable of assessing American intentions than his intelligence officers. While evaluating whether America would decide to invade Iraq in a meeting with his inner circle in October 1990, he said, "The decision is American and requires being alert as politicians even before the intelligence community…I don't want [the Iraqi intelligence agencies] to give me analysis; that is my specialty. I told them they should only give me news so that they don't get distracted and cover their failure with news they don't understand."[3]

We might also look at how decision-makers and intelligence analysts assess the intentions of allies – I suspect that the selective attention hypothesis might still hold, but the variables would take on different meanings given a different background assumption of trust compared to rivalry. These are all likely to be tremendously fruitful avenues for future research, and I hope that my book might contribute in some way to the growing research in that area.

Finally, I second Levy's suggestion of focusing more closely on motivated bias and its effect on assessments of intentions. Theoretically, motivated bias can be integrated into the vividness hypotheses, but as Levy correctly points out, the empirical analysis, at least in the case of Britain and Germany in the 1930s, does not seem to support the proposition that a political leader's preferences for a particular policy option biases their view of events and information related to that policy. One way to incorporate motivated biases into the vividness hypothesis in a testable manner might be to think theoretically about the conditions under which motivated bias should be strongly/weakly associated with assessments about intentions or about the type of individuals who are most likely to be prone to motivated bias in their decision-making. This line of hypothesizing would certainly be very helpful for further refining the argument of my book.

Going forward, I hope that the book will be seen as serving three major purposes. First, it furthers the theoretical debate over long-term intentions as opposed to threat or short-term crisis management. Second, it questions the manner in which international-relations scholars think about diagnostic indicators and the metrics states use to assess each other's intentions and 'types.' In particular, it calls in to question the dominant paradigm of costly signaling to show that certain types of non-costly signaling can carry significant weight. Finally, the book opens the black box of intelligence analysis and the manner in which it feeds into the decision-making process in a way that bridges research in intelligence with broader international relations theorizing.

---

[3] K.M. Woods, D. Palkki, and M. E. Stout, *The Saddam Tapes: The Inner Workings of a Tyrant's Regime, 1978-2001* (New York: Cambridge University Press, 2011), 36.

I am grateful to the reviewers for suggesting important qualifications and extensions to my arguments that will aid in this process.