Predispositions and Foreign Policy Surprises:
Assessing the Impact of Rational and Biased Beliefs on Strategic Decision-Making

by

Scott A. Helfstein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Public Policy and Political Science)
in The University of Michigan
2008

Doctoral Committee:

Professor Robert Axelrod, Chair
Professor Kenneth G. Lieberthal
Professor James D. Morrow
Associate Research Scientist Scott Atran

Acknowledgements

I would like to thank my committee members Scott Atran, Kenneth Lieberthal, James Morrow, and particularly Robert Axelrod for their guidance and constructive criticism. I would also like to thank Steven Drucker, Aleksandar Matovski and Dominick Wright for their helpful suggestions on many chapters. Parts of this project have been presented at the 2007 International Studies Association Annual Meeting, the 2007 Peace Science Society Annual Meeting, and the NIS Seminar at the University Michigan. I would like to thank all of the participants at those meetings.

Table of Contents

List of Figures

List of Tables

Abstract

This project analyzes the relationship between foreign-policy decision-making, surprise and misjudgment using a multi-method approach. In recent years, high profile foreign policy and intelligence failures have captured the attention of policymakers, the media and academics alike. Many retrospective studies of policy failure identify, implicitly or explicitly, psychological biases of pertinent decision-makers as contributing to poor outcomes. By focusing on the internal mechanics of information usage and decision processes, these studies draw from bureaucratic politics theory. An alternative paradigm used to explain state behavior, rational actor theory, focuses on the competition between states and often assumes that the bureaucratic process works perfect. Since each explanation stresses different aspects of the decision-making process, it is difficult to know whether conclusions derived with one paradigm are transportable to the other. The bureaucratic politics explanation often stresses the negative relationship between psychological biases and outcomes, whereas rational actor theory illustrated the benefits of feigning irrationality. This thesis improves our understanding of psychological biases in competitive interstate interactions by incorporating the possibility of misjudgment into a formal model of surprise attack. The project compares an efficient form of learning, Bayesian updating, to an inefficient learning mechanism whereby actors irrationally overweigh their prior beliefs relative to new information. This project then identifies conditions where irrationally entrenched beliefs may be overemphasized as a reason for

policy failure, and situations where irrational over-suspicion can offer competitive advantage by altering another player's behavior. The deductive analysis suggests that psychological biases may be over-used as an explanatory factor in retrospective studies of failure, that psychological bias is neither a necessary or sufficient condition for surprise, and that irrational over-suspicion can have a stabilizing impact on a hostile interaction. These deductive propositions are examined quantitatively using a dataset of international crises from 1918-2002, which suggests that irrationally entrenched beliefs may reduce the incentive to attempt surprise attack. This is followed by a qualitative analysis of the 1962 Cuban Missile Crisis and 1967 Arab-Israeli War, showing that the model with a potentially biased actor seems to outperform the model with perfectly rational actors.

Chapter 1
Introduction

This project is about decision-making and belief biases in the competitive international context, and specifically focuses on the relationship between foreign policy surprises and psychological biases. In the midst of the post-behavioral movement, Robert Jervis (1976) wrote his now canonical book about the role of psychology in foreign policy decision-making. The book was a departure from the dominant realist-liberal debates of the time. Jervis argued that psychology and the nature of decision-making played a central role in international affairs. The first part of the book addressed the importance of perception or beliefs in strategic situations, and later examined how psychological dynamics impact perceptions. Since publication, there have been a number of attempts to refine and expand our understanding of psychology and international affairs, however many of these works focus on decision-making in an internal bureaucratic context rather than the external competitive pressures inherent in international politics. Jervis' book arguably remains the best blending of strategic theory and psychology.

In the introductory lines of the final chapter, Jervis clarifies his position and offers some advice on psychological biases in foreign-policy decision-making:

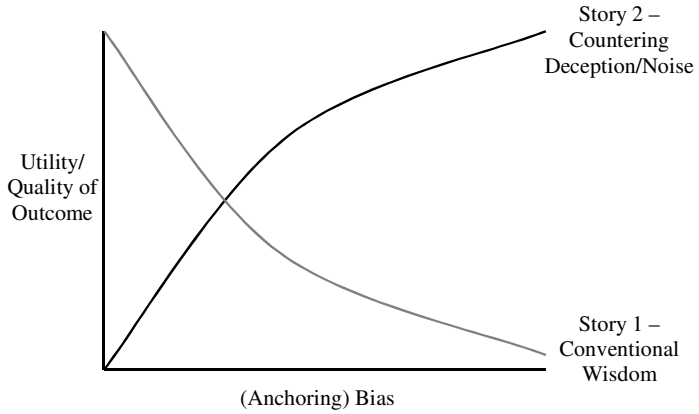
> Although prescription is not the purpose of this book, our analysis suggests a number of measures that would decrease misperception. If decision-makers become aware of common perceptual errors, they may be able to compensate for them. And they can

> adopt safeguards to decrease their unwarranted confidence in prevailing beliefs, make them more sensitive to alternative explanations and images, and thus decrease the amount of discrepant information needed to make them reexamine their views.

Jervis lays out a clear causal chain, which provides one possible perspective on the relationship between psychological biases and policymaking. Though incorrect beliefs cannot be avoided altogether, "common perceptual errors" and the pressures of the international environment lead to misperceptions. These misperceptions lead to faulty conclusions and poor policy. In many ways, Jervis echoes the sentiments of economists and game theorists stressing the importance of accurate beliefs and optimal strategies. However, one is tempted to ask whether psychological biases, specifically the deeply entrenched beliefs he refers to, can actually be beneficial in the competitive context of international politics where surprise and intelligence failures are endemic.

The question above drives the research program that follows. Conventional arguments in the study of foreign policy decision-making and national intelligence stress the negative attributes of psychological processing. Decision-making biases, cognitive biases and bounded rationality all interfere with efficient learning. It follows that this interference is bad and leads to inaccurate images. When these inaccurate images are used as the foundation for policy, the policy is more likely to fail. This argument is convincing, but there are also some counter arguments that have been marginalized due to the dominance of economic theory in political science and world politics. Figure 1.1 illustrates the hypothesized role between belief bias and policy outcomes. The conventional argument, articulated by Jervis represented by Story 1, is the dominant paradigm, but this project is interested in exploring cases where this theory is violated in systemic fashion.

Figure 1.1: Relationships Between Psychological Biases and Outcomes



Summary: There are reasons to believe that actors are worse off when they exhibit anchoring bias and reasons to believe they might be better off. These different hypotheses are examined throughout the project.

Decisions in international politics almost universally involve uncertainty and information constraints. Decision-makers rarely know all of the relevant history, or what might be called the true state of the world. The information constraints are certainly not limited to past events since the results of policies or decisions play out in complicated ways. Ambiguity is a central feature of international politics, it plays large role in decision-making, and opponents often try to manipulate information. Foreign policy decision-makers and intelligence analysts often have to interpret new information using their prior beliefs as a backdrop. When these actors irrationally overweight the importance or quality of their prior beliefs relative to new information, they are suffering from anchoring bias. It is rarely, if ever clear, however, how firmly policymakers should hold to their prior beliefs in the presence of contradictory evidence. Uncertainty and the use of prior beliefs makes it difficult to identify psychological biases, making it especially important to distinguish skeptical learning that is rational from anchoring that is irrational. This is essential when dealing with surprise and deception.

Deception has long played a role in international affairs and war, going at least as far back as the Trojan Horse. One dominant theory of war outbreak stresses the importance of countries' incentive to misrepresent their true interests in an attempt to obtain a better bargained outcome. When countries' misrepresent themselves, there is higher probability that bargaining breaks down and they go to war. Once countries choose to initiate hostilities, they may rely on deception in order to catch their opponents by surprise. Many successful cases of diplomatic or military surprise involved deception operations that cultivated certain beliefs in an opponent. Though a fringe argument, it has been suggested that biased beliefs can improve one's immunity to deception operations. If attempts at deception will be missed or ignored, there is less incentive to utilize that strategy. Just as prior beliefs provide a context for interpreting ambiguous information, it is also important to examine when or how biased beliefs alter the deceptive-character of an opponent's strategy. In short, can psychological biases like entrenched beliefs increase honest communication and deter attempts at surprise?

The Bureaucratic and Strategic Models

This project takes a thorough look at the relationship between misjudgment, surprise and policy failure, integrating two different causal models: the strategic model and the bureaucratic model. The strategic model focuses on the competitive context and information constraints prevalent in international politics, usually concluding that strategic surprise cannot be eliminated. The bureaucratic model focuses on the organization's performance and use of available information, often concluding that

surprise is a problem that can be minimized through specific remedies or policy reforms. Figure 1.2 provides a simplified representation of these two models.

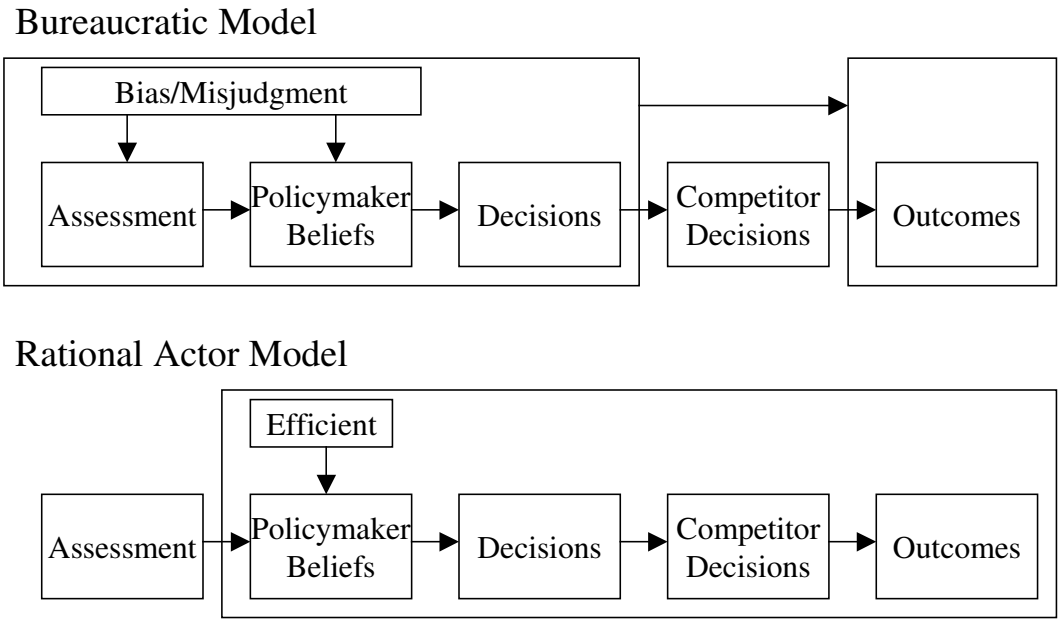


Figure 1.2: Models of Foreign Policy Decision-making
Summary: The bureaucratic model often focuses on how internal issues like group dynamics, assessment process and psychological biases, impact decisions. Rational actor models focus on the competitive processes and often assume that assessment and learning is perfectly rational.

The strategic model of surprise usually assumes that the bureaucratic process works perfectly and assumes away misjudgment. By contrast, the bureaucratic model identifies bureaucratic process and misjudgment as major causes of surprise. These two models tell different stories. Currently, there are few analytically rigorous attempts to sort through the different assumptions, integrate the models and explore the impact of misjudgment in competitive contexts. It is important to ascertain if this lacuna is a critical shortcoming in our understanding of foreign policy and intelligence failure.

This project focuses on the treatment of misjudgment and deeply entrenched beliefs, which are a key difference between the two models. Beliefs about an opponent's

intentions and capabilities play a significant role in many explanations of foreign policy decision-making. Deeply entrenched beliefs inhibit learning from new information and usually lead to inaccurate threat assessments. These assessments are subsequently cited as major reasons for foreign policy and intelligence failures. The bureaucratic model focuses on the assessment process and the impact of inaccurate beliefs on decisions, without explicitly considering the impact of opponents' decisions on outcomes. The strategic model often assumes that assessments and beliefs are objectively accurate, and it focuses on the way optimal decisions interact with opponents' decisions to produce outcomes. By incorporating deeply entrenched beliefs into a strategic model, we are able to reach a more comprehensive understanding of misjudgment and surprise.

The Rationality of Irrationality

Robert Keohane once observed that Thomas Schelling articulated almost every major idea in international relations first (Scott 2007). This projects echoes Keohane sentiment, recognizing Schelling's legacy and the applicability of his work here. During the 1950s and 1960s strategists were trying to adjust to a world with nuclear weapons and the possibility of massive retaliation. For U.S. strategists, the pressing question was how best to contain, or deter, the Soviets in a world were massive retaliation was possible and nuclear war could kill everyone. Schelling laid out a set of arguments that helped guide U.S. policy throughout the Cold War.

One of Schelling's (1960) most notable arguments is referred to as the rationality of irrationality, whereby appearing irrational in a competitive environment could actually be optimal. The argument relied on commitments and the need for commitments to be

credible. One of his basic examples involved a buyer and seller who are trying to transact at one of two prices: low and high. The buyer prefers the low price and the seller the high price. If both players want to complete the transaction at either price, a player that can commit themselves to their preferred price may draw the other into an undesired strategy. Schelling imagines a world were declaring "cross my heart" is a binding commitment. In this world, the buyer can benefit by conveying that he will only pay the low price, and then bind himself to that strategy by saying "cross my heart." The seller then has no choice but to adopt the low price strategy if they want to make the sale, despite a preference for the higher price. One common conceptualization is two drivers playing chicken. One driver throws the steering wheel out the window and the other has no choice but to swerve.

Schelling's real concern was figuring out how the U.S. could use commitment strategies to deter Soviet aggression, particularly against West Berlin, given the possibility that any punishments could lead to nuclear war. While articulating the benefits of commitment, he also recognized that commitments must be credible to be successful. The U.S. could not commit to protect West Berlin with massive retaliation by declaring "cross my heart." If massive nuclear retaliation was an irrational action, as it appeared to be given the possibility of all out nuclear war, then the U.S. commitment to protect Berlin was not credible and of no concern to the Soviets. In more modern economic terms, it would be cheap talk. While Schelling maintained that one possible strategy was appearing irrational, he outlined two strategies whereby a rational actor could seem irrational thereby justifying massive retaliation: burning bridges and "the threat that leaves something to chance."

Burning bridges referred to a strategy that made inaction in the face of aggression impossible such that an actor could justify carrying out their threat. The example he offered was an advancing army who burned bridges to ensure they could not retreat. Within the context of West Berlin, the U.S. tried to credibly commit itself by stationing a contingent of U.S. forces in West Berlin. Should the Soviets break containment and try to take Berlin by force, the U.S. Berlin Brigade would be involved in the ensuing battle. The engagement of U.S. forces, and inevitable casualties, ensured that the U.S. would have to respond.

Schelling's other alternative, the threat that leaves something to chance, relies on randomization of action and expected payoffs in utility theory. He argued that a seemingly irrational act could be justified if it were one of many possible outcomes that are selected based on some probability. Randomization " has been a device to make indivisible objects divisible, or incommensurate objects homogenous. Their 'expected values' are divisible when the objects themselves are not" (Schelling 1960). He argues that firm commitments that involve massive retaliation may not be credible, but randomized threats could be rational based on expect utility calculations. If there is a small chance that government bureaucracy or mechanics randomize outcomes, and one possible outcome is nuclear retaliation, then the threat may or may not be rationally credible.

Schelling's derivation of rational strategies to credibly signal an actor's seemingly irrational commitments, the rationality of irrationality, remains a powerful tenet in international politics and deterrence.

Project Conclusions and Schelling's Irrationality

This project is best considered a reformulation of Schelling's theory that allows both qualitative revision and substantive expansion. Some of the conclusions from this project are substantively similar to Schelling's main conclusions about the rationality of irrationality, but the mechanisms at work are qualitatively different.

There are three main propositions derived from the formal modeling part of the project. The conclusions are derived from comparing a model of surprise attack with perfectly rational actors to an identical model where the target may irrationally anchor on prior beliefs and the initiator knows the probability with which they face a rational or biased target. Anchoring bias or entrenched beliefs, can result in either irrational under-suspicion, where a target believes the initiator is less threatening than they actually are, or irrational over-suspicion, where the target believes the initiator is more threatening than they actually are. There are three primary deductive conclusions: 1) irrational under-suspicion does not increase the probability that a target is surprised since the initiator will attempt surprise as much as possible anyway; 2) the possibility of irrational over-suspicion should deter an initiator from attempting surprise since target over-suspicion reduces the likelihood that the attempted surprise is effective; and 3) targets can derive utility benefits from their own irrational over-suspicion when the likelihood of suffering unnecessary defensive action (false alarms) is sufficiently small.

Many of these propositions bear significant resemblance to Schelling's work, particularly the argument that irrationality can be beneficial and can deter other's hostile actions. This project offers some qualitative differences and extensions. First, Zagare and Kilgour (2000) write that Schelling's argument "turns on the denial of procedural, but not

instrumental, rationality" (Zagare and Kilgour 2000). Procedural rationality is the everyday conception of rational preferences or behavior. Instrumental rationality is a much more narrow concept. Instrumental rationality is an economist's definition of rationality dictating that actors can compare any two possible outcomes to determine a relative preference and that the preference ordering be transitive (i.e. if an actor prefers option *a* to *b* and option *b* to *c*, they cannot prefer option *c* to *a*). By this definition, Schelling's actor is instrumentally irrational despite trying to appear procedurally irrational. Zagare and Kilgour (2000) go on to note that Schelling's player *feigns* procedural irrationality because it is in their interests.

In this project, the player does not feign irrationality for strategic purposes, but may actually have an irrational anchoring bias. Further, the target in this project does not even know whether they have biases that make them irrational. Irrationality, therefore, is not a deliberate strategic decision. This is distinct from Schelling's actors that rationally feign irrationality for their own benefit. This suggests empirically observed psychological biases that naturally impact everyone may have the same effects as strategically feigning irrationality.

The anchoring bias modeled here and Schelling's feigned irrationality yield similar results when 1) anchoring bias results in irrational over-suspicion, and 2) the initiator knows the target is biased. Here, however, we come to another important revision and extension. This project relaxes the assumption that one player knows that the other will play irrationally. While Schelling showed that seemingly irrational commitments could be made credible by randomization or bridge burning, this project shows that the mere possibility of irrational over-suspicion (which everyone displays

sometimes) is sufficient to capitalize on the benefits irrationality. The initiator need not know whether the target is irrationally over-suspicious or not. The mere possibility that the target is irrational is enough for the initiator to change their strategy. An initiator should adopt the same strategy irrespective of whether the target is always irrational or occasionally irrational. This means that the possibility of irrationality is sufficient to derive any possible benefits from feigning irrationality or actually having irrationally anchored beliefs. Schelling's actor relied on in the randomization in the final move of play to make the first stage commitment credible all of the time, but here, the randomization between rational and anchored targets is sufficient to alter an opponent's strategy irrespective of stages or credibility.

The possibility that target could be either a rational or biased type[1] also creates a caveat to the benefits of actual irrationality. Since Schelling's actor was feigning irrationality, there was no concern that the actor might be excessively irrational. Here, a target can benefit from their bias as long as the possibility they are biased is below a threshold value. Once the probability that the target is actually biased (as opposed to being rational) gets too high, the likelihood of unnecessary defensive action (false alarm) gets sufficiently high that it cancels out the strategic benefits from irrationality. This deduction, combined with the result immediately above, means that a target is best off having some probability of being irrational, but not being irrational too often.

The model also extends Schelling's insights by both applying it to surprise attack, and deriving some interesting conclusions about psychological bias in surprise. One

[1] Note that type is game theory term used commonly used when one player is uncertain about the preferences of another. The different preference profiles of an actor correspond to different types. Here, type is also used to distinguish between a target that could be rational or irrational whenever there is some possibility of irrationality. This notion of type is consistent with the usage in game theory since there are fixed preferences to play certain strategies even when the target is irrational.

conclusion about under-suspicion was noted above: irrational under-suspicion does not increase the probability that a target is surprised since the initiator will attempt surprise as much as possible anyway. This differs from many of the arguments made about psychological biases and surprise (Jervis 1976; Betts 1978; Heuer Jr. 1981; Kam 1988). The extension to surprise also shows that feigning irrationality or being irrational may deter surprises as well as challenges, the primary difference being the opaque information and the possibility of deception often associated with surprise attacks. The challenges that Schelling addressed were public challenges, like a Soviet confrontation over Berlin, rather than focusing on the type of confrontation such as surprise or diplomatic act.

A final deduction from the model, also tied to surprise, addresses the benefits and drawbacks of both rationality and irrational over-suspicion. Just as Schelling did, this project shows that actors could benefit from irrationality. There were some additional interesting conclusions based on the possibility that targets could be rational or irrational in any given interaction. Recall from above that the initiator, however, always acted as though the target was biased even when there was only some probability the target was an irrational type (where type represents their method of information updating). In this instance, there are some benefits to being rational and some benefits to being biased. Technically, the rational type has a higher utility. Since the initiator attempts surprise less often than they would when making a rational type indifferent between waiting and defensive action, the rational type benefits from less deception. The downside for the rational type is that they always wait, and they suffer surprise very often as a result. The biased type mixes defensive action and waiting, thereby suffering false alarms too often. While the rational type has a higher utility, they also suffer surprise every time the

initiator attempts surprise. The biased type has a slightly lower expected utility, but they are much less likely to suffer surprise.

Project Organization

The analysis proceeds in a straightforward fashion. Having introduced the primary ideas of interest in this chapter, the next chapter takes a step back to examine prior contributions to foreign policy decision-making and intelligence. There is also a brief discussion of key psychological issues and brief diversion into behavioral economics. The behavioral economics literature helps set the stage for the first of two main theoretical chapters. The third chapter addresses belief updating, the impact of beliefs on decisions, and the relationship between rational and anchored beliefs. It also presents a mathematical approach to representing irrationally entrenched beliefs that is used in the formal model. Chapter 4 is the primary deductive modeling section of the project. The chapter begins by offering a basic model of surprise and then incrementally adds greater complexity, eventually incorporating irrationally entrenched beliefs into a formal model of surprise. The deductive results from a Bayesian rational model are compared to those of the model with a potentially irrational actor in order derive propositions about the impact of irrationally entrenched beliefs on the players' strategies, the likelihood of surprise, and expected utility.

Chapter 5 is the start of the empirical analysis. The chapter statistically examines one of the deductive propositions from chapter four and starts by outlining the primary variables. Since the statistical analysis relies on proxies for key variables like attempted surprise and entrenched beliefs, the model does not provide proof for the deductive

proposition, but the analysis suggests that there is some support for the theoretically deduced claim. Chapter 6 tries to overcome some of the shortcomings of the statistical analysis by examining two cases of foreign policy surprise in detail. The chapter begins by discussing the case study method and its importance for establishing causality in psychological theories of international politics. Then the chapter examines the 1962 Cuban Missile Crisis and the 1967 Arab-Israeli War. Each case study concludes with an explicit comparison of expected results from the model with rational actors against the model with a potentially irrational actor. The model with a potentially irrational actor seems to perform better than the model with perfectly rational actors, providing additional support for the deductions in chapter 4. The project concludes with a review of significant findings, discusses some limitations of the research, and offers some policy implications related to rational and irrational foreign policy decision-making.

Chapter 2
Literature Review

This work draws from and is intellectually indebted to an extensive and diverse collection of scholarship. This project lies at the intersection of literature on intelligence, foreign policy decision-making, psychology and behavioral economics. Anything approximating a comprehensive review of each topic could be a book length work. This chapter focuses on the pieces that were influential in framing this research. The chapter proceeds by introducing the canonical pieces of intelligence and surprise literature. Intelligence literature justifiably focuses on the acquisition of information and analysis. The function of intelligence is the distribution of information to decision-makers in a timely fashion. Since practitioners, in large part, have driven the study of intelligence, this is a logical place to stop. Decision-maker response to the information and its impact on strategic outcomes, however, are important issues in foreign policy and security.

To explore these issues, it is important to move beyond the formative body of intelligence and surprise to the even larger body of work looking at foreign policy decision-making. This is a multifaceted literature that incorporates diverse analytical methodologies and emphasizes different aspects of the decision-making process. For example, game theoretic analyses focus on strategic elements such as national interests, information issues and institutional arrangements, whereas decision-process scholarship tends to focus on small group dynamics, individual interests and issue framing. Since this

project focuses on strategic issues, and is most securely cast within the rational game theoretic approach, this literature is highlighted. The primary component of this section is an analysis of information updating and learning. Since the theory developed in subsequent chapters compares efficient or Bayesian updating with a biased mechanism, it is important to address the role of information processing in different contexts.

It is also important to understand some core psychological concepts and their prior application to international politics. Psychology and the study of international politics have always had a somewhat tenuous relationship. Much of this is driven by the empirical difficulties associated with applying psychological explanations to one-time events. These empirical concerns will be saved for Chapter 5. However, a number of scholars have maintained that psychology plays an important role in determining international outcomes. This work introduces some important challenges to the rationale choice paradigm, but as I argue throughout this project, these analyses need not be mutually exclusive.

The final section, a very brief and narrow review of behavioral economics, serves as a link between the relevant literature already reviewed and the theoretical components of the analysis. This project borrowed some general, but important ideas from the behavioral economic literature, namely the integration of psychological concepts into formal models of strategic interaction. This section will briefly highlight some core ideas associated with behavioral economics and review a few specific pieces of scholarship that were influential here.

Intelligence and Surprise

Intelligence literature can be broadly divided into three categories: practitioners' accounts, official or governmental reports, and scholarly analysis (Maoz 2006). There are a few individuals who span the divide between practitioner and scholar effectively (Heuer Jr. 1981; Lowenthal 2003; Herman 1996), however, the scholarly work is the most theoretically and empirically developed. Practitioner accounts often describe the way in which information is gathered and governmental reports often address instances where the system did not function as intended. Scholarly work on intelligence has focused more broadly on outcomes and is consumed by issues of intelligence failure and strategic surprise. These two issues are of course linked since many instances of intelligence failure involve some level of strategic surprise. Scholarly work in intelligence is not without problems; there exists a bias towards some aspects of intelligence while ignoring others; there is little or no variability in the dependent variable; and there continue to be problems with case selection and research design (Maoz 2006). Having said that, there is still a clear, empirically supported dialogue, worth discussing.

The four common explanations for intelligence failure fault signal-to-noise ratios in noisy environments (Wohlstetter 1962), the possibility of deception (Whaley 1969), organizational constraints and tradeoffs (Betts 1978), and analysts' psychological biases (Heuer Jr. 1981). In truth, each of these explanatory theories is probably sufficient to explain an instance of intelligence failure, but it also possible that multiple causes are operating at the same time. Each one of these presents a formative challenge in isolation, leading most analysts to the conclusion that intelligence failure and strategic surprise are inevitable (Whaley 1969; Betts 1982; Kam 1988).

Intelligence apparatus work in inherently noisy and uncertain environments. Wohlstetter (1962) applied signal transmission theory to the surprise attack on Pearl Harbor in 1941. She noted that states and intelligence apparatuses gather a lot of information. Much of the information is ambiguous or difficult to decipher since analysts get small pieces of information that need to be pieced together. Information that gets corrupted or is too ambiguous is often considered noise. Information that can be used to improve assessments or estimates is considered signals. It can be difficult to identify and separate the signal from the noise. It is also possible that some signals will be dismissed as noise, whereas some noise will be interpreted as signals (Wohlstetter 1962; Lowenthal 2003). The intelligence community often refers to this problem as wheat versus chafe (Herman 1996; Lowenthal 2003)

Deception is another common cause of intelligence failure. In a competitive system, some actors seek to gain an advantage through the use of surprise (Whaley 1969). In recent history, states actively attempt to mask behavior through simulation (showing the false) and dissimulation (masking the real) (Whaley 1982). Analysis suggests that almost all attempts at surprise between 1910 and 1968 were successful, even when there was available warning (Whaley 1969). Deception itself is often further complicated by noise constraints. Information on an opponent may be worthless if the opponent is aware that the signal is being used to forecast intention (Jervis 1972). Simply acquiring information about an opponent is insufficient; one must also assess the opponent's knowledge of these information-gathering efforts. These issues are magnified when one considers the costs associated with strategic response. Any attempt to make the warning system more sensitive will result in more false alarms (Axelrod 1979).

Stepping back from a unitary actor perspective, analysis of organizational constraints in intelligence led to the now conventional belief that intelligence failures are inevitable (Betts 1978). There are three irresolvable issues that plague intelligence: failure in perspective whereby actors fixate on beliefs and miss environmental changes, communication breakdowns, and the inability for organizational reform without tradeoffs. The first criticism stems from the nature of predication itself. Analysts and models that are most often accurate in predicting patterns are most likely to miss large changes and outlaying events (Handel 1989). The second issue, somewhat more self-explanatory, addresses the communication breakdowns that are endemic to large and secretive organizations. The third element, and the primary justification for the inevitability of intelligence failure, is the inability to fix the major problems without sacrificing performance in another area. Analytic barriers to accuracy also play role and include ambiguous evidence, ambivalence of judgment, and atrophy of reforms. Given the presence of ambiguous evidence, analysts can often justify a number of conclusions, and this often leads to analytic ambivalence. Organizational attempts to improve rigor and overcome these constraints tend to whither over time. In this causal explanation, organizational issues and individual limitations in perspective, or psychological biases, are detrimental and a root cause of intelligence failure.

There is an additional explanation of failure that deals with the performance of the individual analyst (Heuer Jr. 1981). This perspective treats individual limitations and psychological biases in a more delicate and balanced fashion. Individuals are influenced by the lens through which they view the world, and that lens is developed through personal and professional experiences. These experiences or perspectives that are

indispensable for interpreting ambiguous evidence may also contribute to intelligence failure. Preconceived notions, lenses, or biases are the backdrop necessary for the analysis of ambiguous evidence and partial information. Of course, these beliefs and biases can also lead an analyst astray, resulting in very inaccurate estimates.

There are also numerous analyses of intelligence failure that offer case specific reasons for intelligence failure such as U.S. surprise when North Korea attacked South Korea in 1950 (Stueck 2001), the Israeli attack that surprised Egypt in 1967 (Oren 2002), the Egyptian attack on Israel in 1973 (Bar-Joseph 2003), The Argentine attack on the Falklands that surprised the U.K (Hopple 1984), and the Iraqi attack on Kuwait that surprised the U.S. (Freedman and Karsh 1993). Analysis of the Argentine invasion of the Falkland Islands suggests that the root of the U.K. failure was in the intelligence-policy link (Hopple 1984). The British policy never signaled serious commitment, which emboldened the Argentines. British analysts never considered that their policy would have that undesired effect. Likewise, there are a multitude of explanations for failing to predict the downfall of the Iranian Shah in 1979. One explanation addresses the culture of the CIA, the lack of peer review, the willingness to rewrite history when it suits a picture, and the under-provision of analytical resources (Jervis 1979). Alternatively, it has been argued that failure stemmed from a lack of human intelligence brought about the Shah's concerns over CIA presence and request that they be removed (Daugherty 2001). Israel's 1973 surprise has also generated different perspectives. The dominant one addresses Israel's belief that the Egyptians would not attack without air superiority. A slightly different perspective maintains that these beliefs were only held by a couple of influential intelligence officers, and the existence of a singular intelligence product drowned lower-

level competing voices (Bar-Joseph 2003). What becomes clear from reading these conflicting explanations is that many of the competing hypotheses have some degree of validity. It is rare that there is any singular cause for intelligence failure.

Analysis of the US failure at 9/11 has tended towards relying on multiple explanations (Parker and Stern 2005); however, others have argued that 9/11 was not a strategic intelligence failure (Pillar 2004). The common denominator is that 9/11 was not a "bolt from the blue." In fact, most cases of surprise are not out of the blue, and almost all involve prior warning or indication of some sort (Whaley 1969; Betts 1982; Levite 1987). Days leading to the Egyptian surprise on Israel were full of military and diplomatic maneuverings, not to mention the prior tension and protracted rivalry. Prior to Germany's surprise attack on the Soviet Union in 1941, Hitler assembled German forces on the Soviet border. Despite rebuffed diplomatic attempts, Stalin refused to interpret the German mobilization as a signal of impending attack. Almost every historical case exhibits some form of warning, whether attack-specific intelligence or a history of tension and conflict. It is important to note that almost all of the warnings of impending surprise could also be interpreted as signs that there is no imminent surprise. For example, Hitler's decision not to engage in diplomatic conversations with the Soviets, and not make any demands, helped to strengthen Stalin's belief that the Germans were not going to attack. The empirical availability, and interpretation, of warning is somewhat problematic for intelligence scholars. It makes it difficult to move beyond the four well-developed explanations: noise, deception, organizational constraints, or psychology.

Often curiously missing from intelligence literature is a discussion of how decision-makers react to the intelligence they receive. After all, a good deal of

intelligence is vague and open to interpretation, even in its finished form. Prior to the 9/11 attacks, there were numerous warnings of increased Jihad traffic and speculation of a large attack. Nonetheless, this intelligence brought no significant policy shifts. When Stalin's generals brought intelligence of the impending attack from the front, Stalin ordered only minor defensive measures for fear of antagonizing Hitler into action.

The intelligence failure literature lacks clear consensus on a definition of failure. Work by, and oriented for, practitioners views intelligence failure as an inability to produce accurate and timely warning. This is the argument for viewing 9/11 as a tactical, but not a strategic intelligence failure (Pillar 2004). The emphasis is on the availability of information and the quality of the intelligence product. Scholars of surprise and warning have tended to take intelligence failure a step further, equating it with policy failure. By this standard, intelligence output is only successful when accurate warnings are shaped into policy. Practitioners have chosen to eschew this approach to failure for obvious reasons: it is their job to provide accurate and timely information, not to set policy. Policymaker's response is out of their hands and beyond their control. The intelligence officer's job is done when the intelligence is put into a manner that the decision-maker finds useful in crafting policy. Those who have gravitated to the broader policy-based definition have obvious reasons as well: cases of surprise attack and the policy failures that result from faulty information are inherently interesting. These events are also readily observable, an element that should not be overlooked in social science (and one that can be particularly troubling in studies of intelligence).

Despite the tendencies for scholars, consciously or unconsciously, to equate policy and intelligence failure, the link between the two has not been developed in

rigorous fashion. Decision-making and policy response are often viewed as deterministic or secondary. When the intelligence is poor, bad policies are expected. When the intelligence is good, it is often assumed that accurate policies are crafted in response. The approach looks at the quality of warning as a key indicator of policy outcome (Levite 1987). However, given the ambiguity associated with international competition, most intelligence analysis involves subjectivity that may involve individual biases. This suggests that the relationship between information in competitive environments, analytical perspectives and the possibility of psychological biases, and policy outcomes should be examined further.

Foreign Policy Decision-Making

The study of decision-making in international politics can be split into two broad categories. The first category addresses strategic decision-making where competitive pressures play an explicit role in policy selection. The second category focuses on the decision process, using internal characteristics to explain why certain policies are selected. These two categories do not capture all of the literature, but serve as a useful organizing tool.

Early strategic modeling treated states as unitary actors with clear national interests (Schelling 1960; Allison 1971, 1999). The models relied on the realist assumptions that all states were essentially identical and acted in their own self-interest (Waltz 1979). Initial realist research (Morgenthau 1948) assumed that all states tried to maximize their power, and later iterations considered that states may seek to maximize their own security (Waltz 1979; Glaser 1997). In formal models applied to international

politics, decision-makers often choose polices based upon each player's preferences over outcomes, the information available, the structure of the game (who gets to move first, etc.), the choices, and their beliefs. Formal models of decision-making rely upon economic notions of equilibrium. Assuming that the individuals are rational, formal models rely upon logical deduction. This approach to decision-making provides an internally consistent method for predicting behavior. Most models are intended to be simplified or abstract versions of general issues. Models never integrate all the nuance of real life, but try to use basic situations to identify expected strategies and outcomes. The primary advantages of formal modeling include the internal rigor, logical conclusions, and ability to identify counterintuitive insights.

Formal modeling with unitary actors has addressed issues as diverse as the balance-of-power (Niou and Ordeshook 1990; Powell 1991), bargaining and war outbreak (Kilgour and Zagare 1991; Fearon 1995), deterrence (Langlois 1991; Nalebuff 1991), alliance behavior (Alfeld and Bueno de Mesquita 1979; Morrow 1991), nuclear proliferation (Beuno de Mesquita and Riker 1982; Wagner 1991), arms races (Powell 1993; Kydd 1997), arms control (Wittman 1989; Kilgour and Brams 1992; Feaver and Niou 1996), and international cooperation (Axelrod 1984; Morrow 1994b; Downs et al. 1996; Fearon 1998). In an attempt to produce theories of the democratic peace, formal modelers dropped unitary actor assumptions. Different institutional arrangements have been employed to gain insights about the role of domestic politics in international decision-making. Models address domestic audience costs (Fearon 1994), opposition signaling (Schultz 2001), and public goods provision (Bueno de Mesquita et al. 1999).

The logic of formal models helps to sort through complicated interactions with diverse interests while still arriving at internally consistent explanations. In short, the approach achieves a degree of parsimony that is difficult to attain otherwise. However, the simplification of complicated interactions is one of the primary challenges to the use of formal models. An alternative approach to foreign policy decision-making relies upon bureaucratic politics. Unlike the logical processes that accompany fixed interests and structures, bureaucratic politics stresses negotiated decision-making where individual roles and conceptions of interest may change. Much of the detail stripped away from formal approaches to decision-making is central to explanation and prediction. The nature of the decision-making unit, the centralization of authority, personal relationships, parochial or organizational interests, personal history, and issue framing can influence outcomes (George 1980). Unlike formal approaches where alternatives and interests are exogenously dictated, alternatives and interests arise from debate among the main players (Allison 1971, 1999).

There are three general approaches to the study of bureaucratic politics. This paradigm has been used to explain the policy preferences that arise among multiple actors (Smith 1985), to understand and predict the final policy outcomes (Vandenbroucke 1984), and evaluate the quality of the outcome to understand whether the process of bureaucratic bargaining was beneficial or detrimental (George 1980). While bureaucratic bargaining has been applied to areas as diverse as arms control (Halperin 1972; Feaver 1993), foreign policy crises and military response (Janis 1972; Allison 1971, 1999; Legro 1995), most of literature focuses on decision machinery. Therefore, much of the research addresses issues like the role of epistemology (Haas 1992), small group dynamics (Fuller

and Aldag 1997), collective risk-taking (Vertzberger 1997), importance of parochial interests (Smith 1985), bargaining power (Hermann and Hermann 1989), and psychology and cognition (Ripley 1995).

Bureaucratic politics approaches have been assailed for their lack of parsimony, adequate specification and empirical rigor (Welch 1998). There are almost infinite variables that could influence policy outcomes, and identifying all of them is nearly impossible. Further complicating matters, identifying the primary variables takes in-depth, detailed analysis that restricts testing across cases. The large number of variables and small number of cases create an inferential problem often referred to as too few degrees of freedom. However, the bureaucratic politics still has intuitive appeal to those with an interest in the policymaking process and its outcomes. In reality, the policymaking process is very complicated and both individual and group elements play an important role (George 1980).

Despite the attention to detail in the decision-making process, the bureaucratic politics approach does not really address issues of information and learning (Stern and Verbeek 1998). Yaacov Vertzberger (1984) examines how the interaction of the decision-making group and the information providers impacts "(1) the search, selection and evaluation of information, (2) the degree of openness to dissonant information, and (3) processes of adaptation to the new incoming information." Groups, organizations and institutions develop coherent values or beliefs that influence members. When individuals belong to different groups, they run the risk of belief conflicts. Vertzberger argues that individuals normally try to alleviate the conflict by negotiating some compromise solution, or adopting the beliefs of one group (usually the small decision-making group

rather than the organization or referent group). It may be difficult to resolve these issues, particularly when faced with a trade-off between parochialism and embarrassment. The conflicts and compromises alter the group's ability to accept new or dissonant information and make adaptation difficult. This is especially true when the decision-makers choose risky alternatives or experience polarization. According this explanation, the individual's beliefs, and the groups they belong to, all influence the ability to process information and adapt.

Alexander George (1980) also examines the role and use of information in bureaucratic contexts, but puts greater emphasis on contextual and individual information-processing issues. The complexities and ambiguity of foreign policymaking have a significant effect upon the decision-makers. The uncertainty often leads to different forms of procrastination. When confronted with time pressure and uncertainty, leaders tend to rely on mechanisms such as bolstering, satisficing, incrementalism, consensus politics, historical analogies, and personal ideology when crafting decisions. Organizational behavior and bureaucratic politics may complicate the decision-making process. Often times, information gathering and processing units will tend towards collecting information benefiting the specific organization. Organizational evaluation of policy suffers from similar parochial interest. Bureaucrats will also try to leverage all available resources, information being central among them, to gain bargaining power. This often results in exaggeration or over simplification of information and analysis. Alternatively, individuals or organizations may withhold information for the benefit of another group, or to avoid the policy problem altogether. Standard operating procedures also affect the information collected and analyzed. These procedures are supposed to

simplify information management and policymaking, however, they can also create blind spots that impact learning or belief formation. Managing these constraints is a difficult process.

Charles Hermann (1990) explores different areas where learning might take place within the governmental system while trying to explain the source of foreign policy change within an existing administration or regime. Examples include Nixon's trip to China, and Anwar Sadat's peace with Israel after the 1973 war. Each of these events involves major foreign policy changes without leadership change. Hermann identified four potential changes agents: leader led, bureaucratic advocacy overcoming inertia, domestic restructuring among political elites, or external shocks. More than one may be operative at the same time, but each involve some sort of learning. In many cases, modest amounts of information are often ignored and only major events or negative feedback motivate policy changes associated with learning. Other work has addressed policy change and strategic assessment in war (Gartner 1997). Here, organizations tend to adopt reasonably simple quantitative indices in accordance with their objectives, what is termed the dominant indicator approach. The indicators provide feedback to decision makers who must decided whether to change the methods by which war is fought or the larger goals of the war. Sudden and dramatic changes in these indicators are more likely to effect policy.

Whereas the bureaucratic politics approach to information and learning presents a myriad of issues including contextual, individual, psychological and organizational forces, formal approaches to decision-making have less to say about information processing and learning. In part this stems from the traditional realist perspective that

states react to the power distribution in the system, and learning is essentially irrelevant (Reiter 1996). In formal treatments of decision-making, information updating is mechanical and often operates efficiently according to Bayes Rule. In many instances relevant actors are presumed to start with common prior, that is to say identical probabilistic beliefs. There are a few notable deviations from the common priors or Bayesian assumptions, but these are exceptions (Myerson 1991; Sakovics 2001; Smith and Stam 2004). Some actors may benefit from acquiring private information, and the rest need to form updated beliefs as the game progresses. This updating process is akin to learning, more specifically diagnostic learning (Levy 1994). The actors with incomplete information use whatever information available to create new probability estimates. Information source might include the behavior of their rivals, signals from third parties, outside advisors or private information. The information available to actors plays a large role in determining expected outcomes. To this effect, it would be wrong to assert that information is irrelevant in formal modeling. However, the information processing in these contexts assumes optimal mechanical learning and excludes many of the individual, psychological and group issues associated with learning.

Other works in international politics that have specifically addressed learning in a decision-making context adopt a bounded rationality approach that is closer to the bureaucratic politics approaches (Reiter 1996; Gartner 1997). Bounded rationality assumes that individuals have constraints on their processing abilities. Analysis of small power's alliance decisions in the 20th century shows that systemic or power explanations are not sufficient, and that state behavior exhibits historical learning (Reiter 1996). Alliance behavior is one of the most important decisions small states can make and it is

not surprising that decision-maker's beliefs play an important role. These beliefs are altered by previous national experience (Reiter 1996). According to organizational theory, learning occurs infrequently, is stimulated by major events, and is targeted to avoid past failures (Levitt and March 1988). Empirical analysis shows that states often learned lessons from their own experience and ignore experiences of other states. There is also no guarantee that states learned the correct lessons about alliance behavior, and learning has led to poor policymaking. One example is Romania in the 1920s and 1930s. Romania made some substantial territorial gains after World War I resulting from its treaty with the Allied Powers. After the war, however, Romania became a status quo power and chose against allying itself with either of major power alliances in the 1920s and 1930s, joining the Little Entente instead. By the late 1930s, Romania found themselves offering significant concessions to Nazi Germany after rejecting alliance offers (Reiter 1996). Reiter (1996) argued that learning model did appear to outperform the standard realist balance-of-threat theory, which had only marginal explanatory power. Systemic threats appeared to have a greater impact than direct challenges, but the statistical relationship was contrary to expectations. These findings support earlier work suggesting that states learn from their own past experiences (Leng 1983; Huth and Russett 1984).

It is worth noting that showing the importance of past history on policy is different from explaining how states actually learn and how the lessons are applied. Jack Levy (1994) attempted to address theoretical and empirical deficiencies in foreign policy learning. Drawing distinctions from other scholars who take an applied or normative perspective, Levy defines "experiential learning as a change in beliefs (or the degree of

confidence in one's beliefs) or the development of new beliefs, skills, or procedures as a result of observation and interpretation of evidence." Learning may not result in policy changes, or be correct in some normative sense and this can make any analysis of learning empirically difficult. Individuals do not necessarily process information in a straightforward or rational manner, and there is a large body of literature that examines how psychology influences foreign policy decision-making. Since individual learning is much more closely tied to cognition and psychology, we turn there next.

Psychology in International Relations

The discussion on learning was decidedly a-psychological. In part this was done in order to review the primary theoretical and empirical contributions to learning in foreign policy contexts. However, almost all of the work reviewed above addressed the psychological and cognitive components of learning. Learning involves the organization of stored information in a metaphoric web, what are commonly called schemata. Learning occurs when additional information is added to the schemata, new relationships or paths are drawn between information within schemata, or new relationships and connections are built between schemata. There is no single or dominant learning model within psychology. Of course, learning represents a small component of psychology and this section focuses on the broader applications in international politics.

The introduction identified Jervis' (1976) work as canonical in international security. In the first section, Jervis showed how different strategic contexts called for drastically different courses of action. The spiral model showed how hard-line policies exacerbated conflict, but the contrary appeasement policies failed in a deterrent world.

This creates an obvious dilemma since the appropriate policy is determined by the strategic context (or state of the world), but the policymaker may be unaware of the context. The best a policymaker can do is rely upon their perception of the state of the world. Since perception plays such an important role in the selection of strategy and the success of the outcome, it is important to understand the root of misperception. The canonical status of the work does not stem from its authoritative proof of any set of propositions, but rather from its theoretical contribution and incredibly broad application of different psychological concepts. Jervis addresses cognitive consistency (the tendency to seek consistency in beliefs, and between beliefs and actions), the evoked set (the information that is readily accessed during decision-making), learning, and attitude change (where attitudes are a mixture of beliefs and emotional associations). These concepts are then applied to common strategic misperceptions including tendencies towards analyzing opponents as highly centralized entities, and overestimating one's own influence upon target behavior and motives. He also addresses the role of motivated biases such as wishful thinking, and the relationship between cognitive dissonance and policy inertia.

Tetlock (1998a) split subsequent psychological research in international politics into two categories: 1) explanations of behavior focusing on cognitive heuristics, and 2) the impact of different kinds of systemic biases. The distinction seems slightly unnatural since heuristics and biases are not clearly separable. Heuristics are decision-making, problem solving, or learning shortcuts. When confronted with a decision, people rarely do all the possible calculations as predicted by bounded rationality. Instead, people tend to rely upon simple rules for making decisions. Heuristics often serve people well, but

may lead to systemic bias (Kahneman et al. 1982). This makes the distinction between the two a little blurry. For example, the representativeness heuristics causes individuals to assume that objects similar in appearance have similar traits. Normally, this heuristic works well enough for identification; however, it can also lead to a bias called the regressive fallacy. When suffering from a regressive fallacy, individuals believe that there is a causal relationship between two independent items. This heuristic has also been tied to the gambler's fallacy whereby individuals believe that their behavior may have an impact on independent events. Despite the relationship between heuristics and bias, the two are often separated in applications to international politics (see Reiter 1996 for an example of tying them together).

The separation is largely a result of the causal processes of interest. For example, Suedfeld and Tetlock (1977) focus on the importance of integrative complexity and the tendency towards heuristic behavior. Integrative complexity is measure of information processing. When there is low integrative complexity, people rely on heuristics, take rigid positions and restrict information use. Alternatively, high levels of processing and cognitive complexity are associated with complexity, flexibility and extended information search. Psychology experiments have shown that people tend toward heuristics when information quantity, time pressure, and threat get higher. Together, this suggests that integrative complexity and heuristics seem well suited to decision-making in international politics. According to this perspective, integrative complexity alters the expected decision-making behavior. Decisions made with higher integrative complexity should result in better policymaking prone to fewer psychological biases. The integrative

complexity is a predictor of applied analytic prowess and bias, which is related to outcomes.

The alternative approach focuses directly the presence of cognitive and judgmental biases. Rather than using integrative complexity, which is exceedingly difficult to observe, scholars rely on different biases in building causal theories. Experiments have shown that people experience numerous different kinds of biases (Nisbett and Ross 1980). Most biases involve causal relationships, information updating, and confidence. Applications of bias in foreign policy descriptively address the impact of the bias on the decision-making unit, and predictive applications usually associate biases with inferior outcomes.

This type of psychological analysis has been applied to numerous areas of interest in international politics. Deterrence failures are equated with overconfidence whereby the deterring party falsely believes that the opponent will not challenge (Lebow 1981). A similar process leads aggressive states to overestimate the probability that an opponent will capitulate. These types of policy failures may also be the result of motivated biases like wishful thinking (Stein 1991). Simplified decision-making and bias has also been used to explain the preference for offensive military doctrines (Snyder 1983). Other work has addressed how decision-makers apply history, focusing on the biases and preconceived notions that lead to inaccurate conclusions (Neustadt and May 1986; Vertzberger 1986). George (1980) relies on psychology in his analysis of individual decision-making, specifically tendencies to avoid ambiguous situations, procrastinate, utilize historical analogy when ill-suited to a situation, and rely upon ideology and beliefs. People do not always find that biases or stress interfere with decision-making.

Sometimes policymakers are open to new information and carefully consider all the options (Stein and Tanter 1980).

Recently, there have been numerous attempts to apply prospect theory to international politics. Prospect theory addresses decision-making in risky situations. The theory, which has a good deal of empirical support in laboratory settings, argues that people react differently to gains and losses. People tend to be risk averse in the zone of gains and risk seeking in the zone of losses. The theory also addressed people's tendency to overweight payoffs associated with low probability events and underweight the payoffs associated with high probability events (Kahneman and Tversky 1979). Prospect theory has been applied to broad phenomenon like bargaining and war (Butler 2007), risk-taking in American foreign policy (McDermott 1998), and great power intervention (Taliaferro 2004). It has also been applied to specific cases, often compared to expected utility models. Mark Haas (2001) looks at satisficing decisions in the Cuban Missile Crisis, and Rose McDermott and Jacek Kugler (2001) examine decision-making in the Gulf War.

One of, if not the, most cited biases in international politics is a tendency towards relying on prior beliefs or preconceived notions. Philip Tetlock (1998b, 1999) has addressed this issue in numerous ways. He argued that experts have a tendency to ignore dissonant information and maintain their prior beliefs by using a complicated set of defense mechanisms (Tetlock 1999). Psychological studies were conducted that confronted professional observers of politics with situations that were similar to the types of problems they might try to solve on a regular basis. The participants were provided different information or cures, and their belief changes were evaluated. People also have

a tendency to maintain belief in the quality of their own counterfactuals. Specifically, there is a tendency to view arguments that are wrong as almost right, but correct assessments are rarely perceived as almost wrong (Tetlock 1998b). The power and inertia of prior beliefs have been used to explain U.S. intelligence failures in Vietnam (Wirtz 1991), the tendency for militaries to use outdated doctrines successful in the prior war (Levy 1984), and miscalculation and misperception in conflict initiation (Jervis 1976; Levy 1994). Recently, there have also been efforts to explain enduring rivalries and protracted conflicts using psychological bias. States tend to have inert views of their role in the international system, even when the system undergoes changes that should facilitate revaluation of roles and identity (Thies 2001). It has also been argued that that the combination of national security conceptions and conflicts leads to a rapid lock-in creating rigid beliefs about the rivalry (Mor 2004). This behavior might be rational given the differential consequences of different types of errors, but the maintenance of these beliefs may or may not be rational given new information. Preconceived notions and prior beliefs have also been identified as one of the major impediments to ending protracted conflicts (Bar-Tal 2000).

Despite a growing literature on psychology in international relations, there is still debate about its merits and explanatory power. These concerns are generated from two perspectives: 1) the possibility and difficulty of addressing alternative explanations and 2) the difficulty associated with empirically evaluating psychological explanations of behavior. Both positions warrant brief discussion here, and will be addressed again later in Chapter 5. Dominant explanations of behavior in international politics center on strategic concerns, and in many cases simple analyses of national interests in strategic

contexts are sufficient to explain behavior. These explanations may abstract away much of the decision process, but their parsimony and predictive ability justify the simplification. Often, when simple strategic explanations are not sufficient, we tend to look for explanations based upon domestic politics or international norms. This often leaves psychological explanations as a marginal competitor to alternative theories. The flip side of this argument is that the psychological dynamics discussed here such as heuristics and biases are almost always operating in decision-making. The primary question is whether incorporating psychology provides additional descriptive and predictive power to current theories.

The second issue addresses the empirical difficulties associated with proving causal arguments involving psychology. Empirical research from psychology is usually based upon laboratory experiments with multiple trials and control groups. In international politics, this type of research design is impossible to achieve. The events of interest occur once. Certainly general events like surprise attacks recur, but they often involve different people, information sets and strategic contexts. Even more problematic is the inability to decisively observe heuristics or biases. Therefore, psychological analyses usually rely upon detailed case studies that try to understand the primary decision-makers beliefs and decision processing. When compared to alternative explanations with observable variables, psychology in international politics has a large empirical hill to climb.

Having said that, many of the existing theories of international politics rely upon psychology in one form or another (Goldgrier and Tetlock 1998). Many neorealist theories rely on miscalculation of forces or misperception of the strategic context. The

same is true for constructivist theories relying on identities and norms. The tacit incorporation of psychology and decision-making processes into the predominant theories of international politics suggest that there is a place for psychological explanations. Arguably, one of the major problems is that psychology is often pit against alternative theories as competition rather than compliment. Much could be learned by carefully incorporating psychology into the relevant strategic contexts. To try and bridge this gap, we provide a brief introduction to behavioral economics, the attempt to integrate psychology into standard economic decision-making theories.

A Brief Introduction to Behavioral Economics

Behavioral economics grew out of prospect theory, its wide acceptance, and the need to address empirical violations of the dominant theoretical models. By showing that people systemically violate certain tenets associated with rationality, specifically in regards to risk-seeking behavior and valuation under uncertainty, prospect theory provided new tools to address empirical puzzles. Some of these empirical puzzles were rooted in approaches to risk, ambiguity (Ellsberg 1962), and bargaining impasses. Others focused on violations of theory in the securities markets. For example, scholars have documented underreactions to public news announcements violating the efficient market hypothesis (Groth et al. 1979; Grinblatt et al. 1984), short-term positive momentum in returns (Jegadeesh and Titman 1993; Daniel et al. 1998), long-term negative reversals in asset returns (DeBondt and Thaler 1985; 1987), and stock price movements in opposite directions than earnings (DeBondt and Thaler 1987). Together, this mix of unexplained

phenomenon led a number of economists toward the integration of psychology into economic reasoning.

The field has progressed along three methodological avenues. Consistent with mainstream economics, behavioral economists continue to rely econometrics and financial market data. Many of the initial puzzles that motivated the movement were just such analyses, and it remains an important empirical component. The second approach common to behavioral economics is the use of laboratory experiments. Borrowing from psychology, scholars have used experiments to identify and address specific decision-making mechanisms. This tool is relatively new to economics, but it has played a significant role the development of prospect theory (Kahneman and Tversky 1979) and bargaining impasse (Babcock and Lowenstein 1992). Most recently, behavioral economists have started formalizing propositions about psychology and decision-making in formal models (Rabin and Schrag's 1998; Compte and Postlewaite 2004). The hope is that these models will yield greater theoretical understanding of processes observed empirically. The application of psychology to formal models of behavior is tricky. Just as one needs to make tradeoffs and simplifications in game structures, the same applies to psychological mechanisms.

Behavioral economic modeling is split between efforts addressing the market anomalies, and those tailored to decision-making. With the exception of the bargaining literature, many of the formal models are decision-theoretic. That is to say, most of the models address the impact of psychological dynamics on a single decision-maker, as opposed to multiple players conditioning their behavior on one another. Behavioral economics has also looked at the behavior of multiple actors in examinations of market

bubbles and panics (Smith, Suchanek and William 1988). Decision-theoretic market-focused models have relied on biases such as representiveness and conservatism (Barberis et al. 1998), or overconfidence (Daniel et al. 1998) to explain the anomalies in investor behavior and market pricing. Both models address how a biased actor will act with incomplete information in decision-theoretic contexts.

Two models proved particularly useful in framing and developing this research. Both works address the role of judgmental biases in non-market decision-making. The first is Mathew Rabin and Joel Schrag's (1998) model of confirmatory bias. Confirmatory bias is the tendency to interpret signals as confirming preexisting views. Rabin and Schrag allow a decision maker to update in biased fashion by incorporating confirmatory bias and show how the bias skews beliefs. They allow for two possible states of the world and assume there is a signal associated with each state. The decision-maker views one signal in each period and uses it to update their beliefs in a Bayesian fashion. However, there is some probability that the individual will misinterpret less likely signals and incorrectly use the information to bolster existing beliefs. An individual may develop strong beliefs about the state even when the signals are random, and these beliefs are strongly history-dependent. One of the deductive results was that a few signals early on could have strong lock-in effects and make the individual blind to the randomness in the environment. In the extreme, decision-makers may continue to maintain and reinforce inaccurate estimates even when there is infinite information available.

The second model, by Olivier Compte and Andrew Postlewaite (2004), addresses confidence biases. The authors focus on two ways that confidence biases manifest

themselves: 1) there is a tendency for individuals to believe they are better then they are at a given activity, and 2) when individuals are more confident, they tend to perform that activity better. Compte and Postlewaite examine the decision behavior of an individual whose performance is altered by confidence. The individual evaluates the probability of success, where the likelihood of success is determined by their previous successes divided by past trials. The model assumes that the individual's capabilities do not improve over time, specifically that the actual skill level is exogenous to the model and held constant throughout repeated decisions and outcomes. Confidence biases leads individuals in inflate this probability. The inflated or biased beliefs about the probability of success leads the individual to accept gambles that would otherwise reject if they had an accurate assessment of previous success. However, the increase in perceived probability of success also raises the likelihood that the individual is successful. They show that the decrease in utility stemming from the sub-optimal gambles is more than offset by the increased performance (due to second order effects). This leads them to conclude that some level of self-confidence bias is optimal and performance enhancing. We now turn to the main theoretical arguments of this project.

Chapter 3
Beliefs, Signals, and Biases

Intelligence failure is a term commonly used to denote breakdowns in the assessment or bureaucratic process that result in surprise. People experience surprise when they are confronted with unexpected outcomes. This implies that people have expectations or beliefs about what is going to happen. Studies of foreign policy and intelligence failure often argue that people experienced surprise because they did not change their beliefs about the future commensurate with the information available. Retrospective studies of intelligence failure often argue that information was discounted, or interpreted in certain ways, ultimately resulting in surprise.[2] To assess this argument, we have to take a closer look at rational and irrational beliefs and belief change. We argue that the distinction between rationality and irrationality is not as straightforward as we ordinarily think.

The chapter begins by defining beliefs in a very specific way and thinking about how those beliefs may change when confronted with new information. This section is intended to familiarize the reader with a rational-actor approach to beliefs and learning. Next, the chapter addresses strategic signaling, which examines how these beliefs and learning mechanisms operate in competitive contexts. This moves beyond beliefs and belief change to explore how other actors may behave to alter one's beliefs. In doing so,

[2] For a full discussion see the intelligence section in Chapter 2.

we hope to build familiarity with some concepts about competitive interaction that are important in later chapters. The chapter concludes by addressing biases and specifically deeply entrenched beliefs. We provide a definition for entrenched beliefs and show how they may differ from rational learning mechanisms.

The central task of this chapter is to distinguish between rationally strong prior beliefs and deeply entrenched beliefs. The previous chapter noted a problem in differentiating the two, and the importance of doing so in order to develop a better understanding of misjudgment in competitive situations. Actors with rationally strong prior beliefs are expected to discount new information, appropriately given certain prior beliefs. If beliefs are rationally derived, there is nothing irrational or biased about this information discounting. Deeply entrenched beliefs refer to an error in judgmental processing where information is discounted irrationally. Drawing this distinction, and specifying a way to differentiate the two, helps set the stage for a more in-depth look at the relationship between entrenched beliefs and surprise.

Rational Beliefs and Learning

Before leaving the house most mornings, people check the weather report to find out whether it is going to rain. If the probability of rain is high enough, then people may choose to carry an umbrella. In these instances people are forming or updating their beliefs about the chance of rain for that day. There is some probability it will rain, and some probability it will not. This probability is what economists call a belief, and specifically a belief about the state of the world (Morrow 1994a). Beliefs arise from uncertainty. We assign some likelihood to different possible states when we are uncertain

about what has or will happen. Technically, a belief is a conjecture about a probability distribution (Watson 2002), summarizing an actor's judgment about was has happened or will happen. This economics-based definition is different from the way we define beliefs in everyday use, which usually conveys the mental acceptance of truth (Heritage Dictionary 2007). In the context here, beliefs may be long held or fleeting. Actors can hold beliefs about the weather, about history, about others' behavior, about preferences, and so on. People can hold beliefs about any number of things, but the common element is that a belief represents a probability distribution over a set of possible states or outcomes that are uncertain.

One's belief about the likelihood of rain plays a large role in the subsequent decision to take an umbrella. If the probability of rain is zero there is no reason to take the umbrella. Contrastingly, we are likely to take the umbrella if rain is a certainty. Between those two extremes lies an entire spectrum of possibilities between zero and 100%. The probability of rain on any given day may be 17%, 43% or 86%. The decision to bring the umbrella is based upon our beliefs about the likelihood of rain, as well as our preferences for not getting wet.

Someone who wants to avoid getting wet at almost any cost may choose to take the umbrella if they believe the likelihood of rain is anything higher than zero. This person has a low belief threshold necessary to justify taking the umbrella (perhaps a 10% chance). That is to say, at any chance of rain greater than this critical belief threshold, they carry an umbrella. At any belief below this critical belief, the umbrella is left at home (Morrow 1994a). A different person, who cares less about getting wet, will have a higher threshold or critical belief. Perhaps this other individual only carries an umbrella

when the chance of rain is above 80%. This illustrates how preferences and beliefs can influence behavior. When the probability is less than your threshold or critical belief you leave the umbrella, otherwise, you take it. We can define a critical belief as the point, or belief, at which you are indifferent between your choices, in this example, bringing or not bringing the umbrella.

We have illustrated how our beliefs and preferences come together to influence our behavior, but we have not addressed how beliefs change. Changing our beliefs about the likelihood of different states is a form of learning (Levy 1994). Learning, however, is a much broader concept. Economists usually limit their statements about belief change to the term belief updating. So, we will begin by providing a basic framework for belief updating.

We discussed the existence of a critical belief in the rain example above, but have not addressed how we come to understanding that our belief is above or below the critical value. This is the essence of belief updating. One could apply different cognitive rules to the example above (Reiter 1996; Tetlock 1999). For instance, you might believe the latest morning weather report verbatim. If the report says there is an 80% chance of rain, that is what you believe. This simple rule, however, ignores any prior beliefs you might have held. Perhaps a weather report the night before indicated there was a 10% chance of rain. Now you have to decide whether you trust the new report, the old report, or some mix of the two. Perhaps you are inclined to trust the new report more since it has the latest data. That would appear to be a rational way to update your beliefs. However, if it were beautifully sunny the entire previous week, and the morning weather predicted an 80% chance of rain everyday, you might be hesitant to trust the report. Now, it might seem

entirely rational to ignore this new report altogether based upon past experience. In actuality, we are likely to use our knowledge of recent weather and the accuracy of recent weather reports to form some new belief about the likelihood of rain.

Social scientists often presume that people update their beliefs in a logically consistent way according to Bayes Rule (Gerber and Green 1999). Bayes Rule, which will be dealt with in more detail in the next section, is based upon conditional probabilities. Conditional probabilities are ways of limiting a pool of information in order to focus on the relevant data (Wonnacott and Wonnacott 1990). Economists often focus on one event occurring given the observation of a prior event (Gibbons 1992). For example, assume there is a bag with four gumballs, two are red and two are blue. The initial probability of drawing a red is 2/4 or 50%. However, if we observe the removal of a blue ball, the conditional probability of subsequently selecting red is 2/3. We can say that the probability of selecting red given that a blue was initially chosen is 2/3. So, our belief about the color of the gumball we select is conditional on the observation of the gumball selected earlier.

Bayesian updating looks at the likelihood of an event contingent upon prior information (Myerson 1991; Gerber and Green 1999). Lets return to the rain example one final time. Lets assume we just moved to London, England and have no idea what the weather is like. We wake up the first morning to find that it is raining. The second day we wake to find a bright sunny day. We note that it rained one of two days (1/2). On the third day, we wake up again to find it raining, and note that it rained two of three days (2/3). For the rest of the week we wake to rain and resolve to living in a place where it rained six of seven days (6/7). Note that each day provided information about the likelihood of

rain, but each observation was not weighted equally. As you get more information about prior events, newer information has less of an effect on your beliefs. For example, the likelihood of rain was 100% after the first observation, 50% after the second, and 67% after the third. Subsequently, in days four, five and six, the rational beliefs were 75%, 80%, and 83%, respectively. By day seven, we are resolved to believe that the Bayesian likelihood of rain on any given day is 86%. This concept is an important part of Bayesian rational updating and worthy of closer examination.

The Bayesian concept above can be represented mathematically as,

$$Likelihood\ of\ \ Rain = \frac{Days\ of\ \ Rain}{Total\ \ Days}.$$

On day three, we began the day believing that the likelihood of rain was 1/2 or 50%. This was the rational prior belief about the likelihood of rain. Upon observing the rain on day three, we updated our belief that likelihood of rain was 2/3 or 67%. This called the posterior belief, and is a rational belief based upon prior and new information. The value of this new information often diminishes with the strength of the prior. For example, on day three, the prior belief of rain was 2/3. If the sun came out on day four, we should update our belief to reflect that it rained 2/4 days or 50%. If we had observed rain six out seven days (6/7) or 86%, observing one day of sun on day eight should only cause us to change our belief to 6/8 or 75%. In fact, on day seven, we would need to observe sun for five straight days in order to rationally believe that likelihood of rain was 50%.

We can set up a more extreme case. Once again, lets assume that our prior belief of rain on day three is 67%, which is to say we observed rain two of three days. If we observe one day of sun, our belief rationally changes to 2/4 or 50%. Lets fast forward and assume that we are on day 300 of our stay in London. In those 300 days, we have

observed rain 200 days, so our rational prior belief about the likelihood of rain is still 67%. If the sun comes out on day 301, we know that it has rained 200 of our 301 day stay. One day of sun changes our beliefs from a 67% chance of rain, actually 66.66…%, to 66.4%. Observing one day of sun on day four and day 301 do not have the same impact on posterior beliefs. Given the strength of our posterior beliefs on day 301, the new information is rationally discounted relative to prior information.

We can generalize the Bayesian concept we have been looking at further, and represent it as,

$$\Pr(A) = \frac{a}{a+b}.$$

This notation simply states that the probability of event A is calculated by looking at the number of instances of *a* divided by the total number of instances *a*+*b* (Gerber and Green 1999). In the example above, the event is the likelihood of rain, where *a* is the number of rainy days observed and *b* is the number of sunny days observed. This highlights that our prior beliefs, and the strength of our prior beliefs, both influence our posterior beliefs and receptivity to new information. When prior beliefs are embryonic, in the early stages of formation, some small piece of new information may have a large effect. Contrastingly, beliefs rationally well supported by history are less vulnerable to small pieces of new information that should have little effect on our posterior beliefs.

This is an important concept that is often overlooked in many studies that address poor learning, irrational information updating, or entrenched beliefs in foreign policy and surprise. Information may be ignored because of faulty updating, but one must also consider whether information was ignored because of the rationally strong prior. The

absence of any attempt to differentiate discounting due to rational priors from biased updating makes any subsequent attempts to attribute causality tenuous at best.

In the middle of World War II, the Germans launched a surprise attack against the Soviets codenamed Operation Blau (Blau 1955). This attack is less well known than Operation Barbarossa, which was the formal initiation of hostilities between the Germans and the Soviets. Despite minimal historical attention, Blau is among the most fascinating cases of intelligence failure during the war. Operation Blau called for the large-scale offensive on the Southern German-Soviet battlefront. The Soviets were convinced the German offensive in spring 1942 would target the Northern front toward Moscow (Whaley 1969; Glantz 1990). A southern attack would help the Germans secure Ukrainian grain and oil from the Caucasus, but Adolph Hitler had taken large risks and losses in an attempt to take Moscow in 1941. Joseph Stalin and many military leaders were convinced that Moscow would be the primary target of the spring offensive.

Operation Blau, the southern surprise attack, commenced on June 28, 1942. The Soviets were not prepared to counter the attack since the majority of troops remained stationed up north to protect Moscow (Glantz and House 1995). Germany's initial attack was a success, however, the Soviets did eventually fight the advancing Germans to a standstill. The fascinating part of the Blau saga actually took place before the German attack. By mid-May 1942, there were signs that Germany might attempt a southern attack in response to a Soviet advance earlier in the month. One such sign was the emergence of German air superiority in the region (Whaley 1969; Glantz 1990). The Germans were careful in their planning for the operations. Many of the combat units that would be used for the southern attack were assembled at the center of the front and were even

distributed maps of Moscow (Blau 1955). Propaganda Minister Joseph Goebbels planted a newspaper story that Moscow was the target on May 15, only to denounce the newspaper and the author on May 20 (Whaley 1969). Goebbels followed up the newspaper article by sending an emissary to Portugal with instructions to leak information about an imminent attack towards Moscow (Whaley 1969). In a final attempt at deception, on May 25, the Germans issued orders to replace the battle-hardened German units with weaker Allied units in the south (Blau 1955). The irony associated with all of the German deception activities was that the Soviets were already convinced that the attack would target Moscow (Glantz 1990), making it seem as though all of this planning was moot.

Then, on June 19, just nine days before the start of Operation Blau, the Soviets downed a small German plane. The plane was carrying a German officer heavily involved with planning the upcoming attack. The officer had violated orders to destroy all the documentation associated with the attack and was carrying the order of battle and tactical plans for the initial stages of Operation Blau (Higgins 1966; Whaley 1969). Soviet field intelligence recovered the plan from the downed aircraft, but did not know what to make of this information. The plans were forwarded to Moscow. The leadership in Moscow examined the plans and almost immediately discounted the authenticity (Whaley 1969; Glantz 1990). The most recent evidence on Soviet intelligence during the war suggests that Stalin and the military commanders were convinced that the battle plan was one more German attempt at deception (Glantz 1990). When the operation commenced on June 28, the early stages of the offensive proceeded almost exactly as the captured plan indicated. It is interesting to note that Stalin, and most of the military

leadership, continued to believe that the plan was intended as misdirection into mid-July (Higgins 1966; Whaley 1969).

As analysts and social scientists, we are compelled to ask whether the Soviet decision to discount the battle plan was reasonable or ill advised. Too often we tend to rely on hindsight and assume that decision-makers must have been suffering from some sort of bias, however, the discussion in this chapter points out the need to distinguish between strong prior beliefs and judgmental bias. Taking a step back, one has to consider the Soviet position prior to finding the battle plan.

The Soviets were caught by surprise by the initial Barbarossa invasion in 1941 (Clark 1965). Military leaders encouraged Stalin to improve defensive positions and call up additional troops, but Stalin was hesitant to provide Hitler with a motive for attacking (Gorodetsky 1999). The Germans, however, were not passive participants in the initial surprise (Whaley 1969; Gorodetsky 1999). Hitler regularly stressed that the German military buildup along the Soviet border were preparations for an invasion of Britain. At the same time, Hitler ensured that Germany maintained diplomatic relations with the Soviets, and dangled the possibility of deepening their ties. Against this backdrop, Stalin interpreted the British Prime Minister Winston Churchill's accurate warnings of the impending attack on the Soviets as misinformation (Jervis 1976; Gorodetsky 1999). By 1942, the Soviets had repeatedly been the targets of German misinformation and deception campaigns.

There are also strategic and tactical issues to consider. During the initial stages of the war between the Germans and Soviets in 1941, Hitler expended significant resources in attempt to capture Moscow (Glantz 1990; Glantz and House 1995). The Germans paid

dearly for this effort. Soviet leaders were convinced that these tremendous costs signaled that Germany was intent on moving toward Moscow in 1942 (Glantz 1990). Months after Operation Blau, as late as November 1942, Stalin still believed that the southern attack was a costly feint and that the real attack would target Moscow Higgins 1966, Whaley 1969).

Looking at the evidence in retrospect, with the benefit of hindsight, it is tempting to argue that the Soviets suffered from entrenched beliefs when evaluating the likelihood of a Southern attack. The Soviets began with the prior belief that Moscow was the target of the next German offensive. When confronted with an accurate battle plan of the impending Southern attack, with sufficient time to take defensive action, the Soviets ignored the information. On the surface, this seems as though it might reflect judgmental bias. However, we must consider Soviet prior beliefs and the possibility that the battle plan was part of a deception campaign. It would be wrong to simply assume that this is evidence of entrenched beliefs without considering the alternative possibility that the Soviets had sufficiently strong priors. When one considers the costs the Germans paid in their attempt to capture Moscow, and the plethora of sophisticated disinformation campaigns targeting the Soviets, it becomes more difficult to argue that the Soviet surprise was the result of entrenched beliefs. It might be just as plausible that the decision to discount the captured battle plan was perfectly rational based upon the strength of prior beliefs. The Soviets might also have calculated that losses along the Moscow line would have been more costly than losses down South. The decision to concentrate efforts along the Moscow line could, therefore, also be viewed as a hedge even if an attack was expected in the South.

In short, Operation Blau represents a situation where the victim of surprise had unusually valuable intelligence about their opponent's strategy, but the information was heavily discounted. It is easy to say that the Soviets made a mistake in retrospect, but it is much more difficult to differentiate rational from impaired belief updating. This is not say that irrational or biased updating is irrelevant or unusual. As we will discuss below, biased updating might be more common than perfectly rational updating. However, it is important to specify a clear distinction between rational and irrational learning if we are use bias as a causal mechanism.

In the next section, it stops raining. We move away from the everyday examples above, and focus on international competition. Instead of worrying about our umbrella, we are worried about the likelihood that our opponent is aggressive and willing to attack. Beliefs, the relationships between beliefs and actions, critical beliefs, prior and posterior beliefs and the belief updating process continue to play an important role, but now we can use these tools to assess our opponent and decide how to respond. The Operation Blau example introduced the nature of competitive pressures, and below we will specify these types of interactions in greater detail.

Strategy, Signals and Beliefs

Prior to Operation Blau, the Soviets were trying to determine whether the Germans would attack north towards Moscow or further south. The primary source of uncertainty was where the Germans would attack, not if. Prior to the Barbarossa attack in 1941, however, the primary source of Soviet uncertainty was whether the Germans would attack or not (Gorodetsky 1999). In short, the Soviets were uncertain whether the

Germans were aggressive or docile. This distinction, aggressive or docile, will be a primary source of uncertainty in subsequent chapters. This section begins by introducing some basic concepts associated with signaling models, then goes on to show how beliefs and Bayesian updating play a role in competitive contexts with uncertainty. We also show how other actors' strategies can provide information, or act as signals, that may be used to reduce uncertainty.

Lets assume that the Soviets were trying to determine whether the Germans were aggressive or docile prior to the 1941 Barbarossa attack. Following convention, we can say that Germany had two possible "types," aggressive and docile. The aggressive type has different intentions, or preferences for attacking the Soviets, then the docile type. Typing, or the use of different types, is a way of introducing uncertainty into a competitive situation (Morrow 1994a). It is common to assume that an actor knows its own type, which is to say that Germany knows whether it is aggressive or docile. The Soviets, however, are uncertain about Germany's type. The actual selection of type is often called a "move by nature," where Germany's type is selected from a probability distribution known to both actors. For example, Germany may be the aggressive type 40% of the time and the docile type 60% of the time. Economic models and game theory often assume that these probability distributions are common knowledge (Myerson 1991). Every player knows this probability, and they both know they know, etc. Therefore, the Soviets know the likelihood that Germany will be aggressive, however, they are uncertain whether the Germany they actually face is aggressive or docile. This uncertainty presents a problem for Soviet decision makers.

We can outline an interaction that includes the sequence of moves that Germany and the Soviets can choose from. After learning its type, Germany can decide whether it wants to make a public demand or prepare for a surprise attack. Some historians argue that Stalin expected Hitler to make demands prior to any military hostilities (Gorodetsky 1999). The Soviets needed to select an appropriate course of action given their beliefs. The Soviets could respond by taking defensive action or waiting. We can assume that Soviets would like to make defensive preparations when Germany is aggressive and wait when Germany is docile. In this sequence of actions, the Soviets can use German demands as a signal, and Germany takes this into account when selecting a strategy. Signaling models like this are useful for studying information transmission, since a player can observe opponents' moves and update their beliefs accordingly. In this interaction, German demands are the signal of interest, however, an opponent's moves may involve any number of behaviors such as actions or statements. Signaling models have been used extensively in international security and particularly in the study of deterrence (Schultz 1998).

We now have a basic framework that describes the dilemma faced by the Soviets prior to the Barbarossa attack: nature determines the German aggressive intentions, Germany decides whether to make a public demand or prepare for surprise attack, and the Soviets then decide whether to take preparatory action or not. Just as we wanted to have the most accurate beliefs about the chance of rain in the previous section, the Soviets want to develop with the best possible assessment that Germany is aggressive. Given Soviet preferences, the Soviets will have a critical belief, similar again to umbrella example. When the Soviets believe that Germany's likelihood of being an aggressive

type is above this critical belief, the Soviets always take action. Alternatively, if the Soviets believe that the likelihood Germany is aggressive is below the critical belief, the Soviets always wait.

This begs the question, how do the Soviets update their beliefs about Germany's type? Many economic and game theoretic models use Bayes' Rule as a learning or belief updating mechanism (Myerson 1991). Bayesian updating is a logically consistent way of revising beliefs based upon the information available. Recall that Soviets know the likelihood that Germany is aggressive, since this is common knowledge. We can label this probability Pr(A), which simply stands for the probability that Germany is an aggressive type. The Soviets want to know the probability that Germany is aggressive given the absence of a public demand. This is a conditional probability that can represented as Pr(A|~D), where A again represents the likelihood that Germany is aggressive and ~D means that Germany chose not to make a public demand. Note that ~ is the logical sign used to indicate the absence or "not," and | is the sign used to denote given that or conditional upon. In order to reach a logical posterior belief given the German signal, Pr(A|~D), the Soviets must have an estimate of the likelihood that Germany will attempt a surprise attack Pr(~D). It is worth noting that the probability of a German surprise attack is a combination of likelihood that the aggressive type attempts surprise Pr(~D|A), the likelihood that docile type attempts surprise Pr(~D|~A), and the prior likelihood that Germany is aggressive in the first place. This can be represented as follows,

$$\Pr(\textit{No Demand}) = \Pr(\textit{Aggressive}) \times \Pr(\textit{No Demand} \mid \textit{Aggressive}) + \Pr(\textit{Not Aggressive}) \times \Pr(\textit{No Demand} \mid \textit{Not Aggressive}).$$

The Soviet Bayesian posterior belief can then be represented as,

$$\Pr(Agressive \mid No\ Demand) = \frac{\Pr(Agressive) \times \Pr(No\ Demand \mid Agressive)}{\Pr(No\ Demand)},$$

which can also be represented with abbreviations as,

$$\Pr(A \mid \sim D) = \frac{\Pr(A) \times \Pr(\sim D \mid A)}{\Pr(\sim D)} = \frac{\Pr(A) \times \Pr(\sim D \mid A)}{\Pr(A) \times \Pr(\sim D \mid A) + \Pr(\sim A) \times \Pr(\sim D \mid \sim A)}.$$

This calculation provides an objectively accurate estimate that the Soviets can use for future decisions. However, we have not said anything about Germany's conditional strategies Pr(~D|A) and Pr(~D|~A). The Germans know that Bayesian Soviets will engage in such information updating, so the Germans can adopt a signaling strategy aimed at manipulating Soviet beliefs in certain ways (Morrow 1994a). Germany's decision to make demands, in this context, can be logically deduced given the actors' preferences. We will look at some basic aspects of strategy selection before addressing entrenched beliefs.

A quick numerical example can help to illustrate the logic of Bayes' Rule. Lets assume that Germany has aggressive intentions towards the Soviets 40% of the time and docile intentions 60%. Lets presume that Stalin figures that Germany makes no demand 50% when they are aggressive. After all, attacking the Soviet Union is still a risky proposition even if the Germans are highly resolved. Stalin further estimates that the Germans make no demand 70% of the time when docile. Here, Stalin presumes that a docile Germany wishes to instigate public standoff most of the time. Using the same abbreviations we did above, we can say that Pr(Aggressive)=40%, Pr(No Demand|Aaggressive)=50% and Pr(No Demand|Docile)=70%. The Soviet posterior beliefs about German aggression given no demand is,

$$\Pr(A \mid\sim D) = \frac{40\% \times 50\%}{40\% \times 50\% + 60\% \times 70\%} = \frac{20\%}{62\%} \approx 32\% .$$

In the scenario above, the Soviet Bayesian rational posterior that Germany is aggressive for no demand is approximately 32%. The Soviets posterior belief that Germany is aggressive is now below the prior belief. In the absence of a public demand, therefore, the Soviets believe the likelihood of facing an aggressive Germany is lower then they initially believed. Under such conditions it might be perfectly rational to avoid taking defensive measures.

The hypothetical example above displayed the mechanics of Bayes' Rule, but it did not address the origins of the signaling strategy. For example, why would Germany opt not to make demands when aggressive 50% of the time versus 70% of the time? This is what we address next.

Germany must decide whether it wants to make a public demand or prepare for surprise, knowing that the Soviets observe the signal and update their beliefs. When the aggressive and docile German types always select different moves, this is called complete separation. Separating signals alleviate all uncertainty since the two different types always send signals that clearly differentiate themselves (Gibbons 1992). For example, the docile type may always make demands and the aggressive type may always prepare for surprise. Given such a separating strategy, the Soviets know that the Germans are docile whenever they observe public demands, and that the Germans are aggressive in the absence of any such demands.

Alternatively, complete pooling occurs when both types always select the same move all of the time (Gibbons 1992). Complete pooling offers no incremental information value since both types always send identical signals. For example, both the

aggressive and docile types may never make demands. In this case, the Soviets observe the absence of demands, but have no idea which type they are facing. The Soviets have no choice but to rely on their initial beliefs about the sender's distribution of types.

Between complete pooling and complete separation are semi-separating strategies (Gibbons 1992). These strategies are probabilistic. An example of such a strategy is a case where the aggressive type makes demands 50% of the time, and the docile type makes never makes demands. After observing the absence of a demand, the Soviets can use the prior belief, and Germany's optimal strategy, to refine the belief they are facing an aggressive type. Mixing strategies makes it difficult for the Soviets to determine whether an attack is truly forthcoming, but Germany's action carries some signaling information that is useful for the Soviets.

Semi-separating or mixed strategies are based upon making other players indifferent between choices. For example, an aggressive Germany should not make demands just often enough that the Soviet's expected payoffs for taking action and waiting are equal. Mixed strategies are deduced from the initial distribution over type and the actor's preferences. If the initial common knowledge probability that Germany is aggressive changes, Germany's mixed strategy also changes in order to ensure that the Soviets remain indifferent.[3] This is important in differentiating strong priors from entrenched beliefs, and highlights a central feature of signaling models that will be important throughout this paper.

[3] The sender always prefers that the target acquiesce, but in the absence of a strategy capable of accomplishing this end, game theory focuses on strategies that are unilaterally enforcing. A unilaterally enforcing strategy is one that, given an opponent's strategy, most reliably provides the best possible outcome on average. In equilibrium neither player has an incentive to alter such a strategy. A strategy that makes the target indifferent means the target does no better or worse by standing firm or acquiescing, and is therefore left free to select a strategy to make the sender indifferent to initiating the challenge.

The Soviets have to estimate Germany's type considering all possible types that the Germans could be (Morrow 1994a). This is the essence of the problem of judgment under uncertainty, and central to the rational choice explanations of intelligence. To this effect, Germany's decision to make no demand is not representative of a single observation of no demand. The signal represents an entire probability distribution of optimal German strategies given types. Therefore, the Soviets are not reacting to the single observation of no demand, but reacting to the likelihood that the aggressive type is among those that choose no demand. When the Soviets are indifferent, they should play a mixed strategy where defensive action is itself probabilistic. At any point or strategy other than indifference, the Soviets strictly prefer either to act or wait all of the time. This way of thinking about signals seems a little abnormal, but it is common to many signaling games.

The strength of the prior once again plays a role in belief updating according to Bayes' Rule. Ordinarily, Germany's conditional strategy will change if prior beliefs change. For this example, we will hold Germany's strategy constant and examine the impact of changing prior beliefs on posterior beliefs. Lets assume that Germany makes no demand 50% of the time they are aggressive and 100% of the time when docile. If the prior belief that Germany is aggressive is 50%, then the posterior belief that Germany is aggressive using Bayes' Rule is 33%. Therefore, the Soviet belief that Germany was aggressive declined from 50% to 33% upon observing Germany's signal. This is a change of 17%. Alternatively, if the prior of the aggressive type was 90%, then Bayes' Rule indicates that the posterior belief should reflect that Germany is aggressive 82% of the time. Under these conditions, the change in Soviet belief between prior and posterior is

8%. This illustrates the impact of strong beliefs in Bayes' Rule. The Soviet beliefs were subject to greater changes at 50% than 90%, which seems perfectly logical. If we are 90% sure that an event will occur, new information should have less of an impact on our beliefs. If we are only 50% sure that an event will occur, which indicates high uncertainty, new information may be very useful.

The example above also illustrates how the Germans may have caught the Soviets off guard. If the Soviets' initial prior belief was that Germany was aggressive 50% of the time, the absence of a demand may have reinforced the belief that Germany was non-threatening. German strategy may have caused a significant change in Soviet beliefs. If the prior likelihood of German aggression were 90%, that same strategy would not have been as effective in altering Soviet threat perceptions. The Soviets would be less likely to believe that the Germans were docile, irrespective of Germany's strategy. It would have been more difficult for Germany to catch the Soviets by surprise.

This leads to a final point in this section. Just as we showed above, beliefs help to determine behavior and strategies. The Soviets select their strategy based upon their posterior beliefs. The Germans select their signaling strategy in order to influence the Soviet's posterior beliefs. This means that both actors adopt strategies that are tied to their prior beliefs about aggression. We would expect changes in the actors' strategies should prior beliefs change.

Now that we have introduced Bayes Rule, shown how prior and posterior beliefs impact strategy and considered the impact of strong priors, we can examine an alternative updating mechanism. In the next section we discuss judgmental biases and deeply entrenched beliefs.

Biases, Anchoring and Entrenched Beliefs

We can think of Bayes' Rule as being a rational approach to information updating. Bayes' Rule is a logically deductive approach to learning that is objectively accurate. When actors have sufficient information to apply Bayes' Rule, the result is a statistically rational forecast of the probability of an event occurring. People, however, are not necessarily rational. Psychologists have gathered extensive data on the ways that people deviate from "rational" learning, assessment or estimation. In this section, we explore some of these deviations and specifically focus on deeply entrenched beliefs. We begin by looking at different types of psychological biases, then focus more closely on judgmental biases. We subsequently focus the discussion on anchoring bias and deeply entrenched beliefs. The section concludes by offering a mathematical representation of entrenched belief updating, shows how it logically varies from Bayesian updating and discusses some possible implication for behavior.

Psychologists and economists have noted that people systemically deviate from rational behavior (Ellsberg 1962; Kahneman and Tversky 1974; Nisbett and Ross 1980). Errors in learning and assessment are often referred to as psychological or cognitive biases. Cognitive biases are systemic and involuntary distortions in the mind that lead to perceptual errors or judgments that deviate from reality. Psychologists have further distinguished between different types of cognitive biases such as judgmental biases, decision biases, social biases and memory biases (Baron 2000). Each form of bias applies to different situations and involves different processing mechanisms. Since we are most concerned with judgmental biases here, we will focus on this category.

Judgmental biases refer to inaccurate estimates of causal relationships (Nisbett and Ross 1980). For example, the gambler's fallacy refers to the tendency to assume individual random events are influenced by previous random events (Estes 1964). For example, a player may roll a six sided dice three times and get a one each time. They then believe the likelihood of rolling one is very low on a subsequent roll, however, the probability of getting one is still 1/6. The actual probability is unchanged. Another judgmental bias is hindsight bias (Fischhoff and Beyth 1975). This is the tendency to see past events as predictable in hindsight, a problem that is almost unavoidable in analysis of surprise and intelligence failures.

Judgmental biases may be motivated or unmotivated biases. A motivated bias refers to an estimative error driven by self-interest or notions of self worth. For example, overconfidence bias leads people to overestimate their own capabilities (Svenson 1981). In these instances, people are motivated to induce a perceptual error. Unmotivated biases refer to estimative errors that arise without any concern for self-interest. Ambiguity effects are not motivated by self-interest. Ambiguity effects are the tendency to avoid selecting or discount options with missing information (Ellsberg 1962). In these instances, lack of information makes it more difficult to estimate probabilities, and such choices appear less attractive even if they are not.

One of the most common forms of judgmental bias is called anchoring bias. The literature review already introduced the basic ideas behind anchoring, but we briefly provide some additional clarity here. Anchoring refers to the tendency to rely too heavily on a single piece or subset of information (Kahneman and Tversky 1974). The subset of information used is the anchor. Once an anchor is set on a specific issue, people

overweight certain information and adjust estimates toward the anchored value. Anchoring bias can arise in both motivated and unmotivated contexts. Motivated anchoring bias occurs when an individual ignores information because they are certain of being correct (Tetlock 1998). In this context, anchoring is tied to a sense of self worth or estimative skill. Other times, anchoring can be completely unmotivated. Research has shown that people may anchor on sets of previously cued random numbers unconsciously (Nisbett and Ross 1980).

Anchoring on initial or prior beliefs is commonly referred to as having deeply entrenched beliefs. The previous section showed how all assessment or belief updating relies upon prior beliefs, however, entrenched beliefs are analytically distinct. Unlike strong rational prior beliefs, entrenched beliefs represent an irrational over-weighting of prior beliefs. Early or initial information, which serves as the basic foundation for inference, exerts excessive influence over subsequent belief revision (Nisbett and Ross 1980). When individuals suffer from entrenched beliefs, new information is discounted or ignored in order to preserve the prior or anchored belief. Entrenched beliefs result in information discounting beyond what would be considered Bayesian rational.

Since psychological biases, and specifically entrenched beliefs, are commonly invoked in studies of surprise and intelligence failure, it is important to specify the concept clearly. That is to say, we must find a way to distinguish between rational and irrational information discounting if entrenched beliefs are to be used as an explanatory variable. This is why it was important to introduce Bayesian updating in detail above. If Bayesian updating is a rational way to integrate new information, we can formalize some alternative that clearly distinguishes between the two.

A convenient way to think of entrenched beliefs is to think of a weighted average between rational prior and posterior beliefs.[4] By using a weighted average of prior and posterior beliefs, we can anchor the beliefs to the prior. The strength of the entrenched beliefs can be determined by the magnitude of actual anchoring bias. For example, we can say that the prior belief is weighted by the anchoring bias and the Bayesian posterior belief is weighted by 1-anchoring bias, where the anchoring is between zero and one. When the anchoring bias is set at the minimum level of zero, the entrenched and the Bayesian rational posterior beliefs are equal. When the anchoring bias is set at its maximum level of one, the entrenched posterior beliefs are equal to the prior beliefs. Technically, we can think of entrenched beliefs as,

$$Entrenched\ \ Posterior = Anchor \times prior + (1 - Anchor) \times Bayesian\ Posterior .$$

At any anchor level greater than zero, the actor using this updating mechanism is exhibiting entrenched beliefs since they are overweighting prior beliefs relative to the Bayesian posterior. This simple specification clearly distinguishes between Bayesian and entrenched beliefs without deviating too far from any of the updating concepts discussed earlier. Its form is also general enough to allow for both irrational under-suspicion and over-suspicion, which we address again below.

We can also represent entrenched beliefs in terms of the German-Soviet example above. Recall that the Soviets wanted to get the best possible estimate that the Germans were aggressive prior to the 1941 Barbarossa attack. The Soviets use Germany's behavior as a signal to update their beliefs about the likelihood of aggression. Germany can choose to make a demand or not. The Germans did not make any demands, and the Soviets

[4] Prior mathematical representations of bias have similarly looked at confirmation bias (Rabin and Schrag 1998) and overconfidence bias (Compte and Postlewaite 2004).

updated their beliefs conditional upon the German strategy. If the Soviets had entrenched beliefs, their beliefs would deviate from the Bayesian example above. Entrenched beliefs could be represented as,

$$Entrenched\ \Pr(Aggressive \mid No\ Demand) = Anchor \times \Pr(Aggressive) + (1 - Anchor)\frac{\Pr(Aggressive) \times \Pr(No\ Demand \mid Aggressive)}{\Pr(No\ Demand)}.$$

We can also use the abbreviated notation where A represents an aggressive Germany, ~D represents Germany's decision to make no demand and *b* represents the anchored value. The entrenched belief about German aggression would be,

$$\Pr(A \mid\sim D) = b \times \Pr(A) + (1 - b) \times \frac{\Pr(A) \times \Pr(A \mid\sim D)}{\Pr(\sim D)}.$$

It is worth noting that the mathematical definition of entrenchment allows entrenched beliefs in either direction. For example, it does not dictate that the Soviets anchor on a high likelihood that the Germans are aggressive. The definition of entrenched beliefs does not indicate whether entrenchment favors aggressive beliefs that result in over-suspicion or docile beliefs that dictate under-suspicion. The definition and associated math simply allow an actor with entrenched beliefs to overweight their prior beliefs, irrespective of what those beliefs actually are. Therefore, the mathematical approach described here is a general definition of entrenchment.

These equations provide a clear distinction between rational discounting of information and irrational entrenchment. We can apply this definition of entrenched beliefs to the Barbarossa example to examine the effects of anchoring bias or entrenched beliefs. It is easiest to return to a numerical example above. Lets assume that the prior likelihood that Germany is aggressive is 60%. We will further assume that Germany's

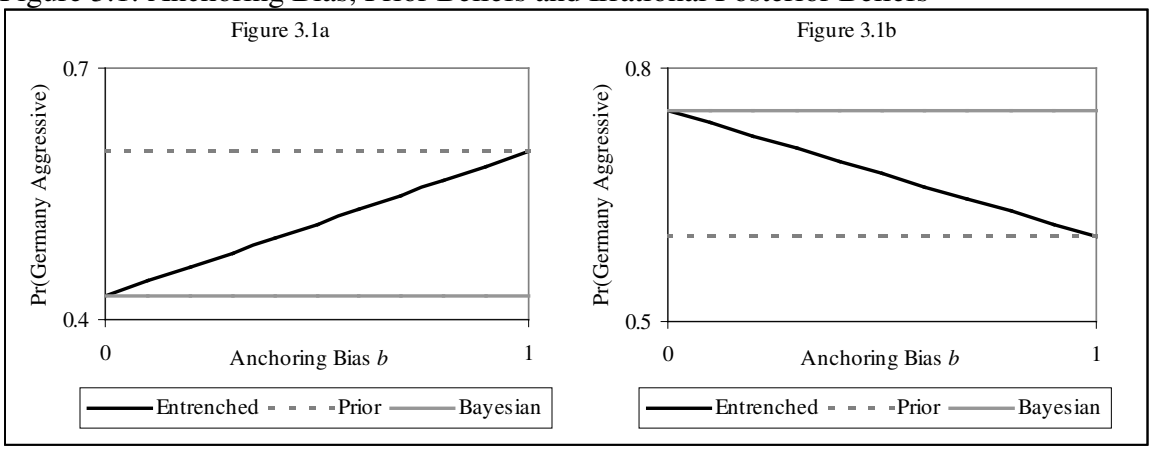
best strategy given the prior belief is no demand 50% of the time they are aggressive and 100% of the time when docile. Under these conditions, the posterior belief that Germany is aggressive using Bayes' Rule is 43%. If the Soviets fell prey to anchoring bias, then their posterior belief would be something between 43% and 60%. If the Soviet anchoring bias were 0.5, then they would weigh the prior belief and Bayesian posterior belief equally. The Soviets would believe that the likelihood of facing an aggressive Germany given no demand was 52% (0.5x60%+0.5x43%). In this case, the Soviet entrenched belief about German aggression is greater than the Bayesian rational belief.

The relationship between prior belief, Bayesian posterior beliefs, entrenched beliefs and anchoring bias are illustrated in Figure 3.1. The entrenched belief ranges between the prior and Bayesian posterior beliefs, and moves from the posterior to the prior belief as anchoring bias gets more severe. In each graph, the dashed line represents the prior belief, the solid grey line represents the Bayesian posterior belief and the solid black line represents the entrenched belief. Figure 3.1a, based upon the example just above, shows what happens when the Bayesian posterior belief is below the prior belief. As anchoring bias increases, the entrenched belief increases from the Bayesian posterior of 43% to the prior belief of 60%. Entrenched beliefs result in irrational over-suspicion.

The effect of entrenched beliefs reverses when the Bayesian posterior belief is greater than the prior belief. This is represented in Figure 3.1b. Assume the Germans do not make a demand 100% of the time when aggressive and 50% of the time when docile. If the prior belief that Germany is aggressive is 60%, then the Bayesian posterior belief is 75%. Under these conditions, increasing anchoring bias results in a decreasing entrenched belief of German aggression. The entrenched belief ranges from 75% with no

anchoring bias to 60% with maximum anchoring, therefore, Figure 3.1b illustrates irrational under-suspicion.

Figure 3.1: Anchoring Bias, Prior Beliefs and Irrational Posterior Beliefs



Note: In Figure 3.1a Pr(A)=60%, Pr(~D|A)=50%, Pr(~D|~A)=100% and the Bayesian posterior Pr(A|~D)=43%. In Figure 3.2b Pr(A)=60%, Pr(~D|A)=100%, Pr(~D|~A)=50% and the Bayesian posterior Pr(A|~D)=75%.

Figure 3.1 clearly shows how anchoring and entrenchment can impact posterior beliefs. When the prior belief is greater than the posterior, entrenched Soviets will overestimate the probability that Germany is aggressive, which we call irrational over-suspicion. Contrastingly, entrenched Soviets will underestimate the likelihood that Germany is aggressive when the prior belief is below the posterior belief, referred to irrational under-suspicion. This again highlights the fact that entrenched beliefs can result in overestimation or underestimation of a threat.

Before concluding this section, it is important consider the competitive or strategic implications now that we have defined and explored entrenched beliefs. To do this, we return to critical beliefs and optimal strategies. The critical belief was a threshold value that helped determine behavior. In the Soviet-German example, the critical belief helps determine whether the Soviets take defensive action or wait. If the Soviets believed the likelihood of German aggression was greater than the critical belief, they took action.

Otherwise, the Soviets preferred to wait. When the Soviets' posterior belief is equal to the critical value, this means that Soviets are indifferent between waiting and acting. Technically, they get the same expected payoff from either choice. The Soviets are free to wait and act with some probability, or mix strategies, at this critical belief. The Soviets can choose a mixed strategy that then makes the Germans indifferent between making and not making a demand. This type of indifference behavior is a typical result or equilibrium.

Germany picks a signaling strategy, mixing between making and not making demands, to make the Soviets indifferent if possible. The strategy that makes the Soviets indifferent occurs when the Soviet Bayesian posterior belief equals their critical belief. This means that the posterior lines in Figure 3.1 equal the Soviets critical belief, assuming that Germany has selected the optimal strategy in each case. We can further deduce that entrenchment results in Soviet posterior beliefs that deviate from the critical belief. In other words, if the Germans adopt the optimal strategy that ensures Soviet Bayesian posterior beliefs are equal to the critical beliefs, then entrenched Soviets will not be indifferent between acting and waiting.

Entrenched beliefs can potentially have a significant impact on Soviet behavior. Figure 3.1a shows that entrenched Soviets are irrationally over-suspicious. This means that Soviets think the Germans are more likely to be aggressive then they actually are, and this belief motivates the Soviets to take action all of the time. Figure 3.1b illustrates the case where entrenched Soviets are irrationally under-suspicious. The Soviets, therefore, underestimate the likelihood of facing an aggressive Germany and choose to wait all of the time. In these instances, Soviet judgmental biases can lead to erroneous

assessment and irrational behavior, provided that the Germans play their optimal strategies.

It is interesting to note that anchoring bias in these competitive contexts can actually lead to overconfidence bias.[5] When the Soviets have Bayesian rational beliefs and the Germans adopt an optimal strategy, then the Soviets know that they will be indifferent between waiting and acting. This allows the Soviets to mix their strategy. Entrenched Soviets, however, are likely to believe that one strategy is preferred to the other. For example, Soviets with entrenched beliefs in Figure 3.1a will believe that waiting is always preferred to acting. The Soviets overestimate the expected payoffs to waiting all of the time because their assessment about German types is inaccurate. The Soviets believe that their average payoff will be higher than it actually is. Recalling the discussion above, overconfidence bias is the tendency to believe that one will perform better than they actually will. Therefore, Soviet overconfidence bias, a biased assessment that they are making the right decision, can be driven by anchoring or deeply entrenched beliefs.

All of this suggests that deeply entrenched beliefs can have a significant, and potentially detrimental, impact on behavior and outcomes. This is a common theme in foreign policy and intelligence literature. However, it is not sufficient to assume that biases, in this case entrenched beliefs, always result in bad outcomes. One central idea running through this chapter was the difficulty associated with differentiating or distinguishing between strong priors and entrenched beliefs. We ordinarily assume that entrenched beliefs are detrimental to decisions and outcomes, but rarely address the strategic or competitive nature of decision-making. For example, what happens when the

[5] See page 12 for a discussion of overconfidence bias.

Germans are aware that the Soviets may have entrenched beliefs? Does this change Germany's behavior? If so, it is unsuitable to simply assume that entrenched beliefs are detrimental to the Soviets. After showing how we can analytically or mathematically distinguish between strong prior beliefs and entrenched beliefs, we need to take a further step to explore the relationship between bias and competitive outcomes. This is the topic of the next chapter.

Conclusions

This chapter provided a definition of beliefs, carefully walked through Bayesian logic and Bayesian updating, introduced the way that Bayesian beliefs play a role in competitive situations and provided a definition of entrenched beliefs that was technically distinct from strong prior beliefs. Each of these concepts plays a crucial role in the next chapter where we present a full game theoretic model of surprise and consider the logical impact of entrenched beliefs.

We conclude by summing up some of the central themes of this chapter. It is important to recognize that information discounting is not necessarily indicative of irrationality or deeply entrenched beliefs. Discounting information based upon strong Bayesian priors is a perfectly reasonable way to update or learn. In fact, failing to discount sufficiently can also result in inaccurate assessment. This problem may be particularly acute in intelligence where deception and counterintelligence efforts produce misinformation. Accepting information too easily runs the risk of falling prey to misdirection. If we want to draw conclusions about the impact of judgmental biases it is

important to begin with clear distinctions between rational and irrational learning. We also examined how a strong prior impacts the belief updating process.

The other central idea introduced in the chapter was the problem of decision-making under uncertainty. We explained how a single observation, such as Germany's demand, is indicative of more than a single observation. Upon observing the demand, the Soviets need to consider the possibility that an aggressive type is among the types that make a demand. German behavior, therefore, is indicative of an entire probability distribution of optimal decisions. This strategic decision is based upon prior beliefs. If prior beliefs change, then strategy is expected to change as well. Since entrenched beliefs are distinct from strong priors beliefs, it is important to consider the impact of entrenched beliefs carefully. In the next chapter, we examine the strategic implications of entrenched beliefs.

Chapter 4
Psychological Biases in a Model of Strategic Surprise

The purpose of this project is to examine the impact of rational and biased beliefs on decision-making, specifically as it applies to foreign policy. Psychological biases and predispositions are often cited as causal variables in cases of foreign-policy surprises such as the Cuban Missile Crisis or North Korea's invasion of South Korea. In this chapter, we introduce a model of surprise attack and systemically evaluate the impact of biased beliefs on outcomes. Scholars often argue that biases, specifically deeply entrenched beliefs, increase the likelihood of suffering surprise, thereby reducing the overall welfare of those with biased or rigid belief systems. Interestingly, Philip Tetlock (1999) has shown that decision-makers, in many cases successful ones, exhibit cognitive consistency that allows them to maintain previously held beliefs despite contradictory evidence. This begs an important question: how is it that these people become successful given their rigid or entrenched beliefs despite the widely held belief that such behavior is potentially dangerous and damaging? To answer this question, we first need to understand how psychological biases or predispositions impact decisions and outcomes.

The previous chapter introduced the one form psychological bias called entrenched beliefs, contrasted them with rational beliefs and showed how biases might impact behavior. This chapter pushes this investigation further by examining the impact of bias in competitive situations. We introduce a series of two-player models, where a

sender decides how to challenge a target, and incorporate the possibility of surprise. The model offers a sender the opportunity to challenge a target, and if they do challenge the target, they can choose to trigger a crisis diplomatically or can attempt a surprise attack. The target must decide whether to take preemptive defensive measures or wait. That decision is largely motivated by the target's beliefs about the sender. When the target believes that a sender holds aggressive intentions, they have incentive to take defensive action, otherwise they prefer to wait.

To explore the impact of biases, specifically irrationally entrenched beliefs, this analysis compares the behavior of perfectly rational actors against those who may exhibit deeply entrenched beliefs, and in so doing identifies the way that entrenched beliefs are expected to influence competitive behavior. Competitive behavior in this context refers to the decisions or strategies of both actors, the sender and target. In this model, the target is the only player with incomplete information that observes signals, and is therefore the only player that learns or updates their previously held beliefs. We begin by looking at a target that updates its beliefs rationally, which creates a benchmark for rational behavior. This benchmark or baseline provides a logically coherent justification for the actors' behavior. Then the analysis proceeds by incorporating the possibility that the target may exhibit biased updating, specifically deeply entrenched beliefs. The general specification of entrenched beliefs allows both under-suspicion and over-suspicion, where direction of bias is determined by the initial conditions. We examine the way that actors respond to the target's deeply entrenched beliefs under different initial conditions and then use the rational benchmark to evaluate the impact of deeply entrenched beliefs on expected outcomes.

The deductive analysis in this chapter shows why and when psychological biases or predispositions should impact expected outcomes. Providing theoretical support for Tetlock's (1999) empirical findings, this chapter shows that decision-makers can benefit from irrationally entrenched or rigid beliefs under certain conditions. Starting with the assumption that leaders often exhibit entrenched beliefs, the analysis here helps to explain why these biases can be advantageous rather than simply assuming they are detrimental. Game theoretic analysis shows that the possibility of bias resulting in irrational over-suspicion may change opponent's behavior, specifically focusing on the way that target bias may change the sender's strategy. Target entrenched beliefs that the sender is friendly, promoting under-suspicion, have no impact on either sender or target behavior or expected utility. Alternatively, senders should change their strategy when an entrenched target is over-suspicious owing to beliefs that the sender is likely to be aggressive. The sender, in this situation, is best off adopting a less provocative strategy. In the model developed in this chapter, a sender is equally likely or less likely to attempt surprise against a biased target. The target, therefore, can benefit from an overly rigid belief system as long as the sender knows the target may be biased.

Together this leads to two important conclusions. Conditions exist where entrenched or overly rigid beliefs can be beneficial to those that hold these beliefs. This chapter will identify those conditions and explain why this is true in greater detail. The second important conclusion provides a different perspective on the recognition and exploitation of others' biases. Literature on surprise, deception and strategic management ordinarily argues that decision-makers either ignore opponent's biases or can take advantage of them (Heuer Jr. 1981). By this logic, one may be able to improve their

outcomes by finding their opponents' blind spots. The analysis here offers a contradictory conclusion: 1) there is a logical reason and natural tendency to take others' biases into account and 2) opponent's biases should make one more conservative for fear of provoking an unwanted response. This suggests that entrenched beliefs or psychological biases should be recognized as risks to manage rather than opportunities to exploit.

This chapter proceeds by introducing a basic model of surprise, and then adds additional complications or refinements. The next section specifies the basic model of surprise that will be used throughout the chapter. After discussing solutions to the basic model, we add incomplete information where the target knows that the sender's intentions are drawn from a probability distribution. The chapter first offers equilibrium solutions for a perfectly rational actor, then examines the way that beliefs and behavior change when a target has entrenched beliefs. Finally, we compare the rational and biased models to get a better understanding of the effects of bias and see who benefits.

The Basic Surprise Attack Model

This section introduces a basic framework for thinking about surprise attack, and establishes a construct that can be used to evaluate the role of beliefs later. Surprise often refers to an emotional state of disbelief resulting from observation of an unexpected event. The concept of surprise attack is more tangible and refers to outcomes rather than psychological states. A surprise attack is a military operation that is intended to catch an opponent unprepared (Betts 1982). When the opponent's military force is caught unprepared at the time of attack, the attempted surprise is successful. A model of surprise must incorporate the possibility of both successful and failed surprises.

Surprise, however, is only possible when there is some degree of uncertainty. If an opposing force knew they were going to be attacked, they would take defensive precautions to thwart the attempt. A surprise attack can only be successful when the opposing force believes there is some chance that the status quo will be preserved and there is no impending attack. The surprise, in essence, is a result of uncertainty over the opponent's plans to attack or not. When an attack is expected, the attempted surprise is unlikely to be successful. Alternatively, the surprise is more likely to catch an opposing force unprepared when that force believes the likelihood of attack is sufficiently small. In summary: surprise is possible because of uncertainty, actors have incentives to take defensive action when they believe an attack is coming, prefer to wait when an attack seems unlikely, and it is this later state that facilitates successful surprise attacks.

The model captures the uncertainty surrounding surprise attacks, and the pressure to take defensive action. The basic surprise model consists of two players, the sender and target, with three total decision stages. As the model is explained, abbreviations of key terms are offered in parentheses. This shorthand notation is useful later in the mathematical sections. The sender begins the game by deciding whether to challenge (C) or preserve the status quo (~C). If and only if the sender chooses to challenge the target, the sender then must decide whether to challenge the target publicly (M) or attempt surprise (~M). A public challenge can take many forms such as diplomatic exchange, a warning, or a major military mobilization prior to hostilities. An attempted surprise, which we call a private challenge, implies that the sender is preparing to attack the target without sending any signals of the sender's plan. In the final decision stage, the target must decide whether to make act (A) or wait (~A).

The three decision stages combine create six possible outcomes to the game. When the sender attempts to preserve the status quo, there is no need to select the type of challenge, and the game ends when the target decides whether to act or wait. If the target waits in response, the status quo is maintained (Q). Should the target choose to take action against a sender attempting to preserve the status quo, the target suffers a false alarm (F). The other four outcomes pertain to senders that choose to challenge. When the sender challenges publicly and the target waits, this is considered acquiescing (E) to the demand. A public challenge that is met by target action ends in a standoff (D). If the sender chooses a private challenge and the target chooses to wait, the sender successfully surprises (S) the target. If the target takes action against the sender's private challenge, the attempted surprise is thwarted (R).

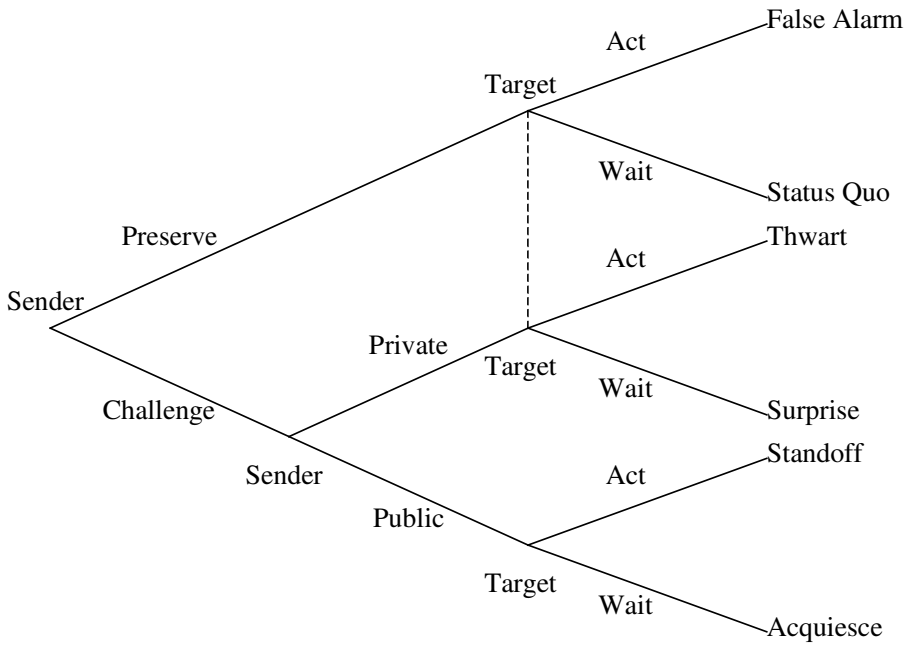
In order to capture the essence of surprise, the model must include some type of uncertainty. Uncertainty in formal models can arise from either imperfect or incomplete information (Morrow 1994a). Imperfect information refers to situations where an actor is uncertain about the previous moves of other players. Alternatively, actors have incomplete information when they are uncertain about the game itself or the characteristics of other players, and this is often used to add uncertainty about opponent's preferences and intentions. Incomplete information plays a large role later in the chapter, but uncertainty is incorporated into the basic model by using imperfect information.

The model assumes that the target is unable to differentiate between private challenges and attempts to preserve the status quo. The target observes public challenges, and knows that the sender has made a demand. In the absence of a public challenge, we assume that the target is unsure whether the sender has decided not to challenge or is

preparing for surprise attack. In other words, the target is uncertain about the sender's previous moves when they do not observe a public challenge. Since the target cannot clearly differentiate between attempts at preserving the status quo and surprise, which we call an ambiguous signal, the target has uncertainty and surprise becomes possible. Technically, the basic surprise model is a game of complete but imperfect information. This does not mean that the ambiguous signal carries no information value. When the target observes the ambiguous signal, they can rule out the possibility of a public challenge and use that information to select an appropriate strategy.

The full game tree for the basic surprise model is found in Figure 4.1, and the dashed line connecting the target's nodes represents uncertainty about the sender's past behavior. The diagram summarizes the moves, outcomes and information sets discussed in the previous paragraphs.

Figure 4.1: Basic Surprise Model



The only aspect of the model that has not been specified yet is the actors' preferences over outcomes. This model uses preference rankings similar to basic deterrence models where actors prefer to secure an opponent's acquiescence without violence. In formal models with discrete outcomes, there may be a number of possible preference orders, and this model is no exception. We solve the model using a few specific preference rules, and later discuss how changes in preferences impact the expected results.

We introduce the sender's preferences first and then discuss the target's preferences. The utilities for each of these outcomes is denoted by $U_S(.)$ and $U_T(.)$, for the sender and target, respectively. All things being equal, the sender prefers that the target wait when they choose to challenge. This means that the sender prefers acquiescing to standoff and a successful surprise to one that is thwarted. We also assume that the sender is indifferent between maintaining the status quo and triggering a false alarm.[6] One possible preference ordering for the sender, and one that is used in this section, can be summarized as follows: $U_S(E)>U_S(S)>U_S(Q)=U_S(F)>U_S(D)>U_S(R)$. This notation says that sender's preference over outcomes is: 1) target acquiescence, 2) successful surprise, 3) status quo and false alarm, 4) public standoff, and 5) a thwarted surprise. This preference order may be complicated by the sender's appetite for violence or public standoffs, and the impact of alternative preference ordering is revisited below.

[6] This assumes that the target's action is defensive and not akin to preemptive or preventative attack. There are two situations that could falsify this assumption. First, a sender might actually derive benefit from the target's false alarm, since the target expends resources on unnecessary defensive measures. Alternatively, a defensive reaction that results in a military strike would be detrimental to a sender seeking to preserve the status quo. However, if we assume that the target's actions are limited to defensive measures, then it is reasonable to assume that a sender will be indifferent since the final result is their desired preservation the status quo. In a repeated game, this assumption should be changed to reflect the fact that false alarms can both increase mutual hostility or desensitize actors to signals in a future round such as Egypt did to Israel in 1973.

The target, consistent with most deterrence models, prefers to maintain the status quo above all other outcomes. The target's least preferred outcome is acquiescence, which is an assumption necessary to derive useful conclusions. If the target preferred giving in to public challenges rather than standing firm, the sender could always challenge publicly and succeed without resorting to violence. This negates the roles of uncertainty and learning, which are two primary areas of interest in this analysis. Presumably, the target also prefers thwarting a surprise to suffering a successful attack. Target preferences in this section, listed from most to least preferred are as follows: 1) status quo, 2) thwarted surprise, 3) false alarm, 4) public standoff, 5) successful surprise, and 6) acquiesce to sender. The summarized notation is $U_T(Q)>U_T(R)>U_T(F)>U_T(D)>U_T(S)>U_T(E)$.

Now we have all the information necessary to derive solutions for this simple game of surprise. The next section introduces the mathematical solutions to the game above. The implications of these solutions are then discussed in the section after.

Solutions. Since the Basic Surprise Model described above is a game of imperfect information, Bayesian equilibrium is an appropriate solution concept. In a Bayesian equilibrium, each player selects a strategy for all possible contingencies given their beliefs, and no player can improve their expected payoffs by unilaterally changing their strategy. Game theoretic equilibria often involve strategies intended to make players indifferent between their available choices. When players are indifferent between possible moves, they cannot improve their expected payoffs by simply changing strategy. The mechanics of the equilibrium will be clarified as we derive the solutions.

We begin by looking at the target's equilibrium strategy since that is easier to derive. When the sender chooses to challenge publicly, the target has perfect information about the history of the game. The target knows that the sender has made some public demand. If the target prefers a public standoff to acquiescence, as indicated in the preference ordering above, the target's dominant strategy is to take action. The target will never wait and acquiesce.

The target's strategy selection, however, does become more complicated when the sender chooses to challenge privately or not challenge at all. In these cases, the target has imperfect information and does not know whether or not they face an attempted surprise. The target's best strategy is to make the sender indifferent between a private challenge and attempts to preserve the status quo. The target makes the sender indifferent when,

$$nU_S(F)+(1-n)U_S(Q)=nU_S(R)+(1-n)U_S(S),$$

where n is the probability that target takes action and 1-n is the probability the target waits. The sender is indifferent when the target mixes between acting and waiting such that,

$$n^*=\frac{U_S(Q)-U_S(S)}{U_S(Q)+U_S(R)-U_S(F)-U_S(S)}.$$

The target will always act when the sender challenges publicly, will act with probability n^* when they have imperfect information, and will wait with probability 1-n^*.

The sender has to decide whether to challenge and then decide whether to challenge publicly or privately. The sender can never make the target indifferent when they challenge publicly since there is perfect information, but the sender can make the target indifferent when they challenge privately or preserve the status quo. The target is indifferent by playing a strategy such that,

$$kU_T(F)+(1-k)U_T(R)=kU_T(Q)+(1-k)U_T(S),$$

where k is the target's belief that the sender preserved the status quo and 1-k is the probability that the sender attempts surprise. The target is actually is actually indifferent when,

$$k^*=\frac{U_T(R)-U_T(S)}{U_T(R)+U_T(Q)-U_T(F)-U_T(S)}.$$

Therefore, the target is indifferent between waiting and acting when the sender adopts a strategy that makes the target believe there is no challenge with probability k^* and private challenges with probability 1-k^*.

The sender's equilibrium strategy requires challenging and selecting challenge type in order to ensure that they preserve the status quo with probability k^*. The probability k^* is the equivalent of saying the probability that sender preserves the status quo given an ambiguous signal. The ambiguous signal, the sender's decision to preserve the status quo or challenge privately, is labeled state B. The conditional probability that the sender preserves the status quo given the ambiguous signal is,

$$\Pr(\sim C \mid B)=\frac{\Pr(\sim C)}{\Pr(\sim C)+\Pr(C)\Pr(\sim M)},$$

where C still represents the decision to challenge and ~M is the decision to challenge privately. We summarize the sender's strategic choices by assuming that the sender challenges with probability x, does not challenge with probability 1-x, opts for a public approach given a challenge with probability y, and opts for a private approach given a challenge with probability 1-y. This means that conditional probability can be represented as,

$$\Pr(\sim C \mid B) = k = \frac{1-x}{(1-x)+xy}.$$

Substituting the value of k^* derived above for k yields an equilibrium strategy where the sender challenges with probability,

$$x^* = \frac{U_T(F) - U_T(Q)}{U_T(F) - U_T(Q) - U_T(R)y + U_T(S)y},$$

or opts for a private challenge with probability,

$$y^* = \frac{U_T(Q) - U_T(F) + U_T(F)x - U_T(Q)x}{(U_T(R) - U_T(S))x}.$$

Unfortunately, this is not yet sufficient to specify a full equilibrium. Close examination of x^* and y^* shows that the two strategies are contingent upon one another. The likelihood of challenge is contingent upon the likelihood of private action, and the likelihood of private action is contingent upon the likelihood of challenge. This means there is no clear indication of equilibrium behavior yet. To complete the equilibrium, we have to identify the combination of x^* and y^* that maximizes the sender's expected utility.

If the sender opts for a strategy that yields an ambiguous signal, the target will play their equilibrium strategy n^*. When the target acts with probability n^*, the sender's expected payoff from the ambiguous signal is,

$$EU_S(B) = \frac{U_S(Q)U_S(R) - U_S(F)U_S(S)}{U_S(Q) + U_S(R) - U_S(F) - U_S(S)}.$$

The sender strategy is dictated by their preference for public standoff ($U_S(D)$) relative to the expected payoff from sending the ambiguous signal ($EU_S(B)$). When the sender

prefers a public standoff, $U_S(D)>EU_S(B)$, then the sender adopts a strategy where they always challenge publicly. This means that $x^*=1$ and $y^*=0$.

Alternatively, when the expected payoffs from playing an ambiguous strategy are preferred to public standoff, $EU_S(B)>U_S(D)$, there is no incentive for the sender to challenge publicly. The sender has incentive to maximize the probability they play strategies that involve the ambiguous signal, or set $xy+(1-x)$ as high as possible. Minimizing the likelihood of a public standoff means never challenging publicly or setting $1-y^*=0$, which also means that all challenges are private and $y^*=1$. When $y^*=1$, the sender challenges in equilibrium with probability,

$$x^* = \frac{U_S(F) - U_S(Q)}{U_S(F) - U_S(Q) - U_S(R) + U_S(S)}.$$

We have identified two possible equilibria given the preference for public standoff. The equilibria are listed in the following order: {the sender's decision to challenge or preserve, the sender's decision to use public or private means, the target's decision to act or not given an ambiguous signal, the target's decision to act or not given a public challenge, and the target's beliefs in equilibrium}. When $U_S(D)>EU_S(B)$, the equilibrium is {C, M, ~ n^*~W $(1-n^*)$W, ~W, $k=0$}. The sender always challenges publicly, the target always acts, and there is no reason to believe the sender ever preserves the status quo. When $EU_S(B)>U_S(D)$, there is a mixed equilibrium {x^*C $(1-x^*)$~C, ~M, n^*~W $(1-n^*)$W, ~W, $k=k^*$}. The sender mixes between challenge and preserve, always uses private means when challenging, the target mixes between acting and waiting, and the target has the critical belief that makes them indifferent. These results are discussed further in the next section.

Discussion. There are two equilibria of the basic surprise model contingent on the sender's value for a public standoff. When the sender has a high value for a public standoff, both actors adopt pure strategies. The sender always challenges publicly, and the target always takes action in response. In this case, the sender never attempts surprise or sends ambiguous signals. If the sender prefers to avoid a public standoff, both players adopt mixed strategies. The sender mixes between preserving the status quo and private challenges, and the target mixes between acting and waiting. The sender never challenges publicly, implying that the target only observes the ambiguous signal and forms some belief about the likelihood of surprise based upon the equilibrium strategy.

The first result discussed above, the pure strategy equilibrium ending in public standoff, is relatively straightforward. If the sender's payoff for a public standoff is greater than the expected payoff for attempting surprise, the sender challenges publicly knowing that the target will take action. It is also worth noting that the payoff for the public standoff is a sure thing, whereas the payoff for the mixed strategy carries uncertainty. A sender who is averse to uncertainty is also likely to prefer public standoffs to mixed strategies that will yield one of four possible outcomes. This result, however, is relatively uninteresting since it predicts deterministic outcomes with complete information. It predicts the same specific course of action for every play of the game, and the result is always a public standoff.

The mixed strategy result is substantively more interesting and the comparative static results are summarized in Table 4.1. In the mixed strategy equilibrium, the sender mixes between attempted surprise and preservation of the status quo, while the target mixes acting and waiting. This means that there is some probability the status quo is

maintained, there are false alarms, there are thwarted surprises and there are successful surprises. Table 4.1 shows the predicted changes in the indifference belief k*, the sender's optimal strategy *x** (the probability the sender challenges privately), and the target's optimal strategy *n** (the probability the target acts). These comparative static results are reviewed at greater length below.

Table 4.1: Predicted Changes in the Imperfect Information Model

	*k** (Threat Threshold)	*x** Pr(Attempted Surprise)	*n** Pr(Act)
$U_T(Q)$	-	+	0
$U_T(F)$	+	-	0
$U_T(R)$	+	-	0
$U_T(S)$	-	+	0
$U_S(Q)$	0	0	-
$U_S(F)$	0	0	-
$U_S(R)$	0	0	+
$U_S(S)$	0	0	+

Note: The table reflects partial derivatives, the changes in beliefs and strategies conditional on an increase in each preference value.

One of the central results from the Basic Surprise Model is that successful surprise is a probabilistic outcome of rational action. Successful surprise is not conditional on breakdowns in bureaucratic machinery or biased analysis. Surprise is the result of uncertainty, in this case imperfect information about the sender's past moves, and the actors' preferences over outcomes. Both players respond rationally, and the target has rational beliefs about the sender's previous play that are objectively accurate according to Bayes' Rule. This basic model shows that mistakes, biases or misinformation campaigns are not necessary to explain surprise. Private information is sufficient to explain surprise (Whaley 1969; Jervis 1970).

The model also offers insights into the likelihood of attempted and successful surprise. The sender is more likely to attempt surprise as the target places higher values on preserving the status quo and suffering surprise. Both of these outcomes occur when the target chooses to wait. As the values for the status quo and surprise increase, the target has more incentive to wait, and the sender has greater incentive to challenge. Alternatively, the likelihood of attempted surprise decreases as the target's preference for false alarms and thwarted surprise increases. This provides the target with greater incentives to act, which reduces the sender's incentive to attempt surprise. A second conclusion derived from the complete information model is that the sender is more likely to attempt surprise when the target has a high value for the status quo and is not particularly concerned about suffering surprise. When the target has a high value for status quo and high value despite suffering surprise, the target has less incentive to take action. The sender is more likely to attempt surprise when the target is less likely to take action since it increases the likelihood that sender's surprise is successful. This means that targets with high values for the status quo and surprise are more likely to face surprise attempts.

Attempted surprise is only one half of successful surprise. The other element that determines successful surprise is the target's decision to take action. The target is more likely to take defensive action as the sender's preferences for successful or thwarted surprises increases. Combined, this increases the sender's incentive to attempt surprise, which unsurprisingly increases the target's incentive to act. The target is more likely to wait as the sender's value for the status quo increases since the sender has less incentive

to challenge. Surprise, therefore, is likely to be successful when the target places a high value on the status quo and surprise, and the sender places a high value on the status quo.

We can move from the abstract players and outcomes in the model towards a realistic assessment of the type of states that are likely victims of surprise. Our assessment draws from the partial derivatives in Table 1 showing how a change in any single preference impacts expected outcomes holding all other preferences constant. States are more likely to be the target of surprise when they value the status quo, have high costs for false alarms, have a low utility for thwarting surprise, and are not severely injured by a successful surprise. Traditionally, great powers are perceived to have a high value for the status quo[7] and are capable of absorbing a surprise attack due to the size and scope of their military force. Major successful surprise attacks targeting great powers, such as the German attack on Soviets in 1941 and the Japanese attack on Pearl Harbor in 1941, shows the resilience of these victims. This means that great powers should also have lower utility for thwarting the surprise since they expect to absorb the attack. Nuclear weapons change this calculus somewhat by empowering smaller rivals that might be capable of imposing significant losses on large militaries and populations. The costs of false alarms and the benefits of thwarting surprise also depend on the costs associated with defensive mobilization. Both large and small states may have high mobilization costs, however, it is particularly costly for countries that rely on large civilian reserves such as Israel. Mobilizations can grind the economy to a halt when reservists are pulled away from their civilian jobs.

These results from the Basic Surprise Model are analogous to existing theories of surprise, but this work points to slightly different causal mechanisms and conclusions.

[7] This point is made in pieces such as Kindleberger 1970, Krasner 1976 and Snidal 1985.

Theories of surprise usually maintain that surprise attacks are used by weaker powers against stronger foes. Surprise becomes a force multiplier giving the weaker state a greater chance for military victory (Gooch and Perlmutter 1982; Handel 1989; Kam 1988). The analysis suggests that force differential establishes a situation where surprise is used, but it does not dictate the likelihood of surprise. Recall that states with a high value for public standoffs, presumably because they possess stronger militaries or international support, do not use surprise at all. Once the public standoff is unpalatable, powerful status quo countries are more likely to be targets of attempted surprise, but the reason is not driven by the force differential. The reason these states are more likely to be targeted is because the surprise is more likely to succeed. The analysis also showed that those facing high costs of mobilization are also more vulnerable to surprise. This leads to the conclusion that great powers and small countries may be particularly attractive targets for attempted surprise.

States are more likely to take defensive precautions against possible surprise when their opponent places high value on attempted surprise, whether they are successful or not. When potential initiators have a high value for status quo, there is less of a reason to expect surprise, and potential targets are less likely take defensive action. Here we see observe a situation typical of the paradoxical nature of military strategy (Luttwak 1987). Surprise attacks are more likely to be successful when the initiator is perceived to be a status quo power. Those with little reason to attack are more likely to achieve surprise. States believed to be revisionist are less likely to achieve surprise, an issue that will be investigated closer at the end of this chapter and empirically in the next chapter.

Surprise attacks are more likely to be successful when they are carried out by content or status quo states that target great powers or small states. Attacks are more likely to be thwarted when carried out by revisionist initiators targeting great powers or small states. Senders are more likely to achieve surprise when they appear or act as though they are content before any attack. Targets are more likely to experience false alarms when they are mid-sized powers facing revisionist initiators, and the status quo is most likely to be upheld when content initiators face mid-sized powers.

The basic model showed how surprise is a product of uncertainty, and the way that uncertainty and state preferences over outcomes can combine to influence the likelihood of different outcomes. This basic exercise has already improved our understanding of the necessary and sufficient conditions for surprise. We showed that surprise was a possible outcome whenever the sender preferred the mixed strategy to a pure public challenge strategy $EU_S(B)>U_S(D)$, the sender attempts surprise as part of the mixed strategy when $U_T(Q)>U_T(F)$, and the target acts as part of the mixed strategy $U_S(S)>U_S(Q)$. In the next section, we introduce a model with additional information constraints. We use this model to evaluate propositions about entrenched beliefs and surprise, and show why targets may benefit from deliberate analytical closure or entrenched beliefs.

The Surprise Model with Incomplete Information

Since the Basic Surprise Model only incorporated imperfect information, the target only formed beliefs about the sender's past behavior. The target had complete information about the sender's preferences and had no need to form beliefs about the

sender's intentions. When we add incomplete information to the model, the target is concerned with the sender's previous behavior and their intentions or type. The target learns about the sender and updates their beliefs accordingly. By adding incomplete information, we can incorporate the possibility of bias where targets are slow to update their prior beliefs, and then examine the impact of these biases on expected outcomes.

Incomplete information is added to the Surprise Model by assuming that there are two possible types of senders. This was discussed at some length in the previous chapter and will be reviewed briefly here. Different types are incorporated to add uncertainty about an actor's preferences. In this model, we assume that the sender can be aggressive (A) or docile (~A). The actual type of sender is drawn from a probability distribution, which is often called a move by nature. After nature selects the type of sender, the game proceeds in the exact same fashion as the basic model above: the sender can challenge or not, the challenge can be public or private, and the target decides whether to act or not.

In this model, as in many models of incomplete information, the probability distribution of the sender's type is common knowledge. This means that both actors know the likelihood with which nature will select an aggressive type, and both actors know they know, *ad infintum* (Morrow 1994a). The target knows this probability, but it does not whether nature actually selects the aggressive or docile sender type. The sender actually observes which type is drawn and knows whether they are aggressive or docile when selecting their strategy. Since the sender knows its type, and the target only knows the probability of the sender's type, there is asymmetric information. The sender has information about its type that is not available to the target. The target uses the common knowledge probability as its prior belief, which is combined with the sender's signal to

form a posterior belief.[8] The likelihood that nature select the aggressive sender, Pr(A), is *p*. Since there are only two types of senders in this model, the likelihood the sender is docile is 1-*p*.

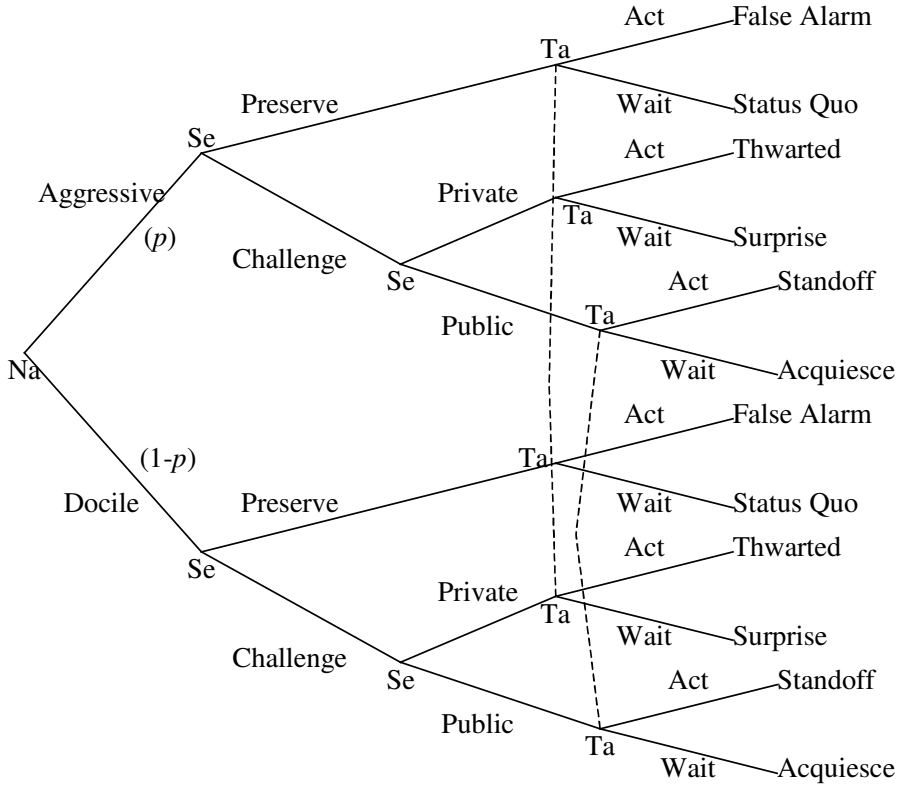
Introducing two types of sender means that we also have to specify the preferences associated with each type. The preferences specified here are intended to reflect a realistic preference structure, while keeping the preference order simple enough to derive parsimonious solutions. Again, later in the chapter we will look at alternative plausible preference structures and the impact on expected outcomes. First, the sender places different values on outcomes when they are aggressive and docile. Outcomes for the aggressive sender are indicated in capital letters, so the preference notation is unchanged from the section above. The docile sender has different preferences, and the values placed on these outcomes are indicated by lowercase letters corresponding to the capital letters above ($U_S(q)$, $U_S(f)$, $U_S(r)$, $U_S(s)$, $U_S(d)$, $U_S(e)$). For example, the aggressive sender's payoff for surprise is $U_S(S)$ and the docile sender's payoff is $U_S(s)$. All of the docile sender's outcome values are less than or equal to the aggressive type's outcomes.

There are two assumptions about the sender's preferences that are helpful for solving the game. We assume that the aggressive sender strictly prefers a public standoff to the status quo. Since the aggressive sender can ensure a public standoff by publicly challenging, the aggressive sender will never preserve the status quo. This means that the aggressive sender always challenges, only deciding whether to challenge publicly or privately. The docile sender strictly prefers the status quo to all other options, which

[8] This was also covered in the previous chapter, but will be reviewed in the solutions section.

means they have incentive to preserve the status quo all of the time.[9] Combining these two assumptions with the preference ordering above is sufficient to solve for the game's equilibria. The incomplete information model is found in Figure 4.2.

Figure 4.2: Surprise Model with Incomplete Information



Solutions. We can solve the incomplete information model using the indifference technique in the previous section. Recall that the equilibrium solutions required that each player select the strategy that makes their opponent indifferent between available moves, if such a strategy exists. In these cases, neither actor can do any better by unilaterally changing their strategy. The primary difference between the incomplete information and

[9] This is an assumption that is revisited later in the chapter. It is not as restrictive as it first appears.

basic surprise models is the role of prior and posterior beliefs. A solution profile needs to dictate the way that both types of senders behave, the target's response and the target's posterior beliefs.

In the incomplete information model, the aggressive sender challenges with probability *X* and attempts surprise upon challenging with probability *Y*. The docile sender challenges with probability *x* and attempts surprise with probability *y*. However, the assumptions discussed above limit the sender's possible strategies. If the aggressive sender always prefers to challenge, then *X*=1. When the docile sender does not challenge, *x*=0, and by extension there is no need for the docile sender to select a type of challenge. The sender completes their strategic profile by selecting public/private challenge strategy *Y* that makes the target believe they are indifferent if possible.

The target knows that the aggressive sender always challenges and the docile sender does not. In the absence of a public challenge, the target is indifferent between acting and waiting when,

$$p^{*}U_{T}(S)+(1-p^{*})U_{T}(Q)=p^{*}U_{T}(R)+(1-p^{*})U_{T}(F),$$

where p^{*} is the critical belief that makes the target indifferent. Solving for the critical belief yields the following condition,

$$p^{*}=\frac{U_{T}(F)-U_{T}(Q)}{U_{T}(F)+U_{T}(S)-U_{T}(Q)-U_{T}(R)}.$$

The constraint for p^{*} shows that threat perceptions is actually endogenous to the model. When the target believes that sender is aggressive more often than this critical value, $p>p^{*}$, the sender is perceived to be threatening and target always acts. Conversely, if the target believes that the sender is aggressive less often then the critical value, $p<p^{*}$, sender is perceived docile and the target prefers to wait. While the target knows the probability

that the sender is aggressive and docile, this probability alone does not warrant a threat. The prior belief and the critical threshold determine the threat level.

The sender adopts a challenge strategy such that target's posterior belief is equal to p^*, if such a strategy is possible. The target begins with the prior belief that the sender is aggressive with probability p, and their Bayesian posterior belief about the sender's type is,

$$\Pr_{Bayes}(A \mid C \sim M) = \frac{\Pr(A)\Pr(C \sim M \mid A)}{\Pr(A)\Pr(C \sim M \mid A) + \Pr(\sim A)\Pr(\sim C \mid \sim A)} = \frac{pXY}{pXY + (1-p)(1-x)}.$$

We can derive the sender's mixed equilibrium strategy by setting the critical belief equal to the Bayesian posterior belief, $p^* = \Pr_{Bayes}(A|C\sim M)$. The aggressive sender attempts surprise with probability,

$$Y_e = \frac{U_T(F)p + U_T(Q) - U_T(Q)p - U_T(F)}{p(U_T(R) - U_T(S))}.$$

Taking defensive action is still the target's dominant strategy given a public challenge. In the absence of the public challenge, the target adopts a strategy aimed at making the aggressive sender type indifferent between public and private challenges. The docile sender is never indifferent since they prefer to preserve the status quo. The target's strategy that makes the aggressive sender indifferent must meet the following condition,

$$nU_T(R) + (1-n)U_T(S) = U_T(D),$$

where n is the probability that the target acts. This means that the mixed strategy equilibrium for the target is,

$$n^* = \frac{U_S(S) - U_S(D)}{U_S(S) - U_S(R)}.$$

As long as the aggressive sender has a higher value for a public standoff than a thwarted surprise, the target's mixed strategy equilibrium exists.

Now we can characterize the set of equilibria for the incomplete information game. There are two standard Bayesian equilibria for this game, one based on pure strategies and the other based on mixed strategies. The equilibria are listed as follows: {the sender's strategy when aggressive; the sender's strategy when docile; the target's strategy when observing a public challenge; the target's strategy when observing the ambiguous signal: the target's belief}. The pooling equilibrium is {C, ~M; ~C; A; ~A: $p<p^*$}. A pooling strategy is a choice set in which the sender transmits the same signal to the target all of the time. In this case, the sender either challenges privately or does not challenge, thereby always sending the ambiguous signal. The pooling equilibrium exists whenever the prior likelihood that the sender is aggressive, p, is below the critical value p^*. In this case, the sender is perceived to be non-threatening, and this assessment is endogenous to the model. When the prior belief equals the critical belief, $p=p^*$, the sender always attempts surprise given a challenge, $Y=1$. Mathematically, $Y>1$ when $p<p^*$, implying that the aggressive sender adopts a pure strategy of private challenge. The aggressive sender always attempts surprise and the docile sender always preserves the status quo. The target always waits since they believe they are more likely to face a docile sender. This pure strategy is also a pooling strategy of sorts since the sender always acts in order to send the ambiguous signal, irrespective of type. The only two expected outcomes are status quo and surprise.

The mixed strategy equilibrium is {C, Y_e ~M, (1-Y_e) M; ~C; A; n^* A, (1-n^*) ~A: $p=p^*$}. The mixed strategy is characterized in terms of the sender's equilibrium

probability of challenging by surprise Y_e, the target's equilibrium probability of acting n^* after observing the ambiguous signal, and the target's *ex post* belief that they are facing an aggressive type p^*. The common knowledge assumption dictates that both the sender and the target know all of the information necessary to derive these three equilibrium constraints. The target does not know what type they are facing, and cannot distinguish private challenges from attempts to preserve the status quo, but they do know the preferences for each type of the sender and the sender knows the target's preferences as well.

In the mixed strategy, the aggressive sender attempts surprise with probability Y_e and challenges publicly with probability $1-Y_e$. The docile sender always attempts to preserve the status quo. The target acts with probability n^*, waits with probability $1-n^*$, and believes that sender is aggressive with probability p^*. In the mixed strategy equilibrium, the game can end with status quo, false alarm, successful surprise, thwarted surprise or public standoff. We examine the likelihood of outcomes and partial derivatives in the discussion section below. It is worth noting that this solution actually becomes a corner solution, making specification of an additional equilibrium unnecessary. A corner solution is one where the strategies are bounded by some value, often on the interval [0,1], and the equilibrium dictates that a player adopts a strategy that permanently bounds them at the minimum or maximum value. When $p=100\%$, then $Y_e=0$. This means that a sender who is always aggressive always challenges publicly, and the target always takes defensive action. Since the semi-separating equilibrium completely characterizes this possibility, there is no need for a further equilibrium.

Discussion. The analysis above identified two equilibria for the incomplete information model. When the sender was perceived to be non-threatening, had a sufficiently low prior likelihood of being aggressive, there was a pooling equilibrium where the sender always conveyed ambiguous signals. The docile sender always attempted to preserve the status quo and the aggressive sender always attempted surprise. The target always waited in response. When the sender was perceived to be a threat, or the prior likelihood of being aggressive surpassed the critical belief, the sender adopted a mixed or semi-separating strategy. The docile sender would preserve the status quo, but the aggressive sender would mix between public and private challenges. In response, the target would play a strategy that mixed waiting and acting. Table 4.2 shows the predicted effects of preference changes on equilibrium strategies based on partial derivatives. The table highlights comparative static relationships that are addressed in greater detail below.

Table 4.2: Predicted Changes in the Incomplete Information Model

	p^* (Critical Belief)	Y_e Pr(Attempted Surprise)	n^* Pr(Act)
p	0	-	0
$U_T(Q)$	+	+	0
$U_T(F)$	-	-	0
$U_T(R)$	-	-	0
$U_T(S)$	+	+	0
$U_S(Q)$	0	0	0
$U_S(F)$	0	0	0
$U_S(D)$	0	0	-
$U_S(R)$	0	0	+
$U_S(S)$	0	0	+

Note: The table reflects partial derivatives, the changes in beliefs and strategies conditional on an increase in each preference value.

A non-threatening sender, one with a sufficiently low likelihood of being aggressive, adopts the pooling strategy because there is nothing they can do, or no signal they can send, that can make the target take defensive action. Since the target starts the game with the prior belief that the sender is sufficiently docile to be perceived as non-threatening, the aggressive sender does not have to make much of an effort to mask their intentions. The target's prior beliefs provide an umbrella of sorts that is sufficient to cover the aggressive sender's planned surprise. If the target has beliefs that prevent it from taking defensive action, any attempted attack will be a successful surprise. The target's decision to wait is a product of the high likelihood that sender is docile and therefore preserving the status quo. Should the target take action, there is a high likelihood that they will suffer a false alarm. The target finds they are better off waiting and suffering the rare surprise than adopting a mixed strategy that ensures many false alarms.

The semi-separating result is substantively more interesting, and it is the equilibrium when the sender is perceived to be a threat, specifically when the prior belief the sender is aggressive exceeds the critical threshold. Before examining the characteristics of the equilibrium, it is worth taking a moment to look at the critical belief that determines whether the players adopt the pooling or semi-separating equilibrium and the target's threat assessment. The critical belief value increases as the target's utility for waiting increases. In other words, as the target's value for the status quo or surprise outcomes increase, the critical belief that makes the sender indifferent increases as well. Conversely, as the values of outcomes associated with taking action, false alarm and thwarted surprise, increase, the critical belief value decreases. This means that states with

a large incentive to thwart a surprise and low costs for false alarms have lower critical belief thresholds. These states are more likely to perceive the sender to be threatening, have greater incentive to take action, and the sender is more likely to adopt a semi-separating strategy. Pooling outcomes are more likely when the target state is not really worried about surprise and have a high value for the status quo. This shows how the target's preferences over outcomes and the sender's prior likelihood of aggression are tied to endogenous threat perception and the different equilibria deduced above.

In the semi-separating equilibrium, the aggressive sender mixes between public and private challenges. Here, the sender is unable to hide behind the target's prior belief that the sender is likely to be docile. If the aggressive sender attempted surprise every time, the target would always have incentive to take action. The sender avoids this by publicly challenging just often enough that the target's posterior belief that sender is aggressive given the ambiguous signal is equal to the critical belief. As the sender's prior likelihood of being aggressive increases, the sender needs to publicly challenge more often to ensure that the target's posterior belief remains equal to the critical belief. Predictably, the sender should attempt surprise less often as the likelihood they are aggressive increases. After all, it should be more difficult for a state that is always aggressive to catch a target state by surprise.

The sender's semi-separating equilibrium strategy is a function of the likelihood they are aggressive and the target's preferences over outcomes. The aggressive sender is more likely to attempt surprise when the target has higher values for the status quo and surprise. In technical terms, the likelihood the sender attempts surprise is increasing in the target's utility for the status quo and false alarms. The logic here is similar to the logic

that applied to the critical belief value. When the target's utility for the status quo and surprise increase, that means the target's value of waiting has increased. The aggressive sender has more incentive to attempt surprise, sending the ambiguous signal, knowing that the target has greater incentives to wait. The sender can attempt surprise more often while keeping the target indifferent between waiting and acting. Conversely, the sender's attempted surprise is decreasing in the target's utility for false alarms and thwarted surprise. When false alarms are not costly, and thwarted surprise is especially beneficial, the sender has to attempt surprise less often in order to keep the target indifferent given the greater incentives to take action. Finally, the sender's likelihood of attempting surprise is decreasing in the prior probability that the sender is aggressive. If the target knows the sender has a higher likelihood of being aggressive, the sender can only keep the target indifferent by attempting surprise less often.

The target's strategy in the semi-separating equilibrium is a function of the sender's preferences over outcomes. The aggressive sender can always opt to challenge the target publicly thereby securing the payoff from a public standoff. The aggressive sender will only mix between public and private challenges when the expected utility of the two challenges is equal. If the expected utility from public challenge is greater than that of private challenges, the aggressive sender will always challenge publicly. The target, therefore, selects a strategy that makes the sender indifferent between public and private challenges. The target accomplishes this by mixing acting and waiting, given an ambiguous signal, such that the sender's expected utility from private challenges is equal to the expected utility in a public challenge. The sender is less likely to take action when the sender's utility for a public standoff increases. The sender has greater incentive to opt

for public challenge, and the target can wait more often as a response. The likelihood that the target takes action in the mixed strategy is increasing in the sender's value for successful and thwarted surprise. When the sender gets greater value for attempting surprise, the target responds by acting more often to increase the likelihood that sender suffers a thwarted surprise. This makes attempted surprise less attractive to the aggressive sender, thereby ensuring they remain indifferent between their policy choices.

The substantive results here are similar to those of the basic surprise model, with the exception that prior beliefs about the sender play an important role. States with a high value for status quo that are not significantly harmed by surprise attacks are more likely to be great powers. Those with a low value for status quo, or are particularly vulnerable to surprise attacks, are more likely to be small and insecure states. These small states, such as Israel, are probably disproportionately affected by false alarms since mobilization is a statewide endeavor that brings a halt to the economy and everyday life. Small or weak states are also likely to place greater emphasis on thwarting surprise since the state may be unable to mount a military response after taking the initial blow. This means that the critical belief value, the belief that makes target states indifferent between acting and waiting, should be higher for great powers than small states. This means great powers are more likely to find themselves in the pooling equilibrium whereby the target never has incentive to act because the initial prior and perceived threat is not sufficient to warrant concern. Small powers, with a lower critical belief, are more sensitive to possible threats meaning they face opponents opting for a semi-separating strategy.

These results extend to the sender's propensity to attempt surprise in the semi-separating equilibrium. Since great powers usually have a higher value for the status quo

and are less vulnerable to initial attacks, challenging states have a greater incentive to attempt surprise. This analysis suggests that the classic analogy, great powers are slumbering giants, may be correct. Challenging states know they can get away with attempting surprise more often while keeping the target indifferent. Small powers with lower costs for false alarms, greater incentives to thwart the surprise, and high costs for successful surprise, should face attempted surprise less often. It is more difficult to lull these states into a sense of security.

The target state is less likely to take action as the sender's preference for public standoffs increases. Challengers are likely to have a high value for public standoffs when they are great powers, since great powers should have more bargaining leverage in the subsequent standoff. Conversely, targets are more likely to act as the challengers value for successful and thwarted surprises increase. Interpreting this result is a little more opaque. Small states are more likely to benefit from successful surprise, since surprise is supposed to be a force multiplier. These same small states, however, are likely to suffer significant costs when these attempts get thwarted. Large or great powers would prefer to avoid a thwarted surprise, but these instances will rarely be disastrous. It seems as though targets are more likely to take action against small powers, those that benefit from surprise, but this generalization may be overly broad given the deductive results.

In this model, the likelihood that a target suffers surprise is tied to the prior probability that the sender is aggressive. The target is most likely to suffer surprise when the prior probability of sender aggression is equal to the posterior critical belief value. As the prior probability declines from the critical belief, a waiting target is less likely to face an aggressive sender in the pooling equilibrium. If the prior likelihood the sender is

aggressive increases, the sender should attempt surprise less often if any attempts at surprise are to be successful in the semi-separating equilibrium. We also argued that senders are more likely to attempt surprise against larger powers, but it was difficult to classify the target's propensity to respond with defensive action.

The next model incorporates the possibility of a biased target. After solving for the equilibria of the model, we compare the results from the Bayesian model discussed above against the potentially biased model, revealing some unexpected features of bias.

The Surprise Model With Bias

The game structure of the potentially biased model introduced here is almost identical to the incomplete information game introduced in the previous section. There are two players, a sender and target. There are two possible types of sender, aggressive and docile, and the actual type is drawn by nature. The sender can choose to preserve or challenge the status quo, and challenges can be public or private. The target then decides whether to take defensive action or wait. The players' preferences over outcomes are the same as the surprise model with incomplete information.

The only difference here is that we include two possible types of targets. Deviating from conventional use of types, where actors usually have different preferences over outcomes, the two types of targets have identical preferences. The difference between the two types is the way they update their beliefs. In this model, the target can be perfectly Bayesian or biased. When the target is Bayesian, they update their beliefs using the Bayesian method in the previous section where posterior beliefs are,

$$\Pr_{Bayes}(A \mid C \sim M) = \frac{\Pr(A)\Pr(C \sim M \mid A)}{\Pr(A)\Pr(C \sim M \mid A) + \Pr(\sim A)\Pr(\sim C \mid \sim A)} = \frac{pXY}{pXY + (1-p)(1-x)}.$$

A biased target, on the other hand, suffers from deeply entrenched beliefs, which were introduced in Chapter 3. A target suffering from deeply entrenched beliefs irrationally discounts or ignores new information relative to their prior beliefs. The posterior belief resulting from this biased updating process is specified as weighted average between the prior belief and the Bayesian rational posterior such that,

$$\Pr_{Bias}(\sim D \mid C \sim M) = bq + (1-b)\Pr_{Bayes}(\sim D \mid C \sim M).$$

The variable *b* represents the strength of the deeply entrenched belief. When *b* equals zero, the target is perfectly Bayesian, and when *b* equals one, the target ignores all new information. We use a point estimate for the level of anchoring bias and introduce uncertainty by offering the possibility of a Bayesian or biased target.

This revised game begins with the selection of the target's type, where the target may update beliefs in an entrenched or Bayesian fashion. With probability *z* the target is entrenched and is Bayesian with probability 1-*z*. The sender knows the probability *z*, but neither player knows which type is selected. This assumption means that the sender has incomplete information about the target's belief updating quality. The sender knows that the target may have entrenched beliefs, but does not know whether the target will actually exhibit entrenched behavior. In other words, the sender knows that there is some probability *z* that the target will have anchoring bias *b*. This assumes that the sender has some knowledge about the receiver's preexisting beliefs and potential biases. While these assumptions simplify the types of uncertainty faced by the sender, this is also a plausible assumption in international politics. Decision-makers are always interested in their opponent's beliefs and preconceived notions, and they are constantly trying to gather additional information on potential opponents. This assumption also ensures that the

target believes they are updating their information throughout the game as best as possible. They are unaware whether they are using the Bayesian or entrenched mechanisms. If the target knew which they were using, they could simply adjust their calculations accordingly. Such adjustments would cancel out any strategic impact of entrenched beliefs.

After the target's type is selected, the sender's type is selected by nature and the actors make their moves. The next section reports the deductive results for the potentially biased model, and then these results are compared to the standard Bayesian results.

Solutions. This section uses the same notation and assumptions that were used in the incomplete information model. As a review, the aggressive sender challenges with probability X and attempts surprise with probability Y. The docile sender challenges with probability x and attempts surprise with probability y. The aggressive sender always prefers to challenge, meaning X=1, and the docile sender does not challenge, so x=0. The sender completes their strategic profile by selecting a public/private challenge strategy Y that makes the target believe they are indifferent. Since there are two types of targets, Bayesian and biased, we have to solve for the indifference characteristics of both types. The notation Y_e is used to reflect the public/private challenge strategy that makes the Bayesian type indifferent, and Y_b is used to denote the challenge strategy that makes the biased type indifferent. In the rest of the chapter we refer to Y_e as the Bayesian-aimed strategy and Y_b as the biased-aimed strategy.

Once again, the target is indifferent between acting and waiting upon observing the ambiguous signal when,

$$p^* U_T(S) + (1 - p^*) U_T(Q) = p^* U_T(R) + (1 - p^*) U_T(F),$$

where p^* is the critical belief. The target's critical belief remains unchanged,

$$p^* = \frac{U_T(F) - U_T(Q)}{U_T(F) + U_T(S) - U_T(Q) - U_T(R)}.$$

Recall the target always waits when their posterior belief $p<p^*$, and always acts when the posterior belief $p>p^*$. The sender's equilibrium is aimed at making the target believe that the sender is aggressive with probability p^*. In the potentially biased model, the Bayesian type updates beliefs as specified in the solution set above. The strategy that makes that Bayesian target indifferent is,

$$Y_e = \frac{U_T(F)p + U_T(Q) - U_T(Q)p - U_T(F)}{p(U_T(R) - U_T(S))}.$$

This strategy, however, cannot make the biased target indifferent. Since the biased target entrenches upon the prior belief about the sender, the strategy that makes the biased target indifferent is,

$$Y_b = \frac{-(U_T(F)(bp - 1) + U_T(Q) - bp(U_T(R) + U_T(Q) - U_T(S)))(p - 1)}{p(U_T(R)(1 + b(p - 1)) - U_T(S) - b(p - 1)(U_T(F) + U_T(S) - U_T(Q)))}.$$

The target once again selects a strategy that makes the aggressive sender type indifferent between public and private challenges such that,

$$nU_S(R) + (1 - n) U_S(S) = U_S(D).$$

This means that the target's mixed strategy aimed at making the sender indifferent is the same for both models, which is not surprising since the result is contingent upon the sender's payoffs. The strategy that makes the sender aggressive indifferent is,

$$n^* = \frac{U_S(S) - U_S(D)}{U_S(S) - U_S(R)}.$$

The equilibrium results here are similar to those deduced in the Bayesian model above. We will specify these equilibria for completeness without addressing redundant details. There are two standard Bayesian equilibria for this game, one based on pure strategies and the other based on mixed strategies. The equilibria are listed as follows: {the sender's strategy when aggressive; the sender's strategy when docile; the entrenched target's strategy when observing a public challenge; the Bayesian target's strategy when observing a public challenge; the entrenched target's strategy when observing the ambiguous signal; the Bayesian target's strategy when observing the ambiguous signal: the target's belief}. The pure strategy equilibrium for the potentially biased is {C, ~M; ~C; A; A; ~A; ~A: $p<p^*$}, and the semi-separating strategy is {C, Y_b ~M, (1-Y_b) M; ~C; A; A; n A, (1-n) ~A; ~A: $p=p^*$}. The rationale is similar, when the prior belief about the sender is below the critical value, the aggressive sender can always achieve surprise. This results in a pure strategy equilibrium where the sender pools on the ambiguous signal. When prior beliefs exceed the threshold value, the sender attempts surprise just often enough to keep the biased target indifferent. Unlike the previous semi-separating equilibrium, there is an additional corner solution contingent on the anchoring bias. We find it is possible that Y_b=0 even when p<100%. When the target is biased there is a threshold on the anchoring value such that Y_b=0 when,

$$b^* = \frac{U_T(F) - U_T(Q)}{p\left(U_T(F) - U_T(Q) - U_T(R) + U_T(S)\right)}.$$

If $b>b^*$, the target is too biased to be indifferent and the sender gets no benefit from attempting deception. If we add bias to the equilibrium, we can characterize the corner solution as {C, M; ~C; A; A; A; ~A: $p>p^*$; $b>b^*$}. Next, we turn to the sender's optimal strategy.

At first glance, it is difficult to compare the sender's optimal strategies in the two models, Y_e and Y_b. If however, we substitute p^* for the prior belief p, the sender's likelihood of attempted surprise is the same, or $Y_e=Y_b$. For any value where the prior probability the sender is aggressive is greater than the critical value, $p>p^*$, then $Y_e>Y_b$. When there is the possibility of target bias, the sender must decide whether to play Y_b or Y_e. Biased targets always mix at Y_b and act at Y_e, and Bayesian target always waits at Y_b and mixes at Y_e. To find the sender's best response, we set $EU_S(Y_e)=EU_S(Y_b)$, and find that $EU_S(Y_b)>EU_S(Y_e)$, whenever,

$$z < \frac{1+U_S(S)Y_b - U_S(D)(1+Y_b)}{1+U_S(S)Y_b + U_S(R)Y_e - U_S(D)(1+Y_b+Y_e)}.$$

The value of the threshold above is always greater than one, and since $0<z<1$, the sender always enjoys a higher utility from Y_b. The sender's decision is summarized in Table 4.3. It shows how each target responds to each of the sender's possible strategies. Since the both targets will play mixed strategies, albeit under different conditions, the sender is better off adopting a strategy that encourages the target to wait.

Table 4.3: Sender Strategies and Target Response

	Bayesian-aimed Strategy (Y_e)	Biased-aimed Strategy (Y_b)
Bayesian Target Type	Target type plays mixed strategy	Target type **always waits**
Biased Target Type	Target type **always acts**	Target type plays mixed strategy

Note: The table shows that the Bayesian-aimed strategy will elicit action form the biased type, while the biased-aimed elicits waiting from the Bayesian type. The sender is always better off adopting the biased-aimed strategy Y_b.

This solution section concludes by examining the impact of entrenched beliefs on the target's utility and likelihood of surprise. To examine the target's change in utility from the possibility of entrenchment, the values for surprise and status quo are normalized. Surprise is the target's least or second least preferred option, so it is assumed that it yields a payoff of zero ($U_T(S)=0$). The target's most preferred option, status quo, is assumed to be equal to one ($U_T(Q)=1$). Under these conditions, the target's expected utility given potential entrenchment is higher whenever,

$$b > \frac{U_T(R)(U_T(D)-1)(z-1)}{(1+U_T(R)-U_T(F))(p-1)(U_T(D)-U_T(R)nz)}.$$

Since $0<b<1$, this condition is only violated when the initial likelihood of being entrenched z and the likelihood of acting in a mixed strategy n are very high in conjunction. In any other circumstance, $EU_{Tb}(p>p^*)>EU_{Te}(p>p^*)$. Alternatively, the target's utility given the possibility of entrenchment is higher than the Bayesian rational outcome when $z<z_U$ where,

$$z_U = \frac{bU_T(D)(U_T(F)+p-pU_T(F)-1)+U_T(R)(U_T(D)-bU_T(D)+bpU_T(D))}{U_T(R)(b(1+U_T(R)-U_T(F)n(p-1))+U_T(D)-1)}.$$

The final issue we want to address here is the likelihood of successful surprise. In the standard Bayesian model, the target is surprised with probability $qY_e(1-n)$. The likelihood of being surprised in the entrenched model is $zqY_b(1-n)+(1-z)qY_b$. Despite the higher utility, the target with potentially entrenched beliefs suffers surprise more often when $z<z^S$ where,

$$z^S = \frac{Y_b + Y_e n - Y_e}{Y_b n}.$$

When the likelihood of entrenchment is high, then the target has a higher likelihood of taking action given the entrenched-aimed signal.

Discussion. This section discusses the primary conclusions, as well as the assumptions and causal processes underlying these results. The solutions and propositions are introduced conceptually, without the mathematical derivations or thresholds. The predicted effects of entrenched beliefs are summarized in Table 4.4. The table shows a set of conditions along the top, and along the side it shows the effects of bias relative to the Bayesian rational outcome.

Table 4.4: Predicted Effects of the Entrenched Beliefs on Outcomes

Impact of Bias on:	$p<p^*$	$p>p^*$ $z<z_U$	$p>p^*$ $z>z_U$	$p>p^*$ $z<z^S$	$p>p^*$ $z>z^S$
Belief Updating	Under-Suspicion	Over-Suspicion	Over-Suspicion	Over-Suspicion	Over-Suspicion
Posterior Beliefs	0/-[1]	+	+	+	+
Target Expected Utility	0	+	-	NA	NA
Pr(Attempted Surprise, *Y*)	0	-	-	-	-
Pr(Surprise)	0	NA	NA	-	+
Pr(Surprise|Target Bayes)	0	+	+	+	+
Pr(Surprise|Target Biased)	0	-	-	-	-

Note: The table shows the difference in expected outcomes when comparing the Bayesian rational model to the possibly entrenched model. Some outcomes are denoted NA since z_U could be larger or smaller than z^S.
1. Entrenched beliefs have no impact on the posterior beliefs in the restricted case due to endogenously determined strategies, but can result in under-suspicion for preferences are revisited below.

We begin by examining the equilibrium strategies, and the likelihood of surprise, when a sender is perceived to be non-threatening. Recall from the discussion above, that targets with entrenched beliefs are supposed to be more susceptible to surprise than those

without. It is important to be clear about this story at the outset. It does not simply assert that decision-makers will not act when they do not perceive a threat; instead, it focuses on the marginal propensity to take action. Targets with entrenched beliefs are assumed to ignore or trivialize an opponent's signals, particularly when preexisting beliefs suggest that the sender is non-threatening. Ignoring signals leads to inaccurate threat assessment, and this inaccurate assessment is hypothesized to increase the likelihood that the target is surprised. However, this proposition is not supported here. The model assumes that entrenched beliefs do impair learning, but the likelihood of surprise given a non-threatening sender is unaltered by entrenched beliefs. Under these conditions, the likelihood of surprise is determined by the strategic context, meaning that the conventional explanation overemphasizes the causal impact of psychological bias. This leads to the first proposition.

Proposition 1: When the Bayesian prior beliefs indicate that the sender is likely to be docile ($\Pr(S|z>0)=\Pr(S|z=0) \forall p<p^*$), a target with potentially entrenched beliefs is no more or less likely to be surprised than a Bayesian rational target.

The rationale behind this proposition actually stems from the competitive context, rather than the assumption that inaccurate beliefs cause bad decisions. When a sender is perceived as non-threatening, it is in the sender's interest to reinforce this belief whether or not they pose a threat. The sender pools on the ambiguous signal, where the aggressive type challenges with surprise and the docile type does not challenge. The aggressive sender never instigates a public crisis because surprise always works. Since aggressive

and docile senders appear to behave similarly, the sender's behavior carries no information value. The sender adopts a pooling signaling strategy that provides no additional clues about the threat level, because the target always has incentive to wait. Entrenched beliefs have no impact on decisions when the opponent's behavior provides no new information, since entrenchment involves ignoring or marginalizing new information. Therefore, decision-makers with and without entrenched beliefs will have the same threat perceptions, and are equally likely to be surprised when the opponent is initially perceived as non-threatening. It is also worth noting that entrenched beliefs have no impact on the sender's decision to adopt a pooling strategy in the first place. This also implies that deception matters most when the target suspects the sender of malign intent.

The nature of signaling in the model plays an important role in the derivation of Proposition 1. The model assumes that the sender's decisions to preserve the status quo and attempt surprise are equally ambiguous. The proposition is contingent upon the sender's decision to pool on ambiguous signals. To get a better understanding of information transmission, it is important to look at the sender's mixed strategy outcomes. When the sender chooses to mix signals, the target can use the ambiguous signals to refine the threat assessment. For example, an aggressive sender might choose a strategy where they attempt deception 50% of the time. If the sender is aggressive with probability 40%, and never attempts deception when docile, a Bayesian target believes they are facing an aggressive opponent 25% of the time they observe the ambiguous signal. The target was able to use the signal to revise their belief it was facing a hostile opponent from 40% to 25%. Therefore, it is inaccurate to argue that the ambiguous signal

never provides useful information. Ambiguous signals can be useful in revising beliefs and estimates.

The lack of new information in the signal that drives the result in Proposition 1 does not stem from the existence of an ambiguous signal, which might make the model appear as though there is no warning mechanism. Nor is it true that the ambiguous signal never imparts useful information for the target, as demonstrated above. The lack of information in the signal that drives Proposition 1 is an emergent property of the sender's optimal strategy. Anytime the sender chooses to pool on a single signal, the sender's behavior cannot be used to refine the target's preexisting beliefs. When the sender mixes their strategy, the ambiguous signal provides information that the target can use to revise their threat estimates. The aggressive sender type will publicly challenge with some probability when playing a mixed strategy. The target can use the probability of a public challenge to refine the likelihood that the sender is aggressive when observing the ambiguous signal. We will examine the impact of additional warning mechanisms below, however, the result above is a product of incentives and competition not simply limited signals.

The sender mixes signals, as illustrated above, when they pose a threat to the target. Technically, the sender adopts a mixed strategy when the likelihood of being aggressive is above a threshold value. In a standard Bayesian game, the sender adopts the mixed strategy that makes the target indifferent between acting and waiting. However, the possibility of entrenched beliefs complicates the sender's strategy selection. The sender can adopt two possible strategies; one aimed at making the Bayesian or efficient target type indifferent (Y_e), and the other aimed at making the entrenched target

type indifferent (Y_b). Neither strategy is ever capable of making both target types indifferent simultaneously.

Since a target with entrenched beliefs undervalues the new signal relative to previous information, the sender actually compensates for this difference. Signals satisfying entrenched targets should have a lower probability of attempted deception, and signals satisfying Bayesian targets are perceived as provocative by entrenched types. Here it is important to recall the discussion about signals in extensive form games. The sender's signal is not merely a single event or observation; it is representative of an entire distribution of signals. Since the entrenched type perceives the signal distribution that makes the Bayesian type indifferent as hostile, the sender should alter the distribution of attempted surprise to make the entrenched type indifferent. If an entrenched target ignores information suggesting that the sender is docile, the target will overestimate the likelihood that the sender is aggressive. The sender's best course of action is to compensate by attempting surprise less often. Analysis shows that the likelihood of attempted deception is strictly lower for the strategy aimed at the entrenched type ($Y_b<Y_e$) in mixed strategy space. The sender must attempt surprise less often than they would if facing a Bayesian target to make an entrenched target indifferent.

A simple numerical example helps to illustrate the impact of entrenched beliefs. In this example, we will call the sender Egypt and the target Israel in 1967. We begin with assumption that there is a high probability that Egypt will act aggressively towards Israel. An aggressive Egypt then decides whether to challenge publicly or attempt surprise to achieve its goal of reclaiming the Sinai Peninsula. A Bayesian rational Israel will be indifferent between acting and waiting given an ambiguous signal when Egypt

attempts surprise just often enough. If Egypt's optimal strategy that makes a rational Israel indifferent is to attempt surprise 50% of the time, then a rational Israel will mix between acting and waiting in response to that strategy. If Israel has entrenched beliefs, on the other hand, the Israelis will perceive Egypt's strategy of attempting surprise 50% of the time as hostile. Under these circumstances Israel will prefer to act whenever they get an ambiguous signal. Alternatively, Egypt can reduce the probability with which it attempts surprise in order to make Israel indifferent when they have entrenched beliefs. By attempting surprise 40% of the time, Egypt can ensure that an entrenched Israel is indifferent and plays a mixed strategy, while a rational Israel always prefers to wait. This illustrates the way that Egypt might alter its challenge strategy in response to the possibility of Israeli entrenched beliefs. Egypt needs to select the strategy that yields the highest payoff, which we address next.

The aggressive sender must choose between these entrenched-aimed (Y_b) and efficient-aimed (Y_e) strategies, and chooses the strategy that yields the highest utility. Despite less attempted surprise, the sender should always take the possibility of the target's entrenchment into account and the optimal strategy requires making the entrenched type indifferent. The reason that the biased-aimed strategy is always optimal for the sender lies in the target's response when they are not indifferent. The target is indifferent between strategies when the entrenched type observes the entrenched-aimed signal and when the efficient type observes the efficient-aimed signal. The target only selects a mixed strategy at these indifference points. When the entrenched target observes the efficient-aimed strategy, it is always optimal for the target to take action. This is what the sender wants to avoid. Conversely, the efficient target always waits when they

observe the entrenched-aimed strategy. If the sender uses the efficient-aimed strategy, they attempt deception more often but always trigger action from the entrenched type. The sender is better off attempting deception less often, and makes up for it by occasionally encountering the efficient target type that always waits. This leads to the second proposition.

Propositions 2: In mixed strategy space, senders should attempt less surprise if they know that there is a possibility that the target has an entrenched belief that the sender is aggressive ($\Pr(C\sim M|z>0)<\Pr(C\sim M|z=0) \forall p>p^*$).

Finally, we examine the impact of entrenched beliefs on the target's utility and find that a positive strategic bias may exist under certain conditions. Less attempted deception should be beneficial to the target, and it is in most cases. However, false alarms can negate this benefit. The target's mixed strategy that makes the sender indifferent does not change, despite less deception. When the target almost always has entrenched beliefs, and plays their mixed strategy in response to the sender's strategy, less deception translates into more false alarms. The target acts too often when the docile sender is attempting to preserve the status quo. Therefore, the target can only achieve increased utility from entrenched beliefs when the expected cost of false alarms is sufficiently low. It is also worth noting that false alarms end the competition in this model, which does not take repeated interaction into account. It is possible that false alarms can increase the tension in a relationship over repeated interaction, what is often called an escalatory spiral (Jervis 1976). The spiral story might argue that the target's false alarm increases

the sender's anxiety about target intentions. A sender that starts out as docile might be frightened into a more aggressive stance after observing repeated false alarms, making the relationship unstable and dangerous.

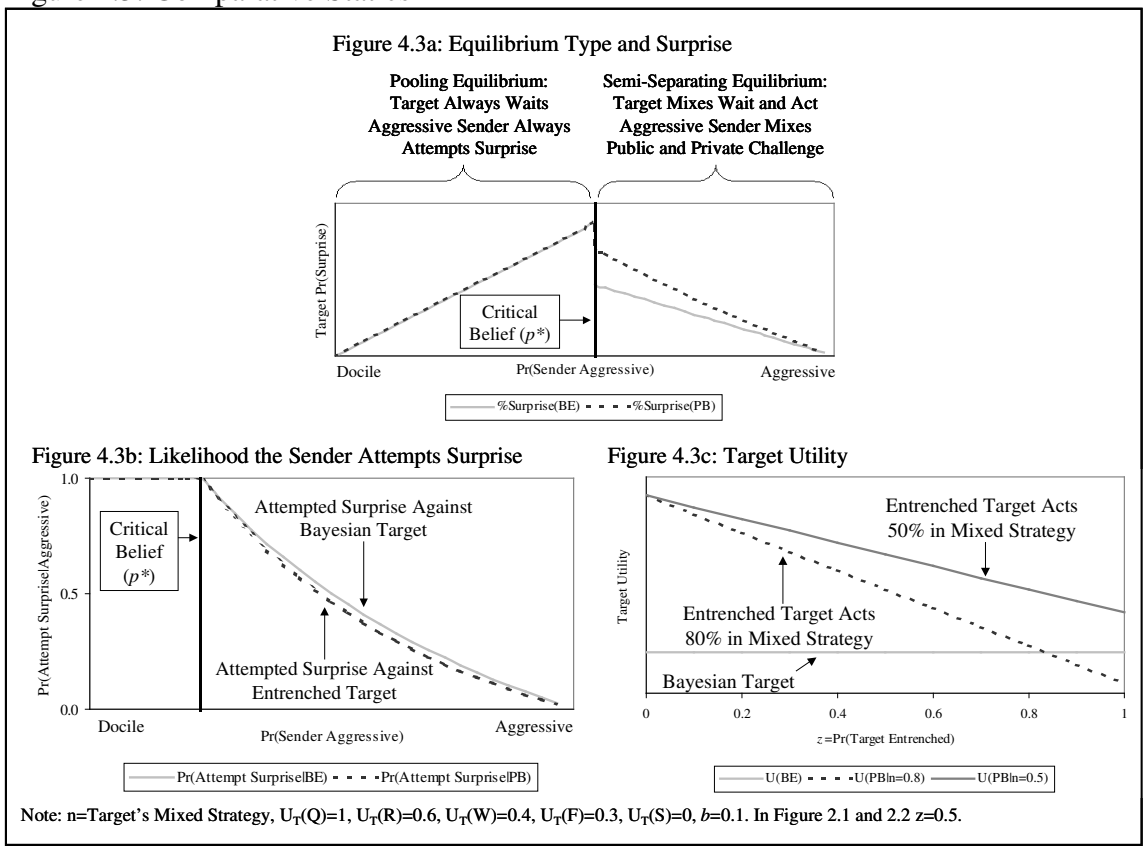
It is also important to recognize that entrenched beliefs are more likely to be neutral or beneficial for the target only when entrenchment is probabilistic ($0<z<1$). If a target always has entrenched beliefs, they run the risk of suffering too many false alarms even when the sender adopts the entrenched-aimed strategy. Together, these insights yield Proposition 3.

Proposition 3: Potentially entrenched beliefs can increase the target's utility as long as the expected costs of false alarms are not too high, and entrenchment remains probabilistic.

Before revisiting some of the preference assumptions, it is worth pointing out a corollary to Proposition 3. The possibility of entrenched beliefs does involve one serious drawback; the likelihood of being surprised may be higher relative to Bayesian equilibrium outcomes in mixed strategy space. Recall that entrenched types mix strategy when the entrenched-aimed signal is sent, but the efficient type always waits. Therefore, efficient types will always be surprised when the sender attempts surprise. This reaction increases the target's likelihood of surprise, but not enough to offset the benefit to its utility. The entrenched type is less likely to be surprised, but suffers too many false alarms. The Bayesian type derives a higher expected utility, but suffers surprise more often. This is displayed graphically in the three panes of Figure 4.3. Figure 4.3a shows the strategy selection contingent upon the likelihood the sender is aggressive, and the

differential in the likelihood of surprise in the semi-separating equilibrium. Figure 4.3b shows the change in the sender's strategy given the possibility of target bias, specifically the reduced likelihood of attempted deception. Finally, Figure 4.3c shows the target's net benefit from potential bias, and the conditions where this benefit is erased.

Figure 4.3: Comparative Statics



Note: Figure 4.3a shows that the pooling and semi-separating equilibria are a condition of the prior belief p and the critical belief p^*. It also shows that bias has no impact on the probability that targets are surprised when $p<p^*$ (under-suspicion), and the probability the target is surprised increases when targets are potentially entrenched and $p>p^*$ (over-suspicion). Figure 4.3b shows that the aggressive sender's strategy is unchanged when $p<p^*$, but the aggressive sender is less likely to attempt surprise in equilibrium when $p>p^*$. Figure 4.3a showed that targets were more likely to suffer surprise when $p>p^*$ and Figure 4.3b showed that senders were less likely to attempt surprise when $p>p^*$. Figure 4.3c shows that targets have a higher expected utility when potentially entrenched than they would if Bayesian rational as long as the probability the target is entrenched and likelihood of acting in the mixed strategy are not too high.

Revisiting Preferences

The equilibrium deduced above involved a few assumptions about the sender's behavior. First, the analysis assumed that the aggressive sender always chooses to challenge the status quo, X=1, where X represented the probability that the aggressive sender challenged in the first place. This is true whenever the sender prefers the sure outcome of public standoff to the status quo, $U_S(W)>U_S(Q)$. When this preference assumption is violated and the status quo is preferred to a public standoff, the sender never challenges publicly. The aggressive sender still mixes strategy, but they mix between challenging with attempted surprise and preserving the status quo. In this case, the possibility of attempted surprise given a challenge (Y) is one, and the sender mixes between challenging and preserving the status quo, X. The results for this mixed strategy mirror those above. The possibility of bias reduces the likelihood of attempted surprise, and the aggressive sender is more likely to maintain the status quo. The target's utility given the possibility of bias remains higher than the standard Bayesian equilibrium as long as the target does not suffer too many false alarms. The target still faces a higher likelihood of experiencing surprise than they would in a standard Bayesian model. Here, the aggressive sender opts to challenge less often, which is beneficial to the target. The target, however, still waits when the efficient type is selected, suffers surprise more often as a result, but may be better off anyway. The impact of bias on outcomes is unchanged, and the only change is to the aggressive sender's decision to mix between preserving the status quo and challenging with surprise.

We also assumed that the docile sender always chooses to preserve the status quo, x=0, where x represents the probability that the docile sender challenges the status quo. A docile sender has little incentive to attempt a public challenge since the target always

acts, however, they may want to attempt surprise if they believe it will be successful. The sender, however, is not doing themselves a favor by adopting a strategy where $x>0$ and $p>p^*$. If docile types attempt surprise sometimes in the semi-separating equilibrium, aggressive types have to reduce the likelihood of attempting surprise to maintain the target's belief $p=p^*$. The target, in this case, is no longer concerned with likelihood that the sender is aggressive given the ambiguous signal, but instead interested in the conditional probability that the sender is attempting surprise given the ambiguous signal. This conditional probability must be equal to the critical belief if the target is to remain indifferent. This means that the likelihood of attempted surprise given an ambiguous signal remains unchanged in the semi-separating equilibrium when $p>p^*$, and the docile sender never attempts surprise as long as $U_S(S)>U_S(s)$, where $U_S(S)$ is the aggressive sender's value for surprise and $U_S(s)$ is the docile sender's value for surprise.

The revision of the pooling equilibrium, which was expected when $p<p^*$, is slightly more complicated. In the pooling equilibrium the aggressive sender always carries out successful surprises since the target's posterior belief is equal to the prior belief and below the critical belief value. In this case, the docile sender can attempt surprise with some probability ensuring that the target's posterior belief equals the critical value. The target, when indifferent, can respond with a strategy that makes the docile sender indifferent between attempting surprise and preserving the status quo. If the target ever takes defensive action in this mixed strategy, which occurs when $U_S(s)>U_S(Q)$, the sender is better off playing the initial pooling strategy where the docile type always waits. The sender can attempt surprise more often, but they pay a cost for doing so. The target takes action with some probability, thereby thwarting some of the surprises when the

sender is aggressive and docile, and the sender is better off avoiding this outcome. There is one refinement to this result, which is important when entrenched beliefs are added in. The docile sender can attempt surprise rarely, just often enough such that the target's posterior belief remains below the critical value. This allows the sender to optimize attempted surprise while ensuring that the target waits. Since this is a razor edge solution whereby the sender adopts a strategy $x^*=x_{indifferent}-\varepsilon$, ε being some very small number, we only look at the substantive effects given the possibility of entrenched beliefs.

The effects of entrenched beliefs in this instance run counter to some of those deduced in the restricted model above. When $p<p^*$, entrenched beliefs did not play a role since there was no new information available. When $p>p^*$, the target anchored on the belief that the sender was threatening. In order to make a biased target indifferent, the sender adopted a mixes strategy where attempted surprise, Y_b, given a biased type was less than it would be for the Bayesian type Y_e ($Y_b<Y_e$). This effect reverses when a docile sender will attempt surprise and $p<p^*$. In this instance, the target anchors on the belief that the sender is non-threatening, so the (docile) sender can attempt more surprise while ensuring the biased type remains indifferent ($x_b>x_e$). There is a catch. The sender cannot extract any benefits from the target's biased belief that the sender is non-threatening. If the target plays the strategy with more surprise, x_b, any time the target is actually Bayesian they will always act and thwart the surprise since, $x_b>x_e=x_{(Bayes\ indifferent)}-\varepsilon$. Just as the sender was better off adopting the strategy with less surprise in the restricted case above (for fear of provoking the biased type), the sender finds themselves in a similar situation (for fear of provoking the Bayesian type). With the unrestricted preference, the target might anchor on the belief that the sender is non-threatening, but the sender is still

best off adopting the strategy with less deception that keeps the Bayesian indifferent ($x^*=x_{(Bayes\ indifferent)}-\varepsilon<x_b$).

This raises one final question about the target assumptions, specifically the target's willingness to act against a public threat. When the target is willing to acquiesce, $U_T(E)>U_T(W)$, then the sender always challenges publicly with the knowledge that the target will not act. Under these conditions, the docile sender may posture as aggressive whenever $U_S(e)>U_S(Q)$. The signals become irrelevant and bias plays no role in altering expected payoffs when the target will not act against a public challenge.

Conclusions

This chapter carefully walked through a series of models that were designed to examine the impact of irrationally entrenched beliefs in competitive and uncertain contexts. Beginning with a basic strategic model of surprise with imperfect information, we showed that surprise could occur in the absence of psychological bias or organizational constraints. Whenever the initiating actor had an incentive to send ambiguous signals, successful surprise was possible. In short, surprise is a perfectly reasonable, even expected, outcome of a simple model with perfectly rational actors. This framework suggests that many instances of policy failure and surprise that are blamed on psychological or organizational mistakes may overlook basic rational explanations. Simple information constraints and differing interests are necessary and sufficient to cause surprise. This is not to say that psychological and bureaucratic studies of policy failure are not useful. Often, these studies help to identify ways to improve information collection, analysis, organizational structure and decision-making processes. It is

important, however, to recognize that perfectly rational actors operating in a world much simpler than that confronting analysts and policymakers still suffer surprise and policy failures without additional complicating psychological or organizational issues.

More complicated models showed that the causal effects of entrenched beliefs seem to be misunderstood. Conventionally, entrenched beliefs and other psychological biases impede accurate assessment thereby leaving actors with biases worse off. Instead, we argued that there are many situations where entrenched beliefs are irrelevant to outcomes and other cases where they can actually be beneficial. When actors begin with the belief that their opponent is non-threatening, there is nothing the initiating state can do to evoke a response even when players are perfectly rational. There is simply no strategic justification for action, and entrenched beliefs that result in irrational under-suspicion do not change that result. Entrenched beliefs are supposed to make these actors more susceptible to surprise, but they are already susceptible so entrenched beliefs have no marginal effect.

Alternatively, states facing opponents perceived to be threatening could actually benefit from entrenched beliefs as long as an initiating state recognizes the possibility that the target may be entrenched. The initiating state has a greater chance of pulling off surprise when adopting more conservative strategies, specifically strategies that involve attempting surprise less often. The initiating state realizes they are perceived to be threatening, and knows that the target is predisposed to maintaining this belief. So, the only way an initiating state can change this belief is by being less provocative. The deductive results showed that potentially entrenched over-suspicious targets were more likely to experience surprise when the sender was perceived to be hostile, but this loss is

often offset by the benefits of less attempted deception. It is possible, and even probable, that actors benefit from entrenched beliefs resulting in over-suspicion. It is also interesting to note that rational actors are more likely to be surprised when initiating states adopt these more conservative strategies. Perfectly rational actors are the first to be fooled, whereby entrenched actors see through the attempted deception.

The chapter ended by looking at alternative preference assumptions showing that the primary deductive conclusions were reasonably robust. The plausible alternate preference assumptions did not change the results pertaining to entrenched beliefs in any meaningful way as the primary conclusions discussed above held against the different orderings.

Chapter 5
Empirical Analysis Part 1:
Quantitatively Assessing Entrenched Beliefs

The previous chapter offered a number of mathematically derived propositions about the relationship between entrenched beliefs and strategic outcomes that challenge conventional logic about psychological biases and decision-making. Since most studies have emphasized the way that biases negatively impact assessment, it is assumed that decisions based upon these inaccurate estimates produce suboptimal outcomes (Jervis 1976). While this is true in decision-theoretic contexts focusing on the quality of an individual decision, these arguments do not necessarily hold up in competitive environments. The deductive logic in the previous chapter provided conditions whereby the conventional story should break down and entrenched beliefs can actually be beneficial. These assertions about the potential optimality of psychological biases deviate sufficiently from the usual causal relationship between biases and outcomes, that empirical testing of these propositions becomes crucially important. The next two chapters are devoted to empirical tests of these counterintuitive ideas.

This chapter, the first of the two empirical sections, uses a large dataset and statistical methodology to seek support of the theoretical propositions. Not all of the propositions lend themselves to statistical tests with the available data, and tests of the other propositions require an innovative approach to variable coding. When this constraint is coupled with the fact that the observation of entrenched beliefs like irrational

under- or over-suspicion in any systemic fashion is difficult, a concept covered in chapter 3, more detailed study of individual cases helps to overcome these issues. The next chapter will address two case studies in greater detail by walking through the causal mechanisms of interest and using more detailed history to trace the impact of these mechanisms. Before getting to the cases, we offer a statistical examination of one proposition in this chapter.

The prominence of statistical inference in political science is a cause of concern for those who advocate psychological explanations (Tetlock 1999). As we will discuss below at greater length, problems like this are particularly difficult to address with statistical studies. First, entrenched beliefs are not necessarily observable, and to date, none of the mainstream datasets common in international politics research includes variables relevant to psychology and information processing. This is also coupled with the fact that most datasets do not easily lend themselves to the study of strategic surprise, which was the substantive area in the deductive model and one where entrenched beliefs are supposed to have a significant causal effect. Here, we attempt to overcome some of these constraints and use a readily available dataset of international crises from 1918-2002, International Crisis Behavior (ICB). The statistical analysis here is not able to provide confirmation of the propositions deduced in chapter 4, but its does illustrate some of the core theoretical deductions.

The quantitative empirical section begins by discussing the propositions from chapter 4. This focuses on the operational specification of the propositions and the testability of each given the available data. The next section discusses the operationalizing of the independent variable used to represent entrenched beliefs and the

underlying logic for the proxy. The third section addresses the rest of the variables used in the analyses. Then we present the results from the empirical test and discuss the implications.

Quantitatively Testable Propositions

In this section, we quantitatively evaluate Propositions 2 and 3 from chapter 4. Testing both Propositions 1 and 3 is particularly difficult given the data available. Proposition 1 stated that targets with entrenched beliefs were no more likely to be surprised than targets without entrenched beliefs when the potential opponent was perceived as non-threatening. Entrenched beliefs, in this instance, results in irrational under-suspicion where a target views the sender as less threatening then they are. In this deductive statement, the likelihood of being surprised is the dependent variable, entrenched beliefs the independent variable, and the relationship is subject to the condition that there is a high likelihood that the sender is docile. Unfortunately, there are no datasets that capture the target's level of surprise when a crisis is initiated. Even if the data can be manipulated to capture a sender's attempted surprise, which we discuss below, there is no way to clearly distinguish between cases where the target expected the surprise or not. Barton Whaley's (1969) unpublished study of surprise attack does try to identify instances where one side surprised the other. While the cases show that attempted surprise attacks almost always successfully catch the target off guard, Whaley notes that the data is not suited for statistical analysis since there is no way to clearly identify aborted surprise attacks. The most visible instances of surprise are successful attacks, and the inability to observe aborted attacks means that the dataset naturally

selects on the dependant variable. While Whaley finds that surprise is often successful in his dataset, simply assuming that the target was surprised whenever the sender attempts surprise eliminates the key distinction between attempted surprise, aborted surprise, successful surprise and thwarted surprise. For this reason, examination of Proposition 1 is not suitable for statistical testing given the available data.

The empirical analysis of Proposition 2 is the only quantitative test in the project. Proposition 2 stated that senders were less likely to attempt surprise against a target with irrational over-suspicion when the sender was perceived to be threatening. The independent variable, the sender's method of triggering a crisis, is observable. The ICB dataset distinguishes between crises triggered with military action and those triggered in other ways. Triggering a crisis by military action may be an imperfect characterization of surprise, but it is adequate to test the proposition. The dependent variable of interest is the possibility that the target has entrenched beliefs. Entrenched beliefs are not directly observable, and as noted above, not included in standard international politics datasets. This problem is more difficult to solve, but not insurmountable. The next section introduces a potential proxy variable, or rather combination of variables, that helps distinguish between states that are more likely to have entrenched beliefs from those that are less likely to have such bias.

Finally, the chapter does not test Proposition 3, because any test would be tenuous at best and possibly unreliable. Proposition 3 argued that targets can actually benefit from the possibility of entrenched beliefs, specifically irrational over-suspicion. That is to say, entrenched beliefs can be welfare enhancing relative to perfect rationality. The target's utility is the dependent variable of interest, and the possibility of entrenched beliefs is the

independent variable. Unfortunately, utility in this context does not directly equate with crisis outcomes. Recall that the benefits from irrational over-suspicion stemmed from the sender's decision to attempt surprise less often (Proposition 2) since the target's irrational suspicion reduced the likelihood of successful surprise. The benefits or utility encapsulated in the model addresses the target's ability to deter surprise, thwart surprise and maintain the status quo. The model did not address the final crisis outcome. For example, a target may successfully thwart a surprise attack, then subsequently suffer setbacks that lead to defeat. Alternatively, a target may reverse the initial losses suffered in a successful surprise and end up victorious. Since the crisis outcome does not equate with welfare increases as defined by the model, crisis outcome is not an appropriate measure. Unfortunately, there is no variable that directly maps to welfare increases as specified in the model.

Operationalizing Entrenched Beliefs for Statistical Analysis

Entrenched beliefs, the tendency to irrationally anchor on prior beliefs, are not directly observable like war casualties or victories. Social scientists, however, have made efforts to create measures to quantify unobservable variables like intelligence quotient and international norms (Gelpi 2002). Since entrenched beliefs are not directly observable, we have to identify observables that indicate the presence of, or are associated with, entrenched beliefs. This section discusses the way that the notion of protracted rivalry can actually be used as a variable representative of entrenched beliefs. Protracted rivalries and entrenched beliefs, however, are not directly comparable and it is

important to think through each concept in order to devise a proxy variable that captures the causal mechanism of interest.

Enduring or protracted rivalries, covered briefly in chapter 2, refers to dyads or sets of states that are locked into a pattern of poor relations. Protracted rivalries often stem from disputes over territory, minority rights, societal values or any number of other issues. These rivalries have drawn significant attention, because states trapped in these long running disputes tend to be more conflict prone than other sets of states (Goertz and Diehl 1993). The early work on this topic identified a number of factors that could contribute to protracted rivalry and the related violence. Protracted rivalries could be the product of strategic, economic, cultural or social issues. One explanation for the violent and protracted nature of these rivalries is partisan bias that develops on each side. This explanation, which is particularly useful for this analysis, was cast aside in the next iteration of protracted rivalry literature. This second iteration of work argued that protracted rivalries could be reduced to territorial conflicts (Vasquez 1993, Thies 2001). While recognizing the importance of issues like national identity, this work showed that the statistical relationship between protracted rivalries and violent conflict was essentially identical to the relationship between territorial disputes and violent conflict. This finding led Vasquez (1993) to argue that scholars should focus on the more concrete territorial disputes than the more abstract concept of protracted rivalry.

More recent work on rivalry and conflict resolution has eschewed the notion that protracted rivalries are simply territorial conflicts (Bar-Tal 2000; Thies 2001). Instead, research has focused on the psychological and social aspects associated with protracted rivalry. Thies (2001) argues that social psychology, and particularly socialization and

competition, play a key role in protracted rivalries. Perceptions of oneself and one's rival might not be the same as the rival's perceptions since these beliefs are often the product of social processes. Since conceptions of national identity help to dictate the nature or terms of interstate relations, firmly held conceptions of national identity may imbue policymakers with inaccurate beliefs about the intentions or capabilities of another state. In this context, beliefs about the opposing state's interests or intentions may be entrenched because of the entrenched notions of national identity. It has also been argued that entrenched beliefs play a significant role in the formation and continuation of rivalries (Mor 2004). Conflicts and threatening events early on in interstate relations can interact with national security conceptions facilitating rapid lock-in into protracted rivalry (Mor 2004). Hassner (2007) examines the relationship between time and protracted rivalries, arguing that these rivalries usually grow more acute with time. Tangible aspects such as investment in territory or the discrimination of minority groups, as well as intangibles aspects such as threat perceptions or national pride, grow more accurate with time. Once entangled in a protracted rivalry, beliefs and conceptions become increasingly entrenched.

If these psychological explanations have some validity, then decision-makers involved in protracted rivalries should exhibit higher levels of irrational over-suspicion when dealing with their rival.[10] Protracted rivals are usually perceived as particularly aggressive or expansionist, and part of the difficulty associated with ending these rivalries involves overcoming entrenched beliefs (Bar-Tal 2000). If this is true, then some existing datasets might prove useful in testing carefully specified psychological

[10] Mutual entrenched beliefs will be dealt with explicitly in future projects. Note that one state with entrenched beliefs may help to deter the aggressive urges of their rival (Proposition 2), so it is possible that mutual entrenchment can actually have a stabilizing impact on the interstate relationship.

explanations. In this case, Proposition 2 specifically contrasts the expected outcomes when targets have irrational over-suspicion with cases absent entrenched beliefs. If states in a protracted conflict believe that their opponent is aggressive, and fixate on that belief, then protracted rivalry may act as a suitable alternate variable. There are, however, some additional issues that need to be addressed before substituting protracted rivalry for entrenched beliefs.

First, the analysis that follows is predicated on the auxiliary assumption that states in protracted rivalries are more likely to exhibit higher levels of entrenchment than they otherwise would. It is worth pointing out that the use of protracted rivalry for entrenched beliefs is not based upon causation but correlation. In other words, it does not matter whether the protracted rivalry caused the entrenched beliefs, or whether the entrenched beliefs led to rivalry. The psychological explanations on enduring rivalries posited different relationships between entrenched beliefs and protracted rivalry. Thies (2001) and Mor (2004), for example, argued that entrenched beliefs play a central role in creating a protracted rivalry. In addition, Bar-Tal (2000) and Hassner (2007) argue that entrenched beliefs cultivate existing rivalries and make them difficult to terminate. There is no clear agreement on the causal relationship. Entrenched beliefs may cause protracted rivalries, or protracted rivalries may promote entrenched beliefs. The empirical test here does not focus on the relationship between entrenched beliefs and protracted rivalry, but the relationship between entrenched beliefs and crisis behavior. There is, therefore, no need to make causal assumptions about entrenched beliefs and protracted rivalry. All that matters for this analysis is that protracted rivalries are positively correlated with the presence of entrenched beliefs.

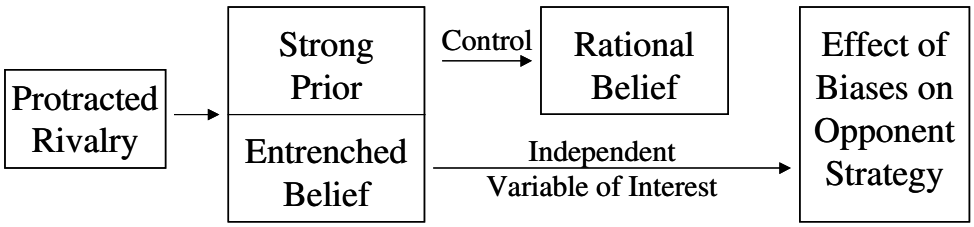
Second, it is important to take a closer look at rivalries, beliefs, and information before proceeding. We argued above that states in protracted rivalries are more likely to be irrationally over-suspicious, but they are also very likely to have been involved in prior crises with their rival. These crises provide first hand experiences that the state can learn from, which is important since prior work on state learning suggests that states are much more likely to learn from their own experience than the experience of others (Reiter 1996). Rivals, therefore, may use the information garnered in prior crises that would be unused by states outside of such rivalries. When actors have more crisis experience or information about an opponent, their prior belief is stronger, and that leads to perfectly rational information discounting. Bayesian rational actors should apply greater weight to information derived from the third crisis than the seventh crisis.[11] In other words, states involved in protracted rivalries should update beliefs modestly, even when completely rational. Protracted rivalry, therefore, captures both rationally strong priors and entrenched beliefs. In order to use the protracted rivalry as a proxy variable, it is necessary that we separate rational beliefs from the effect of entrenched beliefs.

The most straightforward method of separating rational from biased beliefs is by controlling for rational prior beliefs. If protracted rivalry captures both rational and irrational beliefs, and we include a variable that captures the rational beliefs, then protracted rivalry becomes proxy variable for entrenched beliefs. Any explanatory effect of protracted rivalry, when controlling for rational prior beliefs, is capturing variance unexplained by the rational prior. This does not mean that protracted rivalry is only capturing irrational over-suspicion. It may capture other aspects of decision-making such as shifting notions of value or national identity. By controlling for rational prior beliefs,

[11] For a good review of Bayesian updating over individual observations see Gerber and Green (1999).

however, any effect of protracted rivalry explains some unobservable quality associated with the rivalry that is not addressed by the control variables. This process is displayed in Figure 5.1.

Figure 5.1: Protracted Rivalry, Strong Priors and Entrenched Beliefs



Summary: By adding in the Bayesian rational belief about violence in prior crises, any regression results are capturing forces that go beyond the protracted rivals experience and captures irrational beliefs about their opponent's aggression.

Since there is no rational prior belief variable in the ICB data, controlling for rational beliefs requires construction of a suitable variable. The variable needs to represent the actors' beliefs that they are facing an aggressive opponent, and a rational prior belief should reflect the information available from previous crises. Since we are interested in the actors' beliefs about opponent intentions, an appropriate metric is the historical likelihood of violence between the two states. A rational belief variable should reflect the probability that previous crises turned violent. As the proportion of violent crises increases, a rational actor should believe that likelihood of facing an aggressive opponent increases. Violent crises are defined as those that are either initiated violently, what we call a "hostile trigger" or "surprise," or triggered diplomatically and involve violence at some later point. This provides a baseline belief for the actors prior to the new crisis.

The "rational prior" variable developed here to control for rational beliefs is derived using Bayesian logic since the order of events, international crises, matters. The variable itself represents a simple proportion that captures the frequency of violent and nonviolent crises. For any observation within the dataset, the control variable is simply a snapshot of past events. In aggregate, the order of events does not matter in Bayes' Rule and it is tempting to think about the variable as simple proposition rather than a Bayesian learning process. If we are interested in the belief change over time, however, order does matter in Bayesian processes. If two states are involved in four crises and two are violent, the likelihood of violence is 50%, but after two crises, it does matter whether the first two crises or the last two were violent. Order only matters since the series of interest are cropped into time periods, whereas it would be irrelevant in a typical distribution with priors. Since each crisis in the dataset is a single observation that gets aggregated over time, the order does matter. So, the relevant measure at each observation is based on frequency, but that frequency is determined by the updating over individual successive crises making the Bayesian learning process contingent on the sequence over time.

The rational prior variable is computed by taking subsets of prior cases where the sender and the target are the same across the subset. Then the proportion of prior violent cases is calculated during the next crisis observation in the dataset with the relevant actors. The rational adjustment variable is technically calculated as follows,

$$\Pr(Aggression) = \frac{Number\ of\ \ Past\ Violent\ Crises}{Number\ of\ \ Past\ Total\ Crises}.$$

This assumes the actors know about their previous crisis history and use this information when making decisions. The formulation includes two important aspects of a Bayesian belief: 1) a rational belief about dyadic hostility using prior experience, and 2)

rational discounting of new information based on the strength of the prior. An example helps to illustrate the calculation. At the initiation of a fourth crisis, if two of the three previous crises turned violent, the states estimate the probability of violence based upon prior information is 67%. If the fourth crisis is not violent, at the outset of the fifth crisis, the states know that 50% of prior crises turned violent. As discussed above, the information from the fourth crisis will have less effect than the third crisis on how much an actor's beliefs change, and so on. The actors use all previous information on the likelihood of violence when making decisions, but it is weighted based upon the total sum of past experience.

Unfortunately, the coding of rational prior variable is slightly more complicated than the initial calculation might suggest. The major coding difficulty rests with the treatment of the initial instances of crises between any two states. If we could observe prior beliefs about aggression, and the strength of these beliefs, these values could be added to the data. Unfortunately there is no good way to estimate these values. To compensate for this difficulty, we introduce three different measures of rational prior beliefs. The difficulties, and the different ways of dealing with them, will become clearer as we walk through the coding process.

In the initial coding of the rational prior variable, the first crisis between each set of states was coded as such, the first crisis. Each state started with the belief that the history of prior violence between the two states is zero. Both the numerator and denominator are undefined until the first case. After the first crisis, the total number of crises increases to one. If the crisis is resolved without violence, the numerator remains zero, and the rational belief about prior dyadic aggression is zero. Alternatively, if a crisis

that starts or turns violent increases the numerator to one, and the posterior belief about the likelihood of violence (the prior for the next crisis) is one or 100%.[12] There are a number of obvious problems with this coding scheme. First, prior beliefs about the likelihood of violence the first time any two states have a crisis are undefined. This is an unrealistic assumption, but we will see below that it is not as damaging as it seems. The second problem with the coding is that the variable is subject to large swings in initial instances of crisis. States beliefs about the likelihood of violence can easily swing from zero to one after a single crisis. The swing in beliefs is equally difficult to justify. While these problems do smooth themselves over repeated crises, they are significant enough to warrant alternate specifications of the belief variable.

There are two additional specifications used to test the theory, which will be creatively labeled rational prior adjustment 1 and rational prior adjustment 2. Rational prior adjustment 1 addresses the starting belief. Above, we suggested that zero might not be an appropriate initial or starting belief. This can be fixed by choosing an alternate starting value for the belief about violence, but there should some justification for the selected value. Analysis of the ICB data used here shows that the probability a crisis historically turns violent is around 40%. Of the crises included, 40% have military action at some point in the crisis and 60% do not. This would appear to serve as a reasonable baseline for prior beliefs. We assume, therefore, that all sets of states start with the belief that the likelihood of violence in any crisis is 40% (0.4/1.0).[13] This helps to avoid the

[12] Technically, one could argue that this initial coding method is not Bayesian since there are no prior beliefs until the actors experience a crisis. Alternate coding of the variable with different strength prior beliefs is discussed below.

[13] Any subsequent crisis increases the denominator by one and the numerator by zero or one contingent on violence. The formula, therefore, is (0.4+number of violent crises)/(1.0+total past crises).

unrealistic assumption that the initial belief is zero, and also has the effect of minimizing swings in belief.

Rational prior adjustment 1, however, includes the implicit assumption that the communal history is weighed the same as one's own experience in a single crisis. Technically, it assumed that past history was weighted as single instance or observation of crisis. The numerator was 0.4 and the denominator was 1.0 (reflecting a single case). Work on beliefs and learning in international politics suggests that states are much more likely to learn from their own experience than that of others (Reiter 1996). Deterrence research has reported similar findings about direct experience (Huth 1998). The implicit assumption, that communal experience is valued as much as direct experience, does not appear empirically valid. The third specification of rational prior beliefs tries to account for the difference between communal and direct experience. Rational prior adjustment 2 weighs the communal belief less than direct experience. The prior communal belief is weighted as a half of a case (mathematically equivalent to 0.2/0.5).[14] The initial belief still reflects a 40% likelihood of violence in the crisis, but subsequent personal experience has greater leverage over posterior beliefs after each crisis.

There are two additional rational adjustments calculated that represent relatively strong reliance on priors rather than recent experience. In rational prior adjustment 3, the initial prior is equivalent to five crises, and remains set at the dataset average 40% (2.0/5.0). Finally, rational adjustment 4 considers an even stronger prior whereby the prior is based upon 10 crises (4.0/10.0), reflecting the possibility that there is more information needed to alter this belief.

[14] This formula is (0.2+number of violent crises)/(0.5+total past crises).

None of these specifications is perfect, but use of all three counters the weaknesses of each. If the empirical results are similar across each of the three specifications, then the analysis is robust to the variable specification problems. Drastically different results, changes in coefficient signs or significance, would imply that the variable is not a robust measure of rational beliefs. The results section will show that rational prior beliefs play an important explanatory role, and equally as important, the results are robust across specifications.

In essence, the regressions assume that the actors know the prior history reflected by the rational prior variable. Any variance explained by the rational prior belief reflects knowledge that the players have about their past history. Under these conditions, protracted rivalry captures any additional variance. The sender, for example, knows the target's rational prior belief about the likelihood of violence and can adjust their strategy accordingly. Proposition 2 predicted that initiators would be less likely to use hostile crisis triggers when there was a possibility the target had entrenched beliefs that made them irrationally suspicious. According to method of operationalizing entrenched beliefs used here, this means that the predicted effect of protracted rivalry on hostile crisis triggers should be negative. Initiators should be less likely to use hostile triggers against a protracted rival, controlling for the past knowledge or experience of the two states. It seems perfectly rational to assume that initiators would be more likely to initiate crises violently as the likelihood of violence in previous crises increases, meaning there would be a positive correlation between violent crisis triggers and rational beliefs about prior violence, but this is not central to the argument.

Regression Variables

The previous section introduced the method used to specify the primary independent variable of interest, entrenched beliefs. This section discusses the other variables used in the two statistical models presented below. We begin by discussing the two dependent variables and then address the independent variables. A summary of the variables and coding are found in Table 5.1.

Proposition 2 specifically addressed the sender's propensity to use surprise. The ICB data does not specifically code the use of surprise, as discussed above, but it does code the crisis trigger. A crisis can be triggered in any number ways ranging from public statement to large-scale attack. If a surprise is akin to a large-scale military action, specifically a violent military action, then the ICB data is useful for testing the proposition. ICB codes crisis triggers on a scale from 1-9. Lower numbers represent less aggressive trigger mechanisms, and higher numbers are associated with military action and violence. This creates a natural split whereby military activity is similar to the notion of surprise used in the model. Crises triggered violently or by large-scale military actions are coded as "crisis begun by military action." Such crises are considered surprises, and less aggressive crisis triggers are coded as public challenges. Barton Whaley's (1969) finding that states are almost always surprised by major military action at the outset of a crisis, suggests that the crisis trigger variable relying on military action may well be representative of surprise. Militarily triggered crises are coded one, and non-militarily triggered crises are coded zero. Not all crises involve violence or military action. A number of crises are initiated non-violently and resolved without either party resorting to military action. Here, we are primarily concerned with the trigger, or the signal that

initiates the crisis. Since the underlying logic of Proposition 2 was that senders were more likely to trigger crises using public challenges when targets had entrenched beliefs, this variable captures the relationship of interest, the likelihood of triggering a crisis by surprise or military action.

Table 5.1: Summary of Variables

Variable	Description
TRIGGER	Sender's decision to initiate the crisis with military action 0=Diplomatic 1=Military action
POWDIF	The power differential between the target and sender (scaled discretely -3- to 3) -3=Large power advantage for sender 3=Large power advantage for target
GRAVITY	The gravity of threat in the crisis (scaled discretely 1 to 6) 0=Economic threat 1=Limited military threat 2=Political threat 3=Territorial threat 4=Threat to influence 5=Threat of grave damage 6=Threat to existence
RATPRIOR	The probability that historical dyadic crises turned violent (scaled continuously 0 to 1) 0=No prior crises turned violent 1=All prior crises turned violent Note that there 4 adjustments that alter the strength of the prior belief while maintaining a frequency of 40%
PROTRACTED	Whether the states involved in the crisis were protracted rivals 0=Not protracted rivals 1=Protracted rivals

There are also a couple of additional independent variables incorporated into the analysis to control for conventional explanations about the use of force. First, we include a control variable for the power differential. The imperfect information model presented suggested that power status could play a role in surprise decisions. It was deduced that states with higher values for the status quo and lower costs associated with suffering surprise made more attractive surprise targets. We logically argued that large or great powers were more likely to be in that position. However, it is important to recognize assertions about power are relative one's opponent. A great power attacked by another

great power may suffer high costs from surprise. In this case, the target will be more likely to act, reducing the likelihood that any attempted surprise is successful and reducing the incentive for the sender to utilize the tactic. The ICB dataset includes the power status for the crisis target and sender. The power status is coded 1 to 4, with one representing the weakest states and four representing great powers. The power differential variable was constructed by subtracting the senders' power status from the targets' power status. For example, if a greater power were to initiate against a weak state, the power differential would be -3 (1-4=-3). Conversely, if a weak state were to challenge a great power, the power differential would be 3 (4-1=-3). The range of the variable is –3 to 3, where positive numbers reflect a power advantage for the target, negative numbers represent a power advantage to the sender, and values close to zero represent parity.

The second independent variable added to the regression as a control looks at the nature of the threat or good in dispute. When the good in dispute is highly valued, public challenge and negotiation are less likely to overturn the status quo. Senders should be more likely to escalate to violence or use surprise when they believe that the good in dispute is highly coveted by the target, which is another way of saying that the target has high value for the status quo. Less important disputes should be less likely to trigger military reaction, thereby reducing the need to use surprise in order to gain an advantage. The ICB dataset includes a variable called gravity of threat, which codes the nature of the threat to the target. The variable is coded 1 to 6, where lower values cover issue areas such as economic and political threats. Higher gravity numbers are associated with

threats to territory, international influence or survival. When there are multiple issues at stake, the data was coded to reflect the highest threat level.

We use these four variables, as well as the two variables discussed in the previous section to test the formally deduced propositions below.

Empirical Results

The first statistical test examines the likelihood that the sender initiates a crisis with military action, Proposition 2. The proposition stated in chapter 4 is not directly testable. As discussed above, there are no actual observations of surprise in the data, only crisis triggers. Crises triggered with military action and violence is the closest variable to surprise. Second, entrenched beliefs are not directly observable, and instead we are using the combination of protracted rivalry and rational prior historical beliefs. The proposition is restated here as a hypothesis in an operational form that is testable with ICB data.

H1: When initiating a crisis, senders will trigger crises militarily less often given a target state with entrenched beliefs from involvement in a protracted rivalry.

This specifies the causal relationship where the trigger is the dependent variable and the entrenched beliefs tied to protracted rivalry is the independent variable of interest. Since the dependent variable is binary, indicating whether the sender triggered the crisis militarily, this hypothesis is tested with a logit model. The regression results from generalized linear models like logit are not interpreted the same way as one would interpret the output from an ordinary least square (OLS) regression. In the OLS Model, a

variable coefficient represents the change in dependent variable given a one unit increase in the value of the independent variable, holding the other independent variables constant (Wonnacott and Wonnacott 1990). Since there is a natural normalization in the maximum likelihood estimating procedure used for logit models, the coefficients on the independent variables do not directly correspond to the substantive effects. Instead, the substantive effects must be calculated, and are often done so by calculating the predicted using a difference method for an independent variable, holding the other variables constant (Kennedy 2003). The results from a few different specifications are presented below, but the full model takes the functional form,

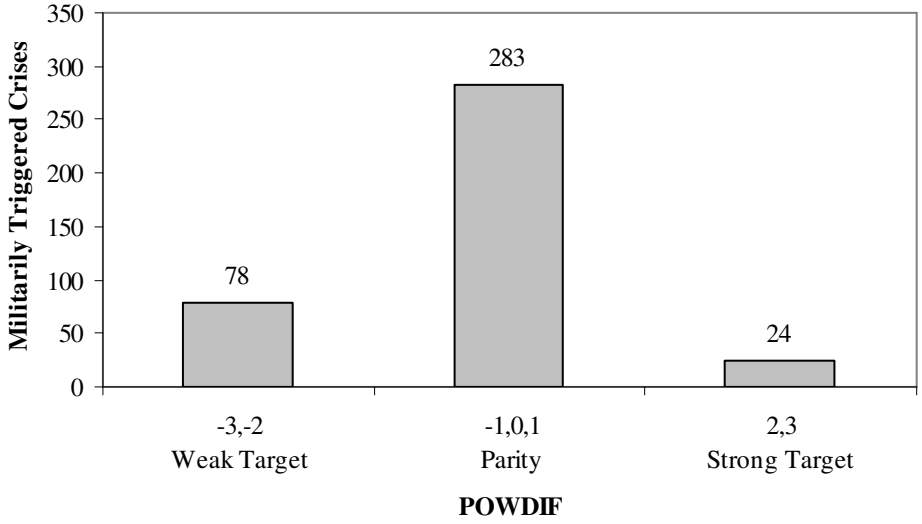
$$\Pr(TRIGGER = 1) = \frac{1}{1 + e^{-(\beta_0 + POWDIF\beta_1 + GRAVITY\beta_2 + PROTRACTED\beta_3 + RATPRIORB_4)}}.$$

The regressions results for five different specifications are found in Table 5.2.

The results presented in Table 5.2 provide strong support for the hypothesis that senders are less likely to use military triggers against protracted rivals controlling for rational prior beliefs. Model 1 limited inquiry to the two strategic variables that are recurrent in surprise literature: power differential and gravity of threat. Neither variable has much explanatory power over the sender's use of military triggers. This is particularly interesting when one considers that many theories of surprise suggest that it is used by weaker states (Handel 1987), meaning that the coefficient on power differential should be positive and significant. The coefficient is actually negative, but not significant. When one takes a closer look at the data, one finds that the correlation between power differential and military trigger is highest at the mid level of the power differential range. If we were to break the power differential into three discrete groups, which we will call relatively weak (POWDIF: -3,-2), relatively balanced (-1,0,1), and

relatively strong (POWDIF: 2,3). The relationship between power differential and crisis triggers is shown in Figure 5.2. Military triggers or surprises are actually employed most against relative equals rather than stronger states. In fact, senders rarely used military triggers to challenge a significantly stronger state. The inverse "U" pattern helps to explain why the power differential is not significant and challenges the idea that only weak parties use surprise against stronger ones. There is no clear pattern between gravity of threat and surprise despite the belief that surprise can be used to secure highly disputed goods as effectively as possible.

Figure 5.2: Militarily Triggered Crises and Power Differential



Note: Shows that states use military triggers against weaker powers approximately 20% of the time, against similar strength powers approximately 74% of the time, and against stronger powers approximately 6% of the time. States use military triggers most often when there is parity with its opponent.

Model 2 keeps the traditional explanations of surprise used in Model 1 and adds the rational prior belief variable. The rational belief of prior violence within a dyad provides an estimate on the probability that a crisis turns violent based upon the past crises in the dataset. The rational belief of violence in prior crises is strongly correlated

with the use of military trigger in future conflicts. This positive coefficient (1.560) is highly significant ($p<0.001$). The higher the probability of violence in previous crises, the more likely states are to trigger crises militarily. This would seem to be a perfectly rational response by the sender. If crises usually result in military confrontation, the sender has an incentive to take the initiative and trigger the crisis militarily. Despite the large size of the coefficient, the marginal impacts are relatively modest. The marginal effects at sample means are shown below in Table 5.3.

Table 5.2: Log Likelihood of Triggering a Crisis Militarily (Proxy for Surprise)

	Model 1	Model 2	Model 3
Constant	0.584	0.134	0.243
Significance	0.003	0.523	0.262
Power Differential	-0.070	-0.106	-0.108
Significance	0.280	0.119	0.113
Gravity of Threat	-0.022	-0.031	-0.011
Significance	0.691	0.583	0.856
Rational Prior Likelihood of Violence		**1.560**	**1.718**
Significance		**0.000**	**0.000**
Protracted Rivalry			**-0.420**
Significance			**0.027**

Note: 611 observations. Significant coefficients ($p<0.05$) are bold.
Summary: A history of protracted rivalry lowers the likelihood of attempted surprise, controlling for the power differential, the gravity of threat, and the rational prior belief.

Model 3 includes the three previous variables, and adds protracted rivalry as a proxy variable for irrational over-suspicion. Recall from the section on entrenched beliefs that since the model includes the rational prior control variable, any explanatory power for protracted rivalry addresses something beyond rational expectations based upon past experience and strong priors. The rational prior variable should control for past experience, leaving protracted rivalry to represent unobservable forces such as

entrenched beliefs about the aggressive intentions of the opponent. The coefficient on protracted rivalry is negative (-0.420) and statistically significant (p=0.027). This is strongly supportive of H1, which predicted that senders were less likely to use military triggers against protracted rivals, controlling for rational updating over previous experience.

Table 5.3: Log Likelihood of Triggering a Crisis Militarily (Proxy for Surprise) With Alternate Coding for Rational Priors

	Model 3	Model 4 Prior = 0.4/1.0	Model 5 Prior = 0.2/0.5	Model 6 Prior = 2.0/5.0	Model 7 Prior = 4.0/10.0
Constant	0.243	**-2.024**	**-1.629**	**-4.557**	**-7.352**
Significance	0.262	**0.000**	**0.000**	**0.000**	**0.000**
Power Differential	-0.108	-0.130	-0.132	-0.120	-0.115
Significance	0.113	0.065	0.062	0.082	0.093
Gravity of Threat	-0.011	-0.003	-0.002	-0.007	-0.009
Significance	0.856	0.961	0.977	0.904	0.883
Rational Prior Likelihood of Violence	**1.718**				
Significance	**0.000**				
Protracted Rivalry	**-0.420**	**-0.427**	**-0.420**	**-0.406**	**-0.386**
Significance	**0.027**	**0.026**	**0.030**	**0.031**	**0.038**
Rational Prior Adjustment 1		**5.796**			
Significance		**0.000**			
Rational Prior Adjustment 2			**4.787**		
Significance			**0.000**		
Rational Prior Adjustment 3				**12.228**	
Significance				**0.000**	
Rational Prior Adjustment 4					**19.263**
Significance					**0.000**

Note: 611 observations. Significant coefficients ($p<0.05$) are bold.
Summary: A history of protracted rivalry lowers the likelihood of attempted surprise, controlling for the power differential, the gravity of threat, and the rational prior belief.

One way to articulate the conclusion from this analysis above is that crises among protracted rivals are less likely to be initiated by surprise military action. Triggering states in protracted rivalries are more likely to initiate crises using non-military or

diplomatic means. Analysis of the marginal of effect of protracted rivalry at the sample means in Model 3 shows that senders in protracted rivalries are 9.5% less likely to attempt surprise to trigger a crisis.[15] Given some of problems of the rational prior belief variable highlighted above, it worth trying the adjusted measures of rational prior beliefs before discussing the results further.

Models 4, 5, 6 and 7 use alternative specifications for the rational prior variable, labeled rational prior adjustment 1, rational prior adjustment 2, rational prior adjustment 3, and rational prior adjustment 4. The output from these two regressions demonstrates that the results are very robust to the alternate specifications, and therefore the coding method in general. The coefficient on the rational prior adjustment 1 variable in Model 4 is positive (5.796) and statistically significant ($p<0.001$). According to this coding, the likelihood of militarily triggered crisis remains positively correlated with the rational belief on likelihood of prior violence, which is distinct from protracted rivalry as the proxy variable for entrenched beliefs. The proxy protracted rivalry variable remained negative (-0.427) and is still statistically significant at conventional levels ($p=0.026$). A similar story holds for the final rational prior variable tested in Model 5. The coefficient on rational prior adjustment 2 is strongly positive (4.787) and statistically significant. Once again, protracted rivalry remained negatively correlated with the use of military triggers (-0.420) and was statistically significant ($p=0.030$). These alternate models increase our confidence that the results are robust and provide support for the assertion in H1.

[15] The sample means are as follows: power differential (-0.303), gravity of threat (3.262) and Bayesian violence (.3522).

The coefficients on Models 6 and 7, with strong prior beliefs, tell a similar story. The coefficient on rational prior adjustments 3 and 4 remained positive (12.228 and 19.263, respectively) and were statistically significant (both had $p<0.000$). The coefficient capturing the effects of protracted rivalry on militarily triggered crises remained negative across the specifications (-0.406 in Model 6 and -0.386 in Model 7), and significant below the 5% level.

Table 5.4: Marginal Effects on the Likelihood Crises are Triggered Violently

	Model 3	Model 4 Prior = 0.4/1.0	Model 5 Prior = 0.2/0.5	Model 6 Prior = 2.0/5.0	Model 7 Prior = 4.0/10.0
Protracted Rivalry (0 to 1)	-9.5	-9.4%	-9.3%	-9.0%	-8.5%
Rational Prior Beliefs					
First Crisis Turns Violent (1/1)	31.7%	30.7%	32.4%	24.2%	21.5%
Second of Three Crises (2/3)	24.0%	23.5%	22.5%	24.2%	23.5%

Summary: The marginal effects of protracted rivalry are very robust to rational prior belief variable specification. Calculated at sample means.

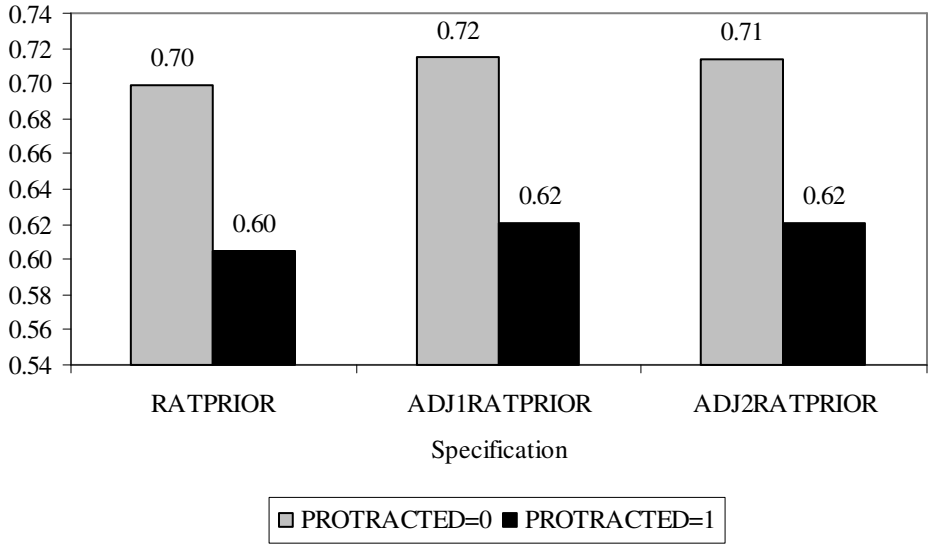
It is important to compare the marginal effects across the different specifications to ensure substantive confidence in the results and check the robustness of the variables. The coefficient value on protracted rivalry was similar across the specifications, but the rational prior coefficient did fluctuate as one might expect. The marginal impacts of the key variables are displayed in Table 5.4. The marginal effects on the rational prior variable are very robust to specification across different historical patterns. The first marginal effect looks at the expected change when the first crisis (1 of 1) turns violent, and the second measure calculates the marginal effects when the second of three crises turns violent (2/3). The results are very robust to specification. This is emphasized below

when we show that the results are robust by comparing the predicted results at sample means for the first three rational prior variables.

The effects of protracted rivalry are particularly robust across the different specifications for rational prior beliefs. Since the primary variable of interest is the effect of protracted rivalry controlling for rational prior beliefs, this helps to provide support for the hypothesized effects of entrenched beliefs. The marginal effects of protracted rivalry are displayed in Table 5.4. The consistency across the different models suggests that the effects of protracted rivalry are not the product of the coding method, but reflective of important empirical patterns within the data. If protracted rivalry and entrenched beliefs are positively correlated, which is to say that protracted rivals are more likely to rely on preconceived notions about their rival, then the analysis supports Proposition 2. This is also supported by Figure 5.2, which calculates the likelihood of violent military trigger, using sample means, across the specifications. The results show that the models are very robust to the different methods of coding rational prior beliefs.

Since the use of proxy or alternate variables risks capturing effects other then those desired, it is important to think about alternative explanations for the observed behavior. The pattern shows that states are less likely to use military triggers against protracted rivals controlling for rational prior beliefs. One alternative explanation is that the coefficient on protracted rivalry is negative because these rivalries peter out over time. This alternate hypothesis would deviate from recent work suggesting that these crises become increasingly entrenched and difficult to resolve over time (Hassner 2007). This explanation must be ruled out if we are to have confidence in the results, and two additional pieces of analysis help to reject this alternative hypothesis.

Figure 5.3: Likelihood of Military Trigger Contingent on Specification



Note: Sample Means for POWDIF=-0.303, GRAVTY=3.262, RATPRIOR=0.352, ADJ1 RATPRIOR=0.501, ADJ2 RATPRIOR=0.524.
Summary: Graph shows that the marginal effects of protracted rivalry, the substitute variable for entrenched beliefs, are almost identical across the specifications. Protracted rivalry decreases the likelihood of military trigger by 9%-10%.

The first of these analyses is a regression that looks at the likelihood of Bayesian violence, the rational prior belief, over historical cases. Recall that the rational prior variable captured all of the crises between a specific sender and target before a given crisis observation in the dataset. If rivalries were to peter out over time, then one would expect that Bayesian likelihood of violence would decrease as the number of historical crises within a dyad increased. A simple ordinary least squares regression, with the rational prior likelihood of violence as the dependent variable, and the number of prior crises between the two states as the independent variable, does not support the idea that rivalries become less violent over repeated interaction. The regression results actually provide strong support for the contradictory argument. The results show that the number of historical cases is positively correlated with rational prior likelihood of past violence, a

coefficient 0.173 with a standard error 0.010, making the effect statistically significant. This provides support for Hassner's (2007) argument that rivalries can grow increasingly entrenched and dangerous with time, since the predicted results of the model show that there is a 17.3% increase in the possibility of violence between a specific sender and target with each additional crisis. A second analysis looked separately at all the protracted rival dyads with six or more crises, and suggests that there is no clear pattern between the duration of the rivalry and the likelihood of violence. Some dyads experience increasing levels of violence, and others experience decreasing levels of violence.

Conclusions

This chapter provided some quantitative evidence using protracted rivalry as a proxy variable for entrenched suspicion, which seemed to support proposition 2 stating that the possibility of irrational over-suspicion should reduce the incentive for initiators to use surprise. Before presenting the statistical analysis, we had to walk though the empirical approach to testing a theory based upon psychological bias and surprise. The absence of clearly observed psychological bias makes testing its effects difficult. This problem was circumvented by careful thought about the types of states that are most likely to be irrationally over-suspicious and the variables that might capture those effects. The methodological approach used to study these effects relied on the hypothesized correlation between entrenched beliefs among protracted rivals and the information available to those rivals. By controlling for prior experience, which was done by using a Bayesian calculation to look at the rational prior likelihood of prior violence between two states, entrenched beliefs capture something more than historical experience. This

provided an innovative way to isolate unobservable effects of rivalry and psychological biases.

Using the proxy variables to control for rational priors and extract an approximation for entrenched beliefs, the statistical analyses provide support for the restated hypothesis. The hypothesis addressed sender's use of military action to trigger crises. The theory suggested that senders facing a possibility of irrational over-suspicion should be less likely to use military triggers for fear of an overly aggressive response from the target. The statistical results provide strong confirmation. Controlling for the rational prior beliefs, senders are approximately 9.5% less likely to use military triggers when the target is a protracted rival. This result was very robust to the coding of the rational prior variable. This does suggest that initiators are more careful when triggering against states that may have entrenched beliefs associated with their rivalry. In short, entrenched beliefs may alter the behavior of an opponent and have the strategic benefits suggested in the theoretical model.

While the statistical approach to measuring or approximating entrenched beliefs is innovative, its limitations are evident. The next chapter tries to balance out these constraints by examining cases in detail. A detailed analysis of two cases permits us to make assertions about bias and its effects that are easier to empirically justify than the aggregate method used here. This is where we turn our attention next.

Chapter 6:
Empirical Analysis Part 2:
Qualitative Case Analysis

This chapter presents two case studies of crisis decision-making to evaluate the deductive propositions from chapter 4. There are three primary deductive conclusions: 1) irrational under-suspicion[16] does not increase the probability that a target is surprised[17] (or its expected utility) since the initiator will attempt surprise as much as possible anyway; 2) the possibility of irrational over-suspicion[18] should deter an initiator from attempting surprise since target over-suspicion reduces the likelihood that the attempted surprise is effective; and 3) targets can derive utility benefits from their own irrational over-suspicion when the likelihood of unnecessary defensive action (false alarms) is sufficiently small. The cases presented here provide support for and illustrate the theoretical relationship between entrenched beliefs[19] and strategic behavior.

The two cases examined, the Cuban Missile Crisis and the 1967 Arab-Israeli War, were selected because they illustrate how the expected outcomes of the model with a potentially biased target[20] (biased model) and the Bayesian model can converge under

[16] Recall that this refers to situations where a target irrationally believes the sender is less threatening than they are.
[17] Recall that surprise is an outcome in the model that results when a sender challenges privately (attempts surprise), and the target chooses to wait rather than take defensive action.
[18] Recall that this refers to situations where a target irrationally believes the sender is more threatening than they are.
[19] Entrenched beliefs referred to the tendency to anchor or irrationally overweigh prior beliefs relative to new information.
[20] In the bias model created in chapter 4, there are two types of targets where some possibility that a target suffers from entrenched beliefs and some probability they are Bayesian rational. The sender knows this

certain conditions and diverge given others. The theoretical models showed that the biased and Bayesian deductive results converge when targets perceive senders as non-threatening, where non-threatening refers to conditions where the prior likelihood the sender is aggressive is below the target's critical belief.[21] The critical belief was endogenously determined by the target's preferences. Even if the target might suffer from entrenched beliefs, their bias results in irrational under-suspicion, and the sender is best off adopting the standard Bayesian strategy where aggressive types attempt surprise all of the time. Conversely, the sender should adopt less provocative strategies when there is some possibility that the target has entrenched beliefs and they perceive the sender to be threatening.[22] These targets suffer from irrational over-suspicion making it more difficult for senders to surprise targets than if the targets were perfect Bayesians. The case studies in this chapter provide illustrative support for these theoretical deductions.

Prior to the outset of the Cuban Missile Crisis, U.S. decision-makers and intelligence estimates reflected the belief that Soviet aid to Cuba in 1962 was defensive and not particularly threatening (Special National Intelligence Estimate: The Military Buildup in Cuba 1962). The theoretical deductions predict convergence between biased and Bayesian models because the sender's endogenously selected strategy dictates that they use surprise irrespective of target biases. Later in the crisis, when the U.S. clearly knows about the missiles, expected outcomes of the biased and Bayesian models should diverge such that the Soviets feel pressured to adopt less provocative strategies. By contrast, Israeli assessments in 1966-67 clearly reflect concern over Egyptian intentions

---

probability. The possibility that a target may be biased or rational in the biased model is significant in the case analysis.

[21] Technically, this refers to situations were $p<p^*$.

[22] Technically, this refers to situations were $p>p^*$.

before the crisis. The Israelis did not think that Egypt would start a full-scale war before 1970, but they did believe that Nasser might move to close the Straits of Tiran or attack the Dimona nuclear facility. Egypt posed a sufficiently large threat to Israel that the expected results from the biased and Bayesian models should diverge throughout most of the case analysis. The biased model predicts that Egypt should be more likely to escalate the crisis in public fashion than the Bayesian model expects, and an unbiased Israel should wait while an entrenched Israel will adopt a mixed strategy of waiting and defensive action.

The case studies in this chapter suggest that the causal mechanisms that drove convergent and divergent results in the theoretical model have the same effect when traced through the decision-making processes in the two crises. Each crisis is broken into stages, the relevant causal conditions are explored, and then the impact of those conditions on outcomes is discussed. The results of this analysis are summarized in tables that explicitly compare the Bayesian and biased models, using statistical analysis to identify whether the one model had better explanatory power than its competitor. In the Cuban Missile Crisis, where certain conditions favored convergence in expected outcomes, the comparable analysis shows some convergence and some divergence, as the pertinent conditions would suggest. The conditions in the Arab-Israeli War case favor divergent results, and the explanatory power of the biased model statistically outperforms the rival Bayesian explanation.

Before examining the actual two case studies, the chapter considers some of the scientific issues associated with quantitative and qualitative empirical methodology. First we consider the relative merits and constraints associated with each research method as

applied to international politics and theories of decision-making that involve psychological bias. Then the chapter examines case study methodology in more detail, stressing a method called process-tracing, which is the historical evaluation of a causal process of interest by identifying a set of initial conditions and translating them into outcomes. The chapter then reviews the key elements of the theory laid out above to emphasize the conditions and results of interest. After a brief section on case selection, the chapter proceeds to the two cases starting with the Cuban Missile Crisis. Each case analysis starts by offering relevant background, an analysis of the relevant case history, a discussion of entrenched beliefs, a summary table comparing the Bayesian and biased models, and a discussion of the coding and results. The chapter ends with some brief concluding thoughts.

Psychology and Research Method in International Politics

The previous chapter identified one, if the not the greatest, difficulty to testing psychological explanations of events in international politics: the unobservable nature of psychological bias and consistent coding for this unobservable variable across cases. We tried to overcome this constraint by introducing an instrumental variable for entrenched beliefs that allowed for consistent coding across a large dataset of crises. This method was based on the idea that entrenched beliefs were more common when states were protracted rivals than when they were not. The statistical research across the large set of cases shows that protracted rivalry is negatively correlated with hostile triggers, the proxy variable used for surprise, but establishing causation is more difficult.

All statistical research, in the absence of deductive models, is inductive reasoning, using results to infer relationships, among variables (Steinberg 2007). The relationships are best when they reflect causality. The problem, which is magnified by the use of a proxy variable, is that it is difficult to know whether the entrenched beliefs associated with protracted rivalry actually influenced the decisions. We know that there is correlation between the dependent and independent variables of interest, but it is difficult to know whether the independent variable is really causing changes in the dependent variable. Case studies provide a method to trace causation in more detail than possible in statistical studies.

Technically, there are three interrelated problems that dictate the necessity of further causal and case analysis: validity, multiple explanations and a lack of repetition (Tetlock 1998a). Validity generally refers to the confidence that researchers have in their causal results. There are two types of validity, external and internal, which are usually perceived as a tradeoff (Campbell and Stanley 1963). The statistical research across the large set of cases was crucial to developing external validity, confidence that the findings are generalizable across a variety of conditions or situations. The use of many cases in this statistical study provides a high degree of external validity since the results are statistically significant across 600 crises over 80 years. We can have confidence that a correlation exists between the instrumental independent variable and the dependent variable.

In contrast, internal validity addresses the likelihood that changes in the dependent variable were caused by changes in the independent variable. In the social sciences, internal validity is usually associated with controlled experiments (Tetlock

1998b). Researchers using these experimental techniques can manipulate the independent variable of interest, while attempting to hold all other conditions constant, improving the confidence that changes in the independent variable caused changes in the dependent variable. One advantage of high internal validity is that the researchers can be reasonably sure that differences in behavior or beliefs are the result of psychological dynamics when they can be isolated or identified. As internal validity increases, it is easier to connect cause and effect, but it is more difficult to generalize the results beyond the controlled environment set up by the researchers.

International politics, by its nature, is a field of inquiry where internal validity is very rare, because it is difficult to run controlled experiments. Researchers have no control over their subjects, whose behavior or beliefs are altered by numerous stimuli at any one time, rather than the single treatment possible in laboratory settings. We know that the events of interest in international politics will apply or function in the real world, since cases are drawn from that world. Establishing internal validity to prove that our variable of interest is a pertinent causal driver becomes much more difficult.

The conflict between internal and external validity alerts us to a second, but related, issue. Internal validity would not be a cause for concern if there were agreement on causal processes. The lack of internal validity becomes a problem in international politics, because there are many plausible theoretical explanations for many observed phenomena. While the study of international politics may suffer from a lack of causal agreement, it does not suffer from a lack of causal theories. The early theories stressed interests and power (Morgenthau 1948; Waltz 1979), but there are large bodies of work that also consider decision-making (Allison 1971, 1999; Jervis 1976), institutions

(Keohane 1984; Axelrod 1984), and social construction (Wendt 1992; Ruggie 1998). Scholars differ on the relative merits assigned to each explanation, and in rare cases the validity of some altogether. Just as it is difficult to rule out potential causes, it is also possible that many causes operate at the same time. Philip Tetlock (1998b) refers to this as the interrelatedness of cases. These debates, combined with the difficulty of establishing internal validity, make it difficult to ascertain causality in a convincing fashion.

The final issue noted above was the problem of repetition, which is also tied to the problems of validity and multiple explanations. A central tenant of the scientific process is that results should be replicable (Keohane, King and Verba 1994). Ideally, social scientists should be able to recreate the conditions of a controlled experiment and get similar results. A problem in international politics, and in fact any field relying on historical evidence, is that the tape of history runs once. There are no opportunities to rerun international crises with different triggers, decision-makers or power differentials. Empirical examinations might try to look at two cases within the same state or political administration, thereby holding some variables constant while allowing others to vary. Statistical study helps to overcome this issue by offering a mathematical method to control for variance across many instances of similar phenomena like crises, but it is not perfect repetition and lacks internal validity (Steinberg 2007). The lack of repetition also motivates the use of counterfactuals to identify what could or would have happened if the case were rerun under different conditions (Tetlock 1998b). Needless to say, this type of counterfactual reasoning is often controversial, and is impossible to prove.

Before moving on to the case method and studies themselves, there are two further issues worth discussing, in large part because they are a recurrent theme of this project. First, analysis of psychological bias in decision-making, or decision-theoretic contexts, often stresses the detrimental effects of bias on decision quality. There is a tendency, therefore, to assume that biases are detrimental in competitive or strategic decision-making contexts. The analysis here showed that this assumption might be inaccurate. Observing a poor decision-making process does not mean that the actual decision produces a poor strategic result. Much of the work in chapter 4 stressed this idea, and showed conditions whereby biased decisions may actually be irrelevant or beneficial in strategic decision-making.

The second point is tightly tied to this idea of placing causation within the proper context (i.e. decision quality versus competitive interaction). The tendency with case study analysis and retrospective studies of policy failure is to observe the poor outcome and reverse engineer a reason for failure. To that effect, we constantly link poor policy outcomes to poor decision-making or assessment, when in fact, that relationship may not exist or may be spurious. In the competitive complex environments that get policymaker's attention, decision-making quality might not necessarily correlate to outcome quality. This chapter challenges this idea in the hopes of improving the rigor used to assess decision-making, strategic outcomes, and the causal mechanisms that link them.

Case Study Methodology

Political scientists debate the relative merits of qualitative approaches, and there is no consensus. Gary King, Robert Keohane, and Sidney Verba (KKV) (1994) use a quantitative framework as a guide to best practice in qualitative analysis. They identify five rules of causal interference that apply to both quantitative and qualitative research. The rules are as follows: 1) construct falsifiable theories, 2) build theories that are internally consistent, 3) select dependent variables carefully, 4) maximize concreteness, and 5) state theories in as encompassing ways as feasible. Some of the lessons gleaned from this approach include careful specification of theory and variables, the use of more cases, risks of selection bias, and concern about measurement bias.

The empirical sections of this paper have sought to meet these requirements as best as possible. The use of formal theory or game theoretic analysis helps in the creation of theories that are internally consistent. This is one of the primary benefits of rigorous rational choice analysis. The dependent variables, which will be discussed at greater length below, and were already addressed in the previous chapter, are specified to ensure that there can be variation in the primary variables of interest. Finally, the theory is stated as encompassing as possible, relying on variables like entrenched beliefs and surprise. This helps to capture as wide an array of phenomena as possible.

Two of the five criteria that KKV identify may be problematic for this research, and it is important to discuss these shortcomings. First, and perhaps most important, is the creation of a falsifiable theory. Logically, a falsifiable theory is one that is specified or worded such that causal propositions can be deemed true or false. This requires creating a theory that specifies dependent and independent variables, and posits some

causal relationship between them. The purpose of this design is to empirically test the hypothesized relationship to establish whether it is in fact true or false. The theory constructed in this project is logically falsifiable. There are clearly specified dependent and independent variables along with causal linkages. In that sense, the theory is perfectly falsifiable and satisfies with KKV's first criteria.

The core empirical problem with this project is the observation of variables such as entrenched beliefs and surprise. This is further complicated by the requirement to establish a causal link such that entrenched beliefs played a role in certain decisions. To the extent that there is agreement on the presence or absence of causal variables, and their causal relationship, the theory is perfectly falsifiable. This suggests that the problem is actually one of concreteness rather than falsifiability. If these variables were clearly observable, we would be able to rely exclusively on the large-n or statistical methodology that KKV believe to be the ideal research design. Causal factors like entrenched beliefs, and their relation to subsequent decisions, are not necessarily concrete variables that are easily discernable. The difficulty of establishing the existence and causal linkages associated with concreteness is one of the reasons that we also rely on a set of case studies to empirically examine the deductive propositions.

In response to KKV's approach to empirical analysis, some researchers have argued that qualitative analysis is not an inferior substitute for quantitative work but a valuable pursuit in itself. These researchers have suggested that KKV's quantitative approach to empirical study need not be applied universally and is too restrictive (Gerring 2001; Steinberg 2007). The purpose of this qualitative study is to help identify the variables of interest, trace the decision process in order to identify the causal

relationships, avoid the use of instrumental variables that were necessary in the statistical chapter and provide further support for the theoretically deduced propositions.

Process-tracing is a qualitative research method recommended by KKV, and advocated by many qualitative analysts. It is a particularly strong tool for within case analysis (Bennett and Elman 2006). Simply put, process-tracing is the historical evaluation of a causal process of interest, often something like a complicated decision or the evolution of an institution, that identifies a set of initial conditions and translates them into outcomes (Pierson and Skocpol 2002). This has the effect of tracing or identifying the causal impact of the conditions, what we might think of as the independent variables, on the outcome or dependent variable. By looking at the details of the causal process this analysis moves beyond mathematical covariance. Statistical analysis is based on establishing a mathematical pattern between variables, but does not clearly distinguish cause and effect from mathematical correlation. By looking at the way that initial conditions contribute to outcomes in detail, we may be able to get a better sense of causation and test a wider range of hypotheses. This makes the purpose of process-tracing something different from statistical approaches of inference or empirical assessment.

There are two key elements involved in process tracing. First is distinguishing between process induction and process verification (Bennett and George 1997). Process induction is the use of process-tracing to identify the relevant causal variables. Process verification involves testing theories from theories developed prior to the empirical examination of the case. This analysis focuses on process verification rather than process induction since the theory was already deduced in prior chapters.

Second, it is important to offer a distinction between historical analysis and process-tracing, since some suggest that process-tracing is nothing more than good historical analysis. There is certainly overlap between the two. The primary distinction is that process-tracing seeks to convert a purely historical analysis or narrative into a causal assessment based upon a carefully specified theory and relevant variables (Bennett and George 1997). The process of converting historical analysis into causal assessment almost always requires that researchers strip away some of the detail, thereby drawing criticism that process-tracing requires simplification. While this almost certainly true, it also speaks to the fundamental difference between historians and political scientists, where the latter focuses on the development and testing of causal theories. Process-tracing is often best accomplished by focusing on micro-level decisions and details, to understand how the variables work in sequence or congruence to influence outcomes. To the extent that analysis focuses on micro rather than macro-level detail, the risks of oversimplification are minimized, though not eliminated.

The final element discussed here is the importance of generating different plausible hypotheses in order to arrive at the most accurate assessment of the causal theory and the underlying mechanism. Two problems arise from this practice. First, any probabilistic theory runs the risk of being an outlier. That is to say, deterministic theories are ideal for case analysis since certain conditions need to be necessary or sufficient to produce an outcome over repeated trials. Much of social science today, evidenced by the prominence of statistics and statistical solutions of game theoretic models, is most correctly treated as probabilistic. This makes it difficult to know whether the hypothesized causal process produced observed outcomes, since the hypothesis suggests

that different outcomes are expected based upon a probabilistic draw. The second issue, related to the first, is the problem of indiscriminate pluralism. Indiscriminate pluralism refers to the inability to eliminate possible explanations. There is no clear solution for this recurrent, and perhaps even naturally embedded, problem. Since process-tracing is usually used to assess complicated decisions and processes, it is difficult to boil down decisions or institutional evolution to a specific factor while eliminating other plausible hypotheses. The best analysts can do is consider a mix of causal drivers and try to rank their importance (Steinberg 2007).

Tracing Entrenched Beliefs and Strategic Behavior

Before actually presenting the two case studies, it is important to clearly layout the analytical method, hypotheses and variables of interest. This guarantees that each case study actually captures the issues and relationships of interest, and helps to ensure that the empirical analysis is consistent across cases. This section briefly reviews the relevant variables and the relationships between them.

Entrenched beliefs are a crucial component of the theory. It is important, therefore, to handle this issue carefully in the empirical section and case studies. Entrenched beliefs, or anchoring bias, have been defined as the tendency to irrationally ignore or marginalize new information in order to maintain initial beliefs. Evaluating beliefs and biases retrospectively is a very difficult endeavor, one that was discussed in chapter 3, because perfect Bayesian actors may discount new information for rational purposes based on the strength of their prior beliefs. Distinguishing between rational and irrational discounting is important for any theory of behavior relying on psychological

bias, but it difficult to do in a rigorous fashion. This analysis tries to go beyond merely identifying a belief, accurate or faulty, and declaring it rational or biased. An entrenched belief arises over time as new information becomes available and is subsequently marginalized. Entrenchment, therefore, is a process that can be traced empirically like other decision-making processes. Like tracing decisions, entrenchment might not always be straightforward, but using the sequencing and process oriented approach provides a higher level of confidence that entrenchment is going on over time rather than declaring it retrospectively. In each case analysis, the information known to the actors will be presented and then assessed to establish the effect on the pertinent belief set over time.

There are three outcomes of interest given the theory specified above: 1) the sender's decision to use public challenges or surprise, 2) the target's response to the sender's challenge, and 3) the outcome of the interaction.[23] To be precise, the target's response can actually be disaggregated further. The target's response refers to the sense of surprise given a challenge, as well as target response. The way the strategic problem is designed or specified, these two elements are tied together. The target was effectively surprised when they did not take preventative action while the sender was attempting surprise. This problem is not particularly acute in the case analyses that follow, but we are interested in the relationship between the belief and the observed challenge.

In fact, we can push this point a little further since we are particularly interested in the way that beliefs influence subsequent decisions for both actors. The primary causal mechanisms are the beliefs, and potential biases, of each actor. The resulting actions, and

[23] Recall that the basic surprise model consisted of three strategic moves. The sender could challenge or maintain the status quo. If the sender challenged they could do so publicly to attempt surprise (where surprise is not clearly differentiable from maintaining the status quo). Then the target had to decide whether to take defensive action or wait.

the combination of actions producing the outcome, are the dependent variables of interest. This means that the three outcomes specified in the paragraph above are the primary causal outcomes or dependent variables associated with each hypothesis. Given the interrelationship between the hypotheses, it seems logical to present them in a form that clarifies each part of the theory while improving the applicability of the theory to the case studies.

The first proposition deduced in chapter 4 argued that target's entrenched beliefs had no impact on the likelihood of suffering surprise. This result was conditional on the assumption that the sender had a high likelihood of being docile. The primary mechanism driving the result was that the sender's behavior was unchanged by the possibility of target bias as aggressive types attempted surprise and docile types tried to maintain the status quo. The strategic dynamic dictated that targets did not take defensive action, and always received an ambiguous signal. The sender, in this instance had no incentive to take the target's entrenched beliefs into account, and played the same strategy irrespective of whether the target was rational or potentially biased.

The second proposition focused the sender's strategy when the initial likelihood that the sender was the aggressive type was above the target's critical belief, making the sender threatening to the target. When the sender had a higher likelihood of being aggressive, entrenched beliefs actually had a stabilizing effect. A sender facing a potentially irrationally over-suspicious target was better off attempting surprise less often since the target was more sensitive to the sender's signals. A signal that made a rational target indifferent compels the entrenched target to take defensive action, thereby nullifying any potential benefit of surprise for the sender. The sender's best response

given the possibility of an entrenched target, therefore, is to adopt a less provocative strategy by using public challenges more often than they would if they were facing rational Bayesian opponent.

These two propositions talk about two sides of the same coin. The first proposition is valid when the sender is initially perceived to be non-threatening. The second proposition is relevant when the sender is initially perceived to be threatening. This means that the first step requires identifying the prior conditions, specifically the initial likelihood that the sender is perceived to be aggressive. When the sender is perceived to be non-threatening, there is no reason for the sender to take the target's entrenched beliefs into account. Even if they did consider the possibility that the target had entrenched beliefs, it would not alter the sender's optimal strategy. Conversely, senders perceived to be threatening should be concerned with the target's potential entrenchment. Sender's perceived to be aggressive should adopt more conservative challenge strategies when they believe that the target may be irrationally over-suspicious. Note they need not believe that the target has entrenched beliefs, they only need to consider it a possibility. Senders perceived to be threatening, who believe they are facing perfectly rational targets, should be more prone to adopt surprise strategies or attempt *fait accompli*.

While the first and second propositions addressed the sender's behavior contingent upon a set of conditions, the first and third propositions speak to the target's behavior. Both propositions, explicitly (as in proposition 1) or implicitly (as in proposition 3), address the target's likelihood of being surprised and taking action. The first proposition stated that targets would never take action, whether they had entrenched

beliefs or not, if the sender was perceived to be non-threatening. In other words, if the target started with the belief that the sender was non-threatening, there is no possible information or signal that would drive the target to take action whether they were rational or entrenched.

The impact of entrenched beliefs on the likelihood of target action is more complicated in proposition 3. This proposition applies when the sender is perceived to be sufficiently threatening. The proposition directly referenced the target's expected utility, but the expected utility is related to the target's susceptibility for surprise. Recall that a potentially biased target may or may not actually exhibit entrenched beliefs. Potentially biased targets that act rationally are more likely to be surprised than a Bayesian rational actor, whereas the entrenched target is less likely to suffer surprise and more likely to act. This means that an entrenched target with the belief that the sender is threatening is more likely to make defensive preparations that foil attempted surprise than the perfectly rational actor given the same signaling strategy from the sender.

It is important to clarify a nuance of the deductive theory. The paragraph immediately above noted that a potentially biased actor could actually be either biased or perfectly rational during any one play of the game. Recall from chapter 4 that the target had two possible types, where type referred to their information updating or learning quality. In the game, the type was decided by a move by nature, which means it is drawn from a probability distribution known to the sender. During any play of the potentially biased game, the target could actually be a perfectly rational type updating their beliefs according Bayes' Rule, or be a biased type that irrationally anchored when updating. In the case analysis that follows, the target's behavior is dependent on the type of target that

is selected in crisis sub-case. The target's belief updating is analyzed in each sub-case, where the default is to code the target as rational unless there appears to be some reasonably good evidence to the contrary. This means it is necessary identify which is the relevant target type in the potentially biased model for each sub-case comparison to the perfect Bayesian model. Throughout the analysis we use terms like "biased model with the Bayesian type" and "biased model with the biased type" to distinguish between the rational target type and irrational target type, respectively, since there is possibility of both in the biased model.

The final outcome is a combination of the sender's strategy and the target's response. The quality of the outcome is not simply the target's decision to act. In reality, strategic interactions, and particularly outcomes, are rarely as simple as an initiation and response. While the case analysis will focus first on the initiation and then the response, it is difficult to limit any discussion of outcomes to the initial stages. To overcome this problem, we evaluate the crisis result and the impact of entrenched beliefs.

Case Selection

Before proceeding to the case studies, a short word on case selection is in order. The section on case methodology above noted the importance of variable selection, specifically the dependent variable. KKV point out that case studies with no variation in the dependent variable are problematic for proving or disproving hypotheses. Comparative case analysis where the outcomes of interest are the same may help to establish necessary or sufficient conditions, but cannot be used to test a causal statement that links an independent variable with a predicted outcome. For this reason, this chapter

will examine one case study where surprise was essential to the initiator's strategy, and a crisis that was triggered in public fashion.

Variation in the dependent variable of interest is one requirement. A second requirement, which is particularly important here, is that there be sufficient primary and secondary source material to address the actors' beliefs and biases in a convincing fashion. The purpose of using case studies, particularly in support of the quantitative study, is the ability to trace the causal impact of beliefs and biases. An inability to capture or identify these variables in a convincing way would obviously offset the benefits of the case analysis. The best way to overcome this constraint is by selecting cases where there is a good deal of information available, both primary and secondary sources. The irony, however, is that cases with significant secondary source material are likely to be controversial, making consensus among the experts less likely. Despite this constraint, it is easier to address beliefs and biases in cases with a large body of literature.

The two cases addressed below are the Cuban Missile Crisis of 1962 and the Arab-Israeli War of 1967. Both of these cases have been written about extensively, and there is some degree of consensus among experts on some features of each case. The Cuban Missile Crisis is an ideal case for this study due to the rich archive of primary source material including transcripts and declassified memos from the U.S. side. The source material from the Soviet side is weaker, but some recent historical works on Khrushchev provide more insight into Soviet decision-making. The 1967 Arab-Israeli War has received a good deal of attention from scholars. There is sufficient secondary source material and there are also memoirs from many involved. Again, decision-making on the Egyptian side is less transparent than the Israeli side, but there were some

biographers that had access to Egyptian leadership before and after the war. Together, these cases offer variation in the variables of interest, and sufficient source material to examine the causal impact of beliefs and biases.

The Cuban Missile Crisis, 1962

There are numerous points of departure that one could use for the study of the Cuban Missile Crisis, owning in part to the voluminous literature on the subject. The most well known and cited of these works remains Graham Allison's (1971, 1999) *Essence of Decision*.[24] Allison applied three different models of decision-making to explain both Soviet and U.S. behavior, and stressed that a model of bureaucratic bargaining offered insights into the crisis that were missed by the more common rational actor and organizational behavior models. One of the book's strengths, however, is its attempt to analyze and identify relevant decision processes and drivers for both crisis actors. We follow this approach here. The theory, as laid out in prior chapters, is inherently strategic. The actors condition their behavior on the beliefs and likely responses of the adversary. We begin by studying U.S. beliefs prior to the crisis, and then examine how that may have influenced Soviet behavior.

The U.S. decision-making process during the Cuban Missile Crisis has been studied extensively, and is often treated as a prototypical case of successful decision-making (Neustadt and May 1986; Herek, Janis and Huth 1987). In an unstructured and egalitarian fashion, President John F. Kennedy and members of the Executive Committee (often referred to as the ExComm) examined Soviet motivations, possible U.S. responses,

[24] The 1969 article is cited 260 times and the 1999 version of the book is cited over 1,800 according to googlescholar. It is the most cited entry for "Cuban Missile Crisis."

and likely Soviet counters (May and Zelikow 2001). Opinions and options were raised, debated, researched further, and often debated again. The debate was unstructured but thorough, sought out positions from all sides and worked towards a consensus solution. In the end, the Excomm chose to blockade Cuba, which provided a window for diplomacy if the Soviets were willing to stand down. While in retrospect the blockade was an optimal decision, it was also a gamble that was partly justified by assumptions and beliefs that turned out to be inaccurate. A detailed examination of the actors' beliefs during the crisis reveals tendencies towards both flexibility and anchoring[25] on different issues. The Cuban Missile Crisis shows how the biased and Bayesian models can converge to common expectations of behavior given certain contextual conditions while diverging under others.

The Cuban Missile Crisis was technically triggered for U.S. policymakers on the morning of October 16 when they were shown photos of medium and intermediate range missiles in Cuba. The roots of the crisis, however, had been growing beneath the surface for sometime (Beschloss 1991, p. 362). The revolution in Cuba and the rise of Fidel Castro's government presented obvious problems, or at least uncertainty, for U.S. policymakers. At first, U.S. policymakers could not tell whether Castro and Cuba would become an ally or adversary (Fursenko and Naftali 1997, p. 16), but Castro's attack on U.S. business interests signaled the inevitability of movement toward the Soviet sphere. The presence of a communist government so close to home was a source of embarrassment and concern for U.S. policymakers. No sooner did Castro become entrenched in his leadership role than the U.S. began to look for ways to undermine his regime (Khrushchev 1974, p. 510).

[25] Recall that anchoring is the tendency to irrationally overweigh prior beliefs relative to new information.

Concerned that direct action against Cuba would embolden the Soviets, possibly setting the two countries on a collision course, the U.S. resorted to covert action (Blight, Allyn and Welch 1993, p. 17). The most visible of these actions was the Bay of Pigs invasion, a covert operation run by the CIA that relied on an invasion force of Cuban exiles. The 1961 invasion, carried out with the help of U.S. forces, was a failure. The invasion became a source of embarrassment for the Kennedy administration despite the fact that many of the details were arranged before they took office. After the Bay of Pigs debacle, the administration softened its position on Castro and Cuba (Freedman 2000, p. 147). Many still wanted to remove the Castro regime from power, particularly Attorney General Robert Kennedy, but they were forced to rely on a sabotage and espionage program that was codenamed Operation Mongoose. The operation did not have the type of success that policymakers hoped for, having little direct impact on the stability of Castro's government.

In early 1962, it looked as though the U.S. would have to live with a communist neighbor. While maintaining its covert efforts to undermine or assassinate Castro, U.S. policymakers also considered the role that Cuba would play in Cold War strategy (Freedman 2000, p. 148). Cuba had relied on the Soviet bloc for weapons supply since the late 1950s. Some of the weapons used during the revolution were funneled through a company in Costa Rica that acted as front for the Czechoslovakian government (Fursenko and Naftali 1997, p. 12). During Castro's 1959 visit to the U.S., his brother Raul was secretly negotiating with the Czech and Polish governments for additional military supplies. Cuba also requested Soviet military and intelligence advisors to help train its fledgling forces. At first, the Soviets rebuffed many of these requests, concerned that

these actions would bring the U.S. into direct conflict with the Soviet Union (Taubman 2003, p. 533).

The Soviet position towards Cuba changed in 1960. KGB agent Aleksandr Alekseev, who arrived in 1959, had formed a strong relationship with Castro. This allowed him to provide Moscow with more information pertaining to Castro's leanings and offered Moscow direct contact with Cuban leadership (Khrushchev 1970 p. 492). That same year, Castro also welcomed a visit from a member of the Soviet Presidium, Anatas Mikoyan. Nikita Khrushchev's decision to offer Soviet aid was influenced by two other events. The first was Castro's challenge to U.S. business interests. Castro had negotiated a deal to import Soviet oil at prices cheaper than U.S. suppliers, meaning that Soviet oil would be refined in facilities owned by U.S. companies. Should the companies refuse to refine the oil, as Castro expected, he would have justification for seizing the refineries (Fursenko and Naftali 1997, p. 48). The second, perhaps more pressing issue, were persistent rumors about the possibility of U.S. military action to remove Castro from power (Khrushchev 2000, p. 482).

Khrushchev responded to concerns about U.S. intervention in Cuba by extending the Soviet nuclear umbrella into the western hemisphere and offering significant military aid to Cuba (Fursenko and Naftali 2006, p. 305). The former was done in a speech in June 1960, in response to rumors that a U.S. invasion of Cuba was imminent (Khrushchev 2000, p. 482; Taubman 2003, p. 534). The U.S. took little notice of these statements since no invasion was planned and any possible invasion was months off (Fursenko and Naftali 1997, p. 52). The arms package, which was negotiated while Raul was touring Europe, included shipments of rifles, tanks and planes (Khrushchev 1970 p.

491). By April 15, 1961, the first day of the Bay of Pigs operation, the Cubans had received 125 of 205 expected tanks, 167,000 rifles, 7,250 machine guns, and much of the promised anti-aircraft and anti-tank weapons (Fursenko and Naftali 1997, p. 92). The Cubans had not received promised artillery and MiG aircraft, but the T-33 jets inherited from the Batista regime were sufficient to repel the invading force.

The persistent rumors of invasion in 1960, in retrospect, were probably related to the covert operations approved by the Eisenhower administration (Freedman 2000, pp. 123-26). Soon after Mikoyan's visit to Havana, U.S. policymakers recognized that they were losing Cuba to the Soviet sphere. Eisenhower approved a number of CIA initiatives that involved the training and supplying of Cuban dissidents, the largest of which was a training facility in Guatemala. The CIA also devised a plan in which Cuban dissidents trained and supplied by the U.S. would invade Cuba and remove the communists from power. The training and preparation would take a number of months, and there were no plans for any imminent invasion. This plan did mature into the Bay of Pigs invasion in 1961, which helped set the stage for the missile crisis (Fursenko and Naftali 1997, p. 92).

Prior to the 1961 invasion, Soviet support of Cuba had been defensive. The Soviets had expressed their willingness to secure Cuba against U.S. aggression, and provided materials to help protect the island (Fursenko and Naftali 1997, p. 87). Intelligence activities showed that the Soviets were still supplying military aid to Cuba, and the U.S. was suspicious of the military buildup (Stern 2005, p. 20). Despite these suspicions, the consensus opinion reflected the belief that this aid was defensive in nature. This position was expressed in a National Intelligence Estimate (NIE) produced in September 1962. Stressing the defensive character of the buildup, "the Soviets evidently

hope to deter any such [U.S. invasion] attempt by enhancing Castro's defensive capabilities and by threatening Soviet military retaliation" (Special National Intelligence Estimate: The Military Buildup in Cuba 1962; May and Zelikow 2001 p. 128).

This sentiment was echoed in a policy paper drafted by Walter Rostow then Director of Policy Planning at the State Department. Rostow's memorandum, dated September 3, 1962, noted that "On the basis of existing intelligence the Soviet military deliveries to Cuba do not constitute a substantial threat to U.S. security" (Rostow 1962). He did warn that the buildup may "constitute a deterrent to certain types of surveillance and a means for improving certain types of Soviet intelligence" (Rostow 1962).

Despite the apparent consensus there was some recognition that the Soviet buildup might extend beyond defensive capabilities. The NIE of September warned "The USSR could derive considerable military advantage from the establishment of Soviet medium and intermediate range ballistic missiles in Cuba, or from the establishment of a Soviet submarine base there." While "the establishment of a submarine base would be the more likely," it concluded that such moves "would be incompatible with Soviet practice to date and with Soviet policy as we presently estimate it" (Special National Intelligence Estimate: The Military Buildup in Cuba 1962). The intelligence community believed the Soviets would not take such big risks in the Caribbean.

By August, CIA Director John McCone had become a dissenting voice. After observing advanced surface-to-air (SAM) missile sites (SA-2s) capable of destroying a U2 reconnaissance plane, McCone believed that the advanced missile units were intended to deter over flights of nuclear weapon facilities (Freedman 2000, pp. 163-64). He reasoned that the Soviets could justify the placement of missiles in Cuba by referencing

U.S. missile facilities in Turkey and Italy. Despite his conviction, the intelligence community concurred with senior policymakers in their belief that Soviets would not take such risks. This is evidenced by the content of the September NIE.

Policymakers wanted to ensure that the Soviets would not attempt placing offensive weapons in Cuba. Rostow, in his September memo, advocated drawing a line to clarify acceptable Soviet behavior in Cuba and specified unacceptable action. “In general, that line should be drawn at the installation in Cuba or in Cuban waters of nuclear weapons or delivery systems, sea or land based. There may be other types of aggressive instruments that we would wish to include in this definition” (Rostow 1962). The military buildup, he reasoned, did not represent a new threat to the U.S. so much as the movement of Cuba towards communism, but offensive weapons could not be tolerated (May and Zelikow 2001, pp. 4-8). Days after Rostow’s memo circulated, this became the official position. On September 11, Kennedy warned that the U.S. would not tolerate offensive weapons in Cuba or the use of Cuba for aggressive purposes.

Transcript evidence suggests that policymakers abandoned this prior consensus assessment quickly on October 16, 1962. Upon seeing the photos, and listening to the explanations provided by the aerial reconnaissance analysts, the ExComm immediately changed their minds (May and Zelikow 2001 pp. 32-41; Stern 2005, p. 46). After getting past the initial shock that the Soviets had placed offensive weapons in Cuba, the ExComm discussions focused on possible Soviet motives. Why the Soviets chose to challenge the status quo and take the risk of putting nuclear weapons in Cuba has plagued historians and political scientists for some time (Allison 1971, 1999; Blight, Allyn and Welch 2000; Haas 2001). Kennedy famously called the incident “one hell of a gamble,”

which is also the title of a book documenting Soviet events leading up to and after the missile crisis (Fursenko and Naftali 1997). At first glance, the Soviet action seems needlessly dangerous and possibly irrational, however, a number of explanations now exist for the missile placement.

Allison (1971, 1999) identifies four possible justifications, or hypotheses, for the Soviets' placement of missiles in Cuba: 1) Cuban defense, 2) Cold War politics, 3) missile power, and 4) Berlin. To this list, more recent scholarship suggests that Khrushchev was concerned about divisions within Cuban leadership and the possibility that Cuba might ideologically or strategically align with the Chinese (Fursenko and Naftali 1997, p. 167; Taubman 2003, p. 534). The Cuban defense explanation, perhaps the most straightforward, explained Soviet action as a product of concerns that the U.S. would seek to remove Castro from power militarily (Khrushchev 1974, p. 512; Taubman 2003, p. 537). The missiles, in this explanation, helped ensure that the island would not be an attractive invasion target. By Cold War politics, Allison was referring to the limited military or diplomatic probes that both the U.S. and the Soviets used to test their opponents' resolve and signal discontent with the status quo in a given arena (Allison 1999, pp. 88-89). Missile power, which is related to Cold War politics, addressed the relative weakness in Soviet strategic forces. The Soviet intercontinental ballistic missile (ICBM) program had developed slower than expected and the Soviets found themselves falling behind the U.S. in strategic capability (Khrushchev 2000, p. 484, p. 484; Taubman 2003, p. 537; Scott 2007, p. 28). Placing medium and intermediate range missiles (MRBMs and IRBMs, respectively) in Cuba would vastly improve Soviet strategic capability. The Berlin hypothesis, which dominated U.S. thoughts during the crisis,

explained the missile placement as an attempt to trade or extort concessions in Berlin for some in Cuba (Allison 1999, p. 99).

Given these five plausible explanations, what drove the Soviets to place the missiles in Cuba? Complex decisions are usually motivated or justified by a complex set of reasons (Vaughan 1996). It would be wrong to eliminate any of these altogether, or focus on any single explanation for the missile placement. Instead, careful analysis helps to distinguish or rank some of these explanations as more crucial than others. After which, we will address how U.S. beliefs may have impacted the Soviet decision.

Initial reports on Soviet decision-making, drawn from Khrushchev's own memoirs, suggested that Khrushchev made the decision on his own during a trip to Bulgaria in March 1962 (Khrushchev 1970, p. 493; Fursenko and Naftali 1997, pp. 177-78). More recent archival work suggests that the decision, and the decision-making process, was more complicated than the initial reports (Khrushchev 2000, p. 487). Khrushchev first considered the possibility of adding nuclear weapons to the Cuban arms package in April 1962 during a conversation with Rodion Malinovsky, the Soviet Defense Minister. Khrushchev wanted to explore the possibility of putting "hedgehogs down the Americans' trousers" (Volkogonov 1998, p. 236). By hedgehogs, Khrushchev meant nuclear weapons, and trousers referred to the Caribbean. He instructed Malinovsky to form a small group to explore the feasibility (Fursenko and Naftali 2006, p. 431). This account is supported by the Defense Ministry's official history, which claims that some staff work had to have been prior to Khrushchev's trip to Bulgaria (Volkogonov 1998, p. 235).

Khrushchev's decision to investigate the possibility of nuclear weapons in Cuba seems to have been motivated by a number of factors. In April 1962, the U.S. deployed the first 18 of 54 Titan missiles in Colorado. By contrast, the Soviets were not in a position to counter. The Soviet R-9 and R-16 rockets had a number of problems (Fursenko and Naftali 2006, p. 429). The fuel used for the R-16 rocket was volatile and highly corrosive, meaning that it had to be fueled immediately prior to use and drained if not used. The R-9 did not fair much better as flight tests revealed a host of flaws (Fursenko and Naftali 2006, p. 430). The weaknesses of the Soviet rockets were not for lack of effort or resources. The Soviets had invested heavily in their rocket programs, and they were supposed to be a key element in Khrushchev's strategy. This poor state of affairs probably made a nuclear base in Cuba seem like a low cost solution to the Soviet's strategic inferiority.

The rocket problems, and U.S. Titan deployment, were not the only issues weighing on Khrushchev in April and May. Events in Asia also weighed on the Soviet leader (Fursenko and Naftali 2006, pp. 433-435). Khrushchev had hoped that Asia would present opportunities, particularly after his inability to force changes in the Berlin situation in 1961. In May 1962, the Pathet Lao, the Chinese and the Vietnamese routed pro-U.S. forces in Nam Tha, Laos (Freedman 2000, p. 347; Beschloss 199, p. 231). To limit the losses and discourage the forces from advancing further, Kennedy decided to send a military force to Thailand. The force comprised 1,800 marines and two air squadrons. Kennedy wished to avoid conflict in Laos and hoped that the presence of the military force would be sufficient to deter Communist rivals. The show of strength was successful, but also had the effect of stirring Khrushchev (Fursenko and Naftali 2006, pp.

431-434; Khrushchev 1970, p. 493). Khrushchev was concerned that the U.S. show of strength was only the first step in a larger intervention, which would once again show that the Soviets suffered from an unfavorable balance of power. To deter possible U.S. involvement, he sent a private message to Kennedy reaffirming Soviet commitment to a neutral Laos ensuring that Communist forces would not advance further (Fursenko and Naftali 2006, p. 434).

While rocket problems and Asia were of immediate concern, there were additional issues motivating Khrushchev. Perhaps the most important was Cuban defense (Khrushchev 1970, p. 492; Khrushchev 1974, p. 511; Khrushchev 2000, p. 482; Taubman 2003, p. 534). Just prior to the Bulgarian trip, Khrushchev had signed a letter to Castro that outlined the terms of the Soviet military aid. The 1961 Bay of Pigs invasion caught the Soviets by surprise. All Khrushchev could do at the time was wait and hope that the Cuban forces would repel the invaders. While Khrushchev had rebuffed earlier Cuban aid requests, he was now prepared to offer Soviet assistance (Khrushchev 1970, p. 493).

Castro had been pushing hard for Soviet aid since he had been warned that the U.S. might try to dislodge his regime again. Khrushchev claims to have received intelligence suggesting an upcoming U.S. operation (Khrushchev 1974, p. 510), but it has also been suggested that there was no intelligence suggesting that a U.S. operation was imminent and actually believed that Cuba was safe for the time being (Fursenko and Naftali 2006, pp. 159-160). He doubted that the U.S. would take any overt action against Cuba in 1962, but he was concerned that Kennedy might be motivated to try again ahead of the 1964 presidential election. Given that Khrushchev was planning to supply Castro with much of the military aid requested, Khrushchev thought it prudent to offer the aid

early in order to assuage Castro's concerns (Taubman 2003, p. 534; Fursenko and Naftali 2006, 427).

Recent events in Cuba's domestic politics also seem to have influenced Khrushchev. In 1962, Castro purged longtime communist Anibal Esclante from the government (Fursenko and Naftali 1997, p. 136). Escalante, unlike Castro and many of the other revolutionaries, was well acquainted with Moscow and a known entity within the Kremlin (Fursenko and Naftali 1997, p. 59; Taubman 2003, p. 534). The purging of Escalante meant that Castro's inner circle was almost exclusively revolutionaries rather than communists, and this fueled concerns that the young revolutionaries may find China a more attractive ideological partner than the Soviet Union (Fursenko and Naftali 1997, 9. 167). While the Chinese lacked the resources to provide comparable aid, Mao Zedong, the Chinese leader and admirer of Stalin, felt that the Soviets had left the revolutionary path. The Sino-Soviet relationship had become strained ever since Khrushchev publicly denounced Joseph Stalin and his brutal methods (Gaddis 2006, p. 141). Khrushchev, therefore, believed it important to deepen his relationship with Cuba in the hopes of ensuring that the Soviet Union, rather than China, would be the Cuban's strategic partner.

A final issue, one that bothered Khrushchev considerably, was the existence of U.S. nuclear missile bases in Italy and Turkey (Khrushchev 1974, p. 512; Khrushchev 2000, p. 484). The Turkish missile deal was negotiated by the Eisenhower Administration and was part of a public defense pact. Eisenhower expected that the deal would upset the Soviets, but he, and Kennedy, chose to go ahead despite concerns that the Soviets might respond aggressively (Stern 2005, p. 14). During the trip to Bulgaria, Khrushchev came to the opinion that "the Americans had surrounded our country with military bases," and

he wanted to give them "a little of their own medicine" (Khrushchev 2000, p. 484). Nuclear weapons in Cuba was a tit-for-tat strategy according to Khrushchev's logic.

Khrushchev may have decided to put nuclear weapons in Cuba during his trip to Bulgaria, but he still needed to present the idea to and get approval from the Presidium in Moscow. Khrushchev presented the plan to the Presidium on May 21, less than 24 hours after his return from Bulgaria. He did not get the response he had hoped for. There was a good deal of dissension, particularly from Mikoyan, and a note taker at the meeting reported that debate went on for quite some time (Khrushchev 2000, p. 487; Taubman 2003, p. 541). Foreign Minister Gromyko joined Mikoyan in his reservations about the plan (Taubman 2003, p. 543). Those that spoke against the plan worried that such a provocation would incite U.S. action and all but ensure that the U.S. would use force to remove Castro (Khrushchev 2000, p. 485). Fearing that his plan would not get sufficient support, Khrushchev ended the meeting before there was a resolution or vote.

There were two strategies available to the Soviets should they attempt to put a missile base in Cuba. They could adopt a strategy similar to the one that the U.S. used when placing missiles in Turkey. The U.S. publicly announced a defense pact in conjunction with their decision to place nuclear missiles there (Beschloss 1991 p. 65; Scott 2007, p. 35). Khrushchev, and Soviet leadership more broadly, feared that the U.S. would try to prevent such a deployment. Given U.S. superiority in the Caribbean, the U.S. would almost certainly be able to block the Soviet move. Khrushchev was confident that the Soviets could get the missiles into Cuba without being detected by the U.S., and believed that secrecy was crucial (Khrushchev 2000, p. 484; Taubman 2003, p. 535). He spent the next three days building support, and the Presidium met to discuss the proposal

again on May 24. This time, no one voiced opposition and the plan passed unanimously (Taubman 2003, p. 544).

The dissention in the Presidium showed that Khrushchev's vision of Cuba, and his attempt to unsettle the U.S., did not necessarily share unanimous support (Taubman 2003, p. 541). This tells a very different story then accounts of Khrushchev's unilateral decision-making, or the possibility that Khrushchev was motivated by military leaders, a possibility discussed at length during the ensuing ExComm meetings (May and Zelikow 2001 p. 67). Instead, this tells the story of deliberative leadership concerned with U.S. response and the beliefs that might motivate different responses (Khrushchev 2000, p. 485; Taubman 2003, p. 535). Khrushchev's belief that surprise would be effective, relied on the assumption that the U.S. believed that Soviet assistance to Cuba was defensive in nature. Had consensus opinion in the U.S. formed around the belief that the Soviets aid to Cuba was offensive or expansionist, any attempt to surprise the U.S. would have a reduced possibility of success.

The missile deployment was accompanied by an active deception plan. In repeated diplomatic exchanges, the Soviets signaled that they had no intention of putting offensive or nuclear weapons in Cuba. In early September, Soviet Foreign Minister Andrei Gromyko told Robert Kennedy that all weapons shipments to Cuba were defensive in nature. On September 4, in part due to Senator Keatings' charge that there were nuclear weapons in Cuba, Kennedy publicly stressed that offensive weapons in Cuba would not be tolerated (Stern 2005, p. 20). The Soviet news agency TASS reported that all the weapons in Cuba were defensive on September 9, echoing earlier signals (May and Zelikow 2001, p. 69). While the Soviets were going out of their way to avoid

signaling an immediate threat, the first shipment of nuclear weapons, unbeknownst to U.S. policymakers, was on its way to Cuba and arrived on October 4, 1962 (Scott and Smith 1994, p. 674, Fursenko and Naftali 2006, p. 461).

Back in Washington on October 16, the ExComm was trying to figure out why the Soviets had chosen to challenge the status quo. They discussed many of the issues raised by Allison (1971, 1999; May and Zelikow, 2001, pp. 54-71) including Cuban defense, Cold War politics, missile power and Soviet attempts to make aggressive moves elsewhere. There was little concern that the Soviets had any intention of using these weapons in offensive fashion against the U.S. Instead, there was widespread concern that missiles could be offensively leveraged to buttress Soviet expansionism in Berlin or Latin America (May and Zelikow 2001, p. 42; Leaming 2006, p. 401). Leveraging the missiles became synonymous with offensive, expansionist, or aggressive intentions. While the ExComm considered possible Soviet motivations, they were not weighted equal. Content analysis of the discussion on October 16 (May and Zelikow 2001, pp. 32-72) reveals that Berlin was mentioned 18 times, U.S. missiles in Turkey were mentioned 7 times and Cuban defense only mentioned 4 times. Policymakers were clearly concerned that the missiles could be used to challenge the U.S. on other fronts.

After seeing the pictures of the missiles sites and discussing possible Soviet strategic intentions, the ExComm spent subsequent days discussing the appropriate policy response (May and Zelikow 2001, p. 75). Broadly speaking, there were four alternatives: do nothing, diplomatic confrontation, an air strike against Cuba, or an invasion of Cuba. Either air strikes or invasion would be accompanied by a blockade to ensure that the Soviets could not rearm the Cuban forces. The first and the last options were reasonably

straightforward, but the air strike option was more complicated. Air strikes could be limited to the missile sites or expanded to include anti-aircraft weapons, planes and other military sites (May and Zelikow 2001 p. 47; O'Brien 2005, p. 659). This meant that air strikes could be comprehensive or targeted. The Joint Chiefs continually advocated for the comprehensive strike, which had the best chance of removing the immediate threat and weakened the Cuban forces such that future air strikes were low risk for U.S. pilots. The Joint Chiefs also suggested following the comprehensive air strikes with an invasion of the island in the following weeks (Freedman 2000, p. 175; May and Zelikow 2001 pp. 47-49).

After the initial meeting, Kennedy was never comfortable with the comprehensive strike solution (May and Zelikow 2001; Stern 2005; O'Brien 2005, p. 661). For one, the strike went way beyond the immediate threat to the U.S. A second concern was that Soviet casualties would be relatively high in the comprehensive strike (May and Zelikow 2001, p. 86). He was also concerned that such overwhelming force against a weaker adversary would send the wrong message to the European allies that had been living with risks of Soviet nuclear forces for years (May and Zelikow 2001, p. 100). The Joint Chiefs further undermined their cause by constantly expanding the scope of the air strikes and the number of sorties that the pilots would fly to cover all of the targets (Freedman 2000, pp. 191-192).

Given the mix of problems associated with the comprehensive air strike, it is not surprising that Kennedy was not enthusiastic about that strategy. Early on in the meetings, Kennedy and many of the ExComm members favored a limited air strike (May and Zelikow 2001, p 50; O'Brien 2005, p. 659). The strike would eliminate most of the

missiles, while inflicting as little damage on Soviet and Cuban forces as possible. After a thorough assessment, Chairman of the Joint Chiefs General Maxwell Taylor told Kennedy that an air strike would probably remove 90% of the threat (Stern 2005, p. 46, 56). This meant that 10% of the known offensive weapons in Cuba could be fired at the U.S. after the strike. This was hardly reassuring. The primary opponent to this plan was Secretary of State Dean Rusk who favored diplomatic confrontation to military options (May and Zelikow 2001, p. 36, 88). The problem, however, was that any confrontation would eliminate the possibility of catching the Cubans by surprise and would provide motivation to speed up work on the missiles sites.

Over the next few days, a consensus formed against any air strike. McNamara cited some of the disadvantages in his October 20 brief (May and Zelikow 2001). The ExComm members were concerned that an air strike would be analogous to the attack on Pearl Harbor, which was the biggest surprise in U.S. history (May and Zelikow 2001, p. 79; Stern 2005, p. 53). They were loath to set a precedent that such attacks were acceptable, and concerned about a Soviet response against U.S. interests, particularly Berlin (Freedman 2000, pp. 188-190). As air strikes and invasion lost appeal, the ExComm slowly settled on a blockade of Cuba.

A blockade of Cuba, technically an act of war, had both positive and negative aspects (Freedman 2000, pp. 182-192). The blockade would be clear response to Soviet activity, and was preferable to doing nothing. It had the advantage of minimizing the likelihood that Soviet and U.S. forces ended up in a hostile engagement, kept the Soviets from adding any further offensive weapons to the island, would keep the Soviets from restocking if an air strike or invasion were necessary, and it could be publicly announced

making accompanying diplomacy possible. The blockade would not, however, neutralize or remove the current missiles from Cuba (Stern 2005, p. 64). McNamara's biggest concern on October 20 was that the blockade would "take a long time to achieve the objective of eliminating strategic missiles from Cuba" (May and Zelikow 2001, p. 128)."

In a televised speech on October 22, Kennedy told America of the offensive weapons in Cuba, announced the blockade of the island, and warned the Soviets that the offensive weapons would not be tolerated (May and Zelikow 2001, pp. 183-189). About four hours before the speech, Soviet intelligence reported an unusual flurry of activity around the White House (Khrushchev 2000, p. 553; Fursenko and Naftali 2006, p. 467). Khrushchev quickly connected the dots fearing that the U.S. had discovered the covert Cuban arms operation and called for a Presidium meeting late that night. These concerns were aggravated further when the U.S. Embassy warned the Kremlin that Kennedy would make a speech, and informed them that they would get an advance copy an hour before it aired. Khrushchev feared that the U.S. had discovered the missiles and would use them as a pretext to invade Cuba (Fursenko and Naftali 2006, p. 470). Malinovsky was more optimistic (Khrushchev 2000, p. 553). He believed that the U.S. position did not reflect the operational capability to carry out the operation, but believed that the U.S. would take some action (Fursenko and Naftali 2006, p. 468). The conversation focused on the worst-case scenario: what to do if the U.S. chose to invade Cuba.

The fears of an imminent invasion were slightly allayed when they received a copy of Kennedy's speech (Khrushchev 2000, p. 559), but they wanted a strategy to ensure Cuba's continuing safety. The Presidium discussed turning nuclear weapons authorization offer to field commanders and even the Cubans. Mikoyan advised caution,

and pointed out that Cuban control of nuclear weapons was likely to magnify U.S. fears (Fursenko and Naftali 2006, p. 472). The Presidium agreed and also chose to maintain authority over the strategic weapons, but gave authorization for tactical nuclear weapons to the field commander (Taubman 2003, p. 562). The order was sent on to the field commanders just before Kennedy addressed America. Cuba appeared safe for the moment, and the Presidium adjourned.

When the Soviets met at the Kremlin on the morning of October 23, the conversation had a different tone. The Soviets still wanted to ensure Cuba's safety, but they also wanted to keep the crisis from escalating (Fursenko and Naftali 2006, p. 477). Any provocative act at this delicate juncture, particularly against U.S. forces in the Caribbean, risked starting nuclear war. The Soviets immediately recalled weapons supply ships in the Mediterranean and advised ships in the Atlantic to stop (Fursenko and Naftali 2006, p. 477). Only a few ships close to Cuba were instructed to continue on in the hopes of arriving in Cuba before the blockade. They drafted a public statement denying the existence of the missiles and appealed to the U.N. on Cuba's behalf (Fursenko and Naftali 2006, p. 477). Khrushchev decided to take a different tone in his personal correspondence to Kennedy stressing that the weapons had solely defensive purposes (May and Zelikow 2001, p. 203). The Soviets also deliberated the fate of four Foxtrot class diesel submarines in route to Cuba (Fursenko and Naftali 2006, p. 478). Malinovsky argued that the submarines could arrive in Cuba undetected, against the recommendation of Mikoyan. Eventually, the Presidium sided with Mikoyan, agreeing that the submarines would not be able to reach Cuba undetected given U.S. anti-submarine capabilities in the area (Fursenko and Naftali 2006, p. 479). For some reason, the submarines were never

issued orders to stop, setting the stage for a tense standoff during the blockade (Fursenko and Naftali 2006, p. 480).

Curiously, or perhaps conspicuously, missing from these Soviet debates was any serious discussion about action against Berlin. Sergei Khrushchev (2000, p. 560) notes that Vasily Vasilyevich Kuznetsov proposed exerting pressure on Berlin, and Khrushchev reportedly responded with surprising sharpness, "Keep that kind of advice to yourself… We don't know how to get out of one predicament and you drag us into another." Fears of Soviet retaliation against Berlin, for blockade or strike, had dominated U.S. deliberations from October 16 to 22. Between October 23 and 24, the Soviets never discuss retaliating with a blockade of Berlin, focusing on ensuring Cuba's defense while deescalating the crisis (Fursenko and Naftali 2006, p. 473). The Soviets believed that the U.S. would look for any reason to invade Cuba, and believed the best hope of keeping Cuba safe and avoiding war was caution and prudence rather than retaliation.

Before examining the Bayesian and biased models as applied to the Cuban Missile Crisis, it is important trace the beliefs of the target, the U.S., at key junctures in the crisis. The analysis focuses on the prior beliefs, the evidence that becomes available over time, the impact of that evidence on beliefs in the hopes of identifying the existence of entrenched beliefs, the issues on which beliefs become entrenched, and the strength of the evidence used to identify entrenched beliefs. Tracing beliefs as a process helps to identify instances where new evidence should have been sufficient to alter beliefs, potentially isolating situations where the decision-makers may have exhibited bias.

It is difficult to argue that U.S. policymakers suffered from any type of entrenched beliefs at the outset of the crisis, in fact, they changed their beliefs 180

degrees in an hour. Prior to October 16, we have argued that U.S. decision-makers believed that Soviet aid to Cuba was defensive in nature. The aerial reconnaissance photos showing Soviet medium range missiles in Cuba was sufficient to cause a major change in U.S. beliefs. After seeing the photos, the ExComm came to believe that Soviet military aid was offensive or expansionist in nature aimed at gaining leverage to alter the status quo in Berlin and South America (May and Zelikow 2001, p. 67-68). There is no evidence to suggest that U.S. decision-makers clung to their prior belief that the aid was defensive, and no evidence of entrenched beliefs during this initial stage of the crisis.

After the initial crisis stage, where the decision-makers first learn of the missiles, there is some evidence that they anchor on a couple of beliefs throughout the crisis. The strongest evidence of entrenched beliefs comes from the assessment of Soviet capabilities already in Cuba. The initial intelligence assessment on October 16 revealed the existence of four missile MRBM missile launchers at three sites (Freedman 2000, p. 1999). Analysts were unable to find any nuclear storage facilities or any evidence of nuclear warheads at these sites, but they did warn that the missiles would have little use without nuclear warheads (May and Zelikow 2001, p. 34). The assessment that there were no nuclear ordinances in Cuba becomes a vital component of the early ExComm discussion. On October 17, McNamara articulates support for two of five military options: an air strike aimed at all the offensive weapons and an invasion. He also warns that "All of these plans are based on one very important assumption: That we would attack with conventional weapons, against and enemy who is not equipped with operational nuclear weapons. If there's any possibility that the enemy is equipped with operational nuclear weapons, I'm certain the plans would have to be changed" (May and Zelikow 2001, 82).

Over the next few days, the intelligence assessments of Soviet missile capability continued to change (Freedman 2000, p. 1999). On October 17, the assessment was increased to eight MRBM launchers, in two to three sites, with 16-32 missiles operational in one week's time. The assessment was refined on October 18, maintaining the existence of eight MRBM launchers in two sites with 16 missiles. On October 19, the assessment of capability was expanded to include two IRBM launchers in two sites in addition to the MRBM sites. The hardened IRBM sites would not be ready until December, but the discovery of IRBMs was a source of tension for the ExComm. The assessment was expanded again on October 20: 16 MRBM launchers with 32 missiles in four sites, along with the eight IRBM launchers with 16 missiles in two sites. The other addition to the October 20 assessment was the possible sighting of a nuclear storage facility under construction at an IRBM site (May and Zelikow 2001, p. 126). The assessment indicated that the facility would not be completed until December along with IRBM site. The October 22 assessment expanded the MRBM force to 23 launchers with 33 missiles in six sites (Freedman 2000, p. 199), and also indicated that there could be one to three nuclear storage facilities, probably under construction (May and Zelikow 2001, p. 166). The assessment also warned that aerial reconnaissance might never be able to identify storage facilities or warheads with certainty.

Taken together, this means that the October 16 and 17 assessments were changed incrementally to reflect almost four times as many MRBM launchers, almost eight times as many MRBM missiles in twice as many sites, along with a previously unobserved sizeable IRBM force, and increase in possible nuclear storage facilities from zero to three. Though confidence in observing nuclear storage facilities was not high, the change

in estimates over six days might have triggered a major reassessment of Soviet commitment to Cuba. While the ExComm did consider the Soviet intentions, the military planning seemed static in light of the new evidence. The Joint Chiefs continued to advocate an invasion without considering the possibility that U.S. forces might face tactical nuclear weapons (Gaddis 2005, p. 77). They did revise the air strike option, increasing the number of sorties as the number of weapons and sites increased, but never assumed that the revised assessments were indicative of a larger error. This could reflect a difference of opinion between McNamara, who suggested that military options be reevaluated if assessment of nuclear capabilities changed, and the Joint Chiefs, but it reflects a fundamental inconsistency since neither actor ever explicitly revisited this core assumption about Soviet capabilities later in the crisis.

McNamara and the Joint Chiefs were not the only parties to fall prey to this bias. On October 20, the ExComm met to discuss the latest intelligence assessment and policy recommendations. There is no transcript of the meeting since it took place in the White House residence under the auspicious of Kennedy's false illness that brought him back from Chicago early (May and Zelikow 2001, p. 125). During that meeting, each participant discussed their position and offered a brief justification. Both Director McCone, the skeptic, and Undersecretary of State George Ball offered policy assessments relying on the assumption that there were no nuclear weapons in Cuba. Ball argued that "if an effective blockade was established, it was possible that our photographic intelligence would reveal that there were no nuclear warheads in Cuba" (May and Zelikow 2001, p. 129). McCone reinforced the prevailing view later in the discussion acknowledging, "that we do not know positively that nuclear warheads for missiles

deployed had actually arrived in Cuba. Although we had evidence of the construction of storage places for nuclear weapons, such weapons may not yet have been sent to Cuba" (May and Zelikow 2001, p. 133).

McCone's statement is telling because it is indicative of the prevailing attitude across the national security policymakers: in the absence of clear evidence, they should operate under the assumption that there might be a limited nuclear capability in Cuba. This perspective seemed to have solidified despite mounting evidence that the U.S. had fundamentally miscalculated and perhaps even underestimated the Soviet investment in Cuba.

The assessment turned out to be horribly wrong. At the time of the US blockade, there were at least 36 nuclear warheads for the MRBMs, 12 Luna tactical nuclear weapons (Scott 2007, p. 41) and 42 FKR cruise missiles with nuclear warheads (Fursenko and Naftali 2006, p. 469). The Luna missiles had a range of 31 miles and a 2-kiliton yield, which would irradiate everything within 1000 yards of detonation. The FKR cruise missiles had a 100-mile range and a yield that varied from five to 12 kilotons. Both weapons, particularly the cruise missiles pointed out to sea, were intended to destroy naval vessels (Fursenko and Naftali 2006, p. 473). The largest estimate of nuclear armament at the time of the blockade have put the number of nuclear weapons as high as 162: 60 for MRBM and IRBM missiles, 12 FROG-7 warheads, 80 tactical warheads for cruise missiles, 6 gravity bombs, and 4 naval mines (Scott and Smith 1994, p. 674).

McNamara had initially supported air strikes and invasion early on under the assumption that there were no nuclear weapons in Cuba (May and Zelikow 2001, p 84). Had the U.S. known the true status of Soviet nuclear capabilities on the island, it would

have been hard to justify taking military action. The presence of nuclear weapons on the island diminished value of all military options including the invasion, air strike and blockade. The utility of the blockade, since it was executed at some distance from the island, would have been negatively impacted the least. At a radius of 500 miles, the ships should have been just out of the reach of the FKR cruise missiles. None of the publicly available discussions between Kennedy and the military leadership reflect concern about Soviet tactical nuclear weapons (May and Zelikow 2001). Had the military leadership seriously considered the possibility of tactical nuclear weapons in Cuba, it would have been much more difficult to justify an invasion.

The value of the blockade relative to other options was also bolstered by the inaccurate assumption that there were probably no nuclear weapons on the island. If there were no nuclear weapons for the missiles, a blockade would keep the offensive weapons from becoming fully operational. The most the Soviets could do, under these assumptions, was fire conventional rockets. If there were no weapons, the blockade was both a signal of resolve and defensive tactical measure. If the nuclear warheads were already in Cuba, its value as a tactical defensive measure dropped since it would only keep the Soviets from building up the arsenal further or restocking it if the existing nuclear weapons were used or destroyed. An accurate assessment would have made the blockade seem more valuable than the other military options, but it would not have prevented nuclear warheads from entering Cuba and would have been viewed as a risk given the tactical nuclear weapon stockpile. The inaccurate U.S. belief helped to justify the blockade in the hopes that it was the lower risk option that might keep nuclear weapons from reaching the island.

When one examines the drastic changes in the intelligence assessment of Soviet strategic capability over six days, it is hard to imagine not revisiting the basic assumptions that the ExComm formed on October 16. Even if the U.S. was operating under the belief that there were MRBM warheads (Scott and Smith 1994, p. 674), the Soviets weapon shipments to Cuba far exceeded the initial estimates. Despite this obvious fact, no one clearly revisited the assumptions on the disposition of strategic nuclear warheads or tactical nuclear weapons. This belief was both convenient, and probably helped facilitate decision-making, but it was at odds with new information and inaccurate. Tracing the beliefs and available information reveals that the decision-makers probably exhibited some sort of anchoring or entrenchment on their initial beliefs developed on October 16. While it appears that there were no entrenched beliefs prior to October 16, it is more difficult to argue that the decision-makers displayed rationality thereafter.

There is one other area where U.S. decision-makers may have displayed entrenched beliefs, but this argument is not as strong as the one above. After the U.S. decision-makers viewed the missile photos on October 16, they quickly abandoned their view that the missiles were defensive for the belief that the missiles were offensive or expansionist in nature (Leaming 2006, p. 401). Even though the ExComm did not believe that the missiles were going to be used against the U.S. in the immediate future, they assumed that missile would probably be used as leverage to expand Soviet influence in Berlin or Latin America (May and Zelikow 2001, pp. 90-91). Lawrence Freedman (2000, p. 171) writes that U.S. decision-makers suffered from a common bias "to think that with offensive missiles came a more offensive purpose." He goes on to note that offensive

weapons are often used in defensive campaigns, just as defensive weapons are crucial in offensive campaigns. The bias Freedman is talking about is somewhat different from entrenched beliefs in that entrenched beliefs develop over time. The belief that the missiles are intended for offensive or expansionists purposes only represents an entrenched belief if it hardens over time in the absence of any new evidence or the presence of new conflicting information.

The weakness of this second claim stems, in large part, from fewer incremental signals during the crisis. Fewer incremental signals or cues make it more difficult to establish that a belief was irrationally entrenched. The best, and perhaps only, way to establish this claim is tracing the sentiments offered during the ExComm deliberations. The index of Kennedy transcripts edited by Ernest May and Philip Zelikow (2001, p. 506) notes six instances where the ExComm discuss the "offensive vs. defensive" nature of the Cuban missiles. In the first instance, on October 16, the participants discuss a range of possibilities including Cuba's defense, U.S. missiles in Turkey, and particularly Soviet desires in Berlin (May and Zelikow 2001, p. 61). The next two indexed conversations take place on October 18. In the first instance, the U.S. discusses the possibility that Soviets might want to trade to get U.S. missiles out of Turkey (May and Zelikow 2001, p.100). Soviet desires in Berlin then dominated the second conversation on October 18 (May and Zelikow 2001, p. 106). There are also two conversations indexed on October 19. The first focused on attempted Soviet blackmail in Berlin and Latin America, the second focused on Berlin ((May and Zelikow 2001, p. 117). There is a brief discussion of defensive rationale, but Kennedy negates this responding, "Of course, that's how they define these weapons, as defensive" (May and Zelikow 2001, p. 120). The final indexed

instance is on October 22, which again focuses on Berlin (May and Zelikow 2001, p. 159).

The raw count reveals that Berlin, and thereby offensive or expansionist intentions, dominated the discussions. During these six conversations, blackmail over Berlin is cited five times, Turkish missiles and strategic capability twice, blackmail in Latin America once, and the defense of Cuba once. What is more telling is the trend over time. On the first day, they consider Berlin, Turkish missiles and defense of Cuba. On the second day they focus on Berlin and Turkish missiles. On the third day they focus on Berlin and blackmail in Latin America. Then three days later they only discuss Berlin. The trend over time shows that the participants found themselves leaning increasingly toward offensive rather than defensive rationales, despite the fact there was no clear or new evidence of Soviet intentions over this time. The belief becomes entrenched despite the lack of tangible intelligence on Soviet motivation visa vie Cuba.

The entrenchment seems based on two common decision-making phenomena. First, U.S. policymakers seemed to engage in mirror imaging, which is the tendency to assume that others know what you do or think the same way about the same set of facts (Jervis 1976, p. 354). They knew there were no plans to invade Cuba and assumed that the Soviets had the same information. Policymakers never seriously considered that the aggressive U.S. stance toward Cuba was responsible for Soviet concerns about Cuban defense (O'Brien 2005, p. 658). The Soviet reaction to Kennedy's speech tells a different story. Their first priority was avoiding war, followed by defending Cuba. There were no discussions about retaliating against Berlin. The second effect that seems present is a groupthink or echo effect whereby small group concerns or beliefs bounce off one

another and amplify over time (Vertzberger 1997, p. 303). This process often leads groups toward taking larger risks than they would if making individual decisions. It is worth noting that the tangible intelligence available during the crisis was the actual force size in Cuba, and the estimate changes showing a larger and larger force almost certainly fueled increased concern about Soviet motives. This might justify the U.S. position, but it is important to recognize that the U.S. position about Soviet motives hardened over time.

There is a reasonably strong argument that U.S. beliefs regarding nuclear ordinances in Cuba reflected entrenchment as evidence mounted that the Soviet aid package was much more substantial than initially believed. The belief that the Soviet missiles had offensive or expansionist purpose could be rationally justified by the intelligence changes, but could also reflect entrenchment over the course of the crisis.

The final step in the case analysis is the comparison of Bayesian and biased models based upon the sub-cases. Table 6.1 summarizes the performance of the models side by side. The coding on model success, 0 and 1, is based on the capability of model to explain the outcome and which model better explains the outcome. For example, if the Soviet Union attempted to surprise the U.S. and model A predicts that the Soviets would mix their strategy and model B predicts they always surprise, then model B is coded a 1 and model A gets a 0. If both models predict a deterministic result, both theories are scored 1 if the case behavior matches the deterministic results, or coded 0 otherwise. If both models predict mixed outcomes, the analysis focuses on the differential between the deductive results. If model A predicted a mixed strategy where public challenge was more likely than it was in model B, model B would coded 1 and model A 0 if the Soviets attempted surprise. If both strategies predict an identical mixed result, such as model A

predicts a 50% chance of surprise and model B predicts a 50% chance of surprise, there is no way to evaluate the models relative to one another in a meaningful way so both are coded zero.

The top part of Table 6.1 shows conditions where we expect the predictions of the Bayesian and biased models to converge. Given prior U.S. beliefs, both models indicate that the Soviets were better off attempting surprise than making a public announcement. Both model predictions are pure rather than mixed strategies. Both models, therefore, explain the crisis initiation equally well. This does not mean that Khrushchev's decision to secretly deploy nuclear weapons to Cuba was necessarily rational, but it does show that the behavior was consistent with the deductive results of a Bayesian rational model. Kennedy actually voiced his opinion that the missile deployment was a reasonable move for Khrushchev (May and Zelikow 2001, p. 111; Leaming 2006, p. 403). The two models also converge on the target's optimal strategy, which is waiting. Since the Soviet intentions in Cuba are perceived to be defensive, evidenced by the intelligence assessment prior to October 16, the target always waits whether biased or Bayesian. The U.S. did not take any action to impede the Soviet arms shipments prior to October 22. There is convergence between the Bayesian and biased models, and since both accurately predict the same behavior, neither model offers an explanatory advantage.

Table 6.1: Cuban Missile Crisis Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
Prior to October 16:					
Target's Belief			Aid Was Defensive		
Evidence of Entrenched Beliefs			No		
Details of Entrenched Beliefs			Rapid Belief Change		
Sender Concern About Target Beliefs			Yes; Deception Effort		
Proposition 1 ($p<p^*$):					
Sender Use of Surprise[1]	Always When Challenging	Always When Challenging	Soviets Attempt Surprise	1	1
Target Response[2]	Always Waits	Always Waits	Target Unprepared	1	1

Note: The table shows that both models score well in instances where they have identical or convergent predictions. In instances where they diverge, the biased appears to have better explanatory power. See the footnotes or text for details on coding explanations.

1. Both models predict that the sender will always attempt surprise, and the Soviets did attempt surprise, so both models receive a score of 1.
2. Both models predict that the target will always wait, and the U.S. did not take any action to prevent the arms shipments when they believed them to be defensive, so both models receive a score of 1.

Table 6.1 (cont'd): Cuban Missile Crisis Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
US Secretly Debates, October 16-22:					
Target's Belief			Aid Was Offensive		
Evidence of Entrenched Beliefs			Yes		
Details of Entrenched Beliefs			Offensive Nature of Missiles; No Nuclear Warheads in Cuba		
Sender Concern About Target Beliefs			Unaware of Target Knowledge		
Proposition 2 (U.S.: $p>p^*$; Soviets: $p<p^*$):					
Sender Use of Surprise[3]	Surprise or Violent Act Likely	Surprise or Violent Act Likely	Continued Surprise Attempt	1	1
Proposition 3 ($p>p^*$):					
Target Response when Bayesian[4]	Mix	Mix	Blockade	0	-
Target Response when Biased[4]	-	Always Act	Blockade	-	1

3. Since the Soviets are operating under the belief that the U.S. is still unaware of the missiles and allowing the defensive arms shipments ($p<p^*$), both models predict that the sender will always attempt surprise. The Soviets continue their attempted surprise, so both models receive a score of 1.

4. The U.S. knows of the missiles and changes their beliefs about Soviet intentions ($p>p^*$). If the Soviets continue to play the Bayesian strategy (as they should since they do not know that the U.S. knows), the target in the Bayesian model will adopt a mixed strategy, the Bayesian type in the biased model will adopt a mixed strategy, and the biased type in the biased model will always act (since the sender adopts too aggressive of a signaling or surprise strategy). The analysis above argued that the U.S. starts exhibiting entrenched beliefs at this time, so the biased type in the biased model is the relevant comparison to the Bayesian model. Since the U.S. dismisses the possibility of waiting, preferring to take immediate action, the biased model better captures U.S. behavior.

Table 6.1 (cont'd): Cuban Missile Crisis Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
Post October 22:					
Target's Belief			Aid Was Offensive		
Evidence of Entrenched Beliefs			Yes		
Details of Entrenched Beliefs			Offensive Nature of Missiles; No Nuclear Warheads in Cuba		
Sender Concern About Target Beliefs			Sender Aware That Target Knows		
Proposition 2 ($p>p^*$):					
Sender Use of Surprise[5]	Surprise or Violent Act Likely	Public Act Likely	No Soviet Retaliation (Particularly Against Berlin)	0	1
Proposition 3 ($p>p^*$):					
Target Response when Bayesian[6]	Mix	Wait	Blockade	0	-
Target Response when Biased[6]	-	Mix	Blockade	-	0
Total				3	5

Statistical Analysis on All 6 Observations		Statistical Analysis on 3 (Divergent) Observations	
t-Statistic	-1.581	t-Statistic	-2.000
P(T<=t): Two-Tail	0.175	P(T<=t): Two-Tail	0.184
t Critical Value: Two-Tail	2.571	t Critical Value: Two-Tail	4.303

5. The Soviets learn the U.S. knows about the attempted surprise ($p>p^*$), and if they believe there is some possibility that the U.S. is irrationally over-suspicious, they should adopt a less provocative strategy. They do not take any action (against Berlin) in response to the blockade, and they try to deescalate the crisis. The biased model predicts more conservative strategies than the Bayesian model, meaning it better captures Soviet behavior, and receives score of 1.

6. Neither model receives any points. The Bayesian model predicts a mixed strategy, while the biased model with the biased type (since the U.S. is still exhibiting signs of entrenched beliefs according to the argument above) predicts an identical mixed strategy. Neither model can claim to provide a better explanation than the other.

The final convergent prediction in the case extends into the second stage of crisis wherein the U.S. knew of the missiles, but the Soviets were functioning under the belief that the deployment was still secret. The biased model shows that there is no reason to take a target's anchoring into account if the sender is perceived to be non-threatening.[26] Since the Soviets were still operating under the assumption that U.S. did not know of the missiles, both models predict that aggressive Soviets should always attempt surprise. This result to similar the one discussed immediately above. Once again, both models offer the same prediction, the case actually matches the deductive conclusions, and both models function equally well. It is interesting to note that in the three instances where the two model predictions converge, both offer predictions that are in line with the case events.

The Bayesian and biased models diverge after October 16 when the U.S. knows about the Soviet missiles in Cuba, but it is uncertain about Soviet intentions or next steps. The models accurately predict that Soviet behavior should remain unchanged since they believed the missiles were still a secret. The U.S., however, had additional information and now believed that the Soviets had expansionist or aggressive intentions associated with Cuban military aid. Given this revised belief, both the Bayesian model, and the biased model where the target is actually Bayesian, predict that the U.S. would adopt a mixed strategy. Contrastingly, the biased model predicts that a biased type of target would always act. Above, there seemed to be some evidence to support the assertion that the U.S. had entrenched beliefs about the weapons in Cuba and the Soviet rationale for placing them. If this were true, the biased model with a biased actor is the pertinent comparison to the Bayesian model, and the biased model appears to be a better fit. The

[26] Note that the technical definition of non-threatening is $p<p^*$.

ExComm discussed the possibility of doing nothing or taking diplomatic measures, but there was a strong conviction that some military steps were required to signal resolve. Since there was strong consensus for some sort of action throughout the crisis, the biased model seems to have better explanatory performance than the Bayesian model.

There is also divergence in the results for the third stage of the crisis after October 22. The ExComm discussions show that the U.S. decision-makers worried that action against Cuba would trigger Soviet retaliation, and were particularly concerned about Berlin. They repeatedly reasoned that a blockade of Cuba would lead to a Soviet blockade of Berlin. The Soviets' reaction differed from the ExComm's expectations. The Soviets never considered escalating the crisis by acting against Berlin and focused on how to deescalate the crisis while ensuring that Cuba was safe from U.S. action. When the biased and Bayesian models are compared, the Bayesian model predicts that aggressive strategies such as surprise are more likely than more benign crisis strategies like public statements. The deductive models highlight that in uncertain situations where opponents are potentially over-suspicious, the only way to avoid antagonizing the opponent into overreaction is by being more conservative. Since the game is a "one-shot" interaction, which is to say it does not explicitly consider future interactions beyond the exogenous payoffs, its logic can apply to any single crisis stage or interaction (initiation, escalation, ongoing war) where there is uncertainty and preparing military action or surprise is possible. Thus, the Bayesian model predicts that the Soviets would be more likely to adopt a provocative escalatory strategy, whereas the biased model predicts that there is a higher relative likelihood of adopting more conservative strategies. Given the Soviet deliberations and final strategy, the biased model, predicting that the Soviets were

likely to adopt a less provocative strategy, outperforms the Bayesian model predicting a relatively more provocative approach.

The final proposition tested in the table addresses the target's strategy after October 22. The U.S. maintained the blockade and relied on diplomacy to deescalate the crisis. The Bayesian model predicts that the target would mix its strategy, the biased model with the Bayesian type predicts that the target would always wait, and the biased model with biased type predicts that the target would adopt the same mixed strategy as the Bayesian. If the U.S. had some degree of entrenched beliefs as argued above, it is appropriate to compare the Bayesian model to the biased model with the biased type. Since both predict that the target mixes their strategy, and the probability with which both mix their strategy is identical, neither one outperforms the other or wholly captures the U.S. behavior after October 22.

A raw count reveals that the Bayesian model offered accurate deterministic predictions in three cases, yielding a total score of three. The biased model offered accurate deterministic predictions in the same three cases and outperformed that Bayesian model in two cases, yielding a total score of five. It is not clear that the results suggest that one model is superior to the other statistically since there is a random chance that the biased outperformed the Bayesian model. The way to eliminate, or at least minimize, the possibility that the biased model was randomly superior is a statistical analysis comparing the two models. The studentized t-statistic is an appropriate test for small samples such as this. A studentized t-statistic analysis of the two groups shows that the probability that the results of biased model are statistically different from the Bayesian is 82.5% on a two-tailed test. This does not meet the normal level of statistical significance of 95%.

However, an analysis looking at differences between the models should not include the three explanations where the models have convergent predictions. Given that the three predictions are the same, and both models score highly in these instances, it would be surprising if the model results were found to be statistically distinct. To evaluate the relative quality of the models, it makes sense to focus on instances with divergent predictions. There are three divergent cases in which the Bayesian model scores zero and biased model scores two. The statistical analysis on the three cases shows that the probability that the results of the biased model are statistically different from the Bayesian is 81.6% on a two-tailed test. Once again, we find that the test does not meet the standard 95% level for significance, but the scores are still reasonably high. One major constraint to getting a statistically significant result is that there are so few cases that it is difficult to distinguish the two models. If there were two additional observations coded at the sample means (0 for the Bayesian model and 0.666 for the biased model), the statistical analysis would reveal that the models were different with a 97.5% chance with a two-tailed test. These results meet the standard requirement for significance. The model comparison, therefore, is slightly hamstrung by the fact that there are convergent results and only a few divergent predictions.

Broadly speaking, the biased model of decision-making developed here seems to perform relatively well. It is successful in instances were its predictions converge with the Bayesian model, and seems to outperform the Bayesian model given divergent predictions, but we cannot eliminate the possibility that the superior performance was random. The analysis suggests that the biased model performs at least as well as the

Bayesian model, and it may have greater explanatory power than the Bayesian benchmark.

The Arab-Israeli War, 1967

Historians and political scientists have also extensively studied the Arab-Israeli conflict that played out between April and June 1967. Scholars have used it as prototypical case of escalatory spirals and psychologically motivated conflict (Stein and Tanter 1980; Lebow 1981; Jervis, Lebow and Stein 1985; Stein 1988), group decision-making and rational decision-making (Laqueur 1969; Bar-Zohar 1970; Mor 1991). Others have focused on the personalities of the key leaders, particularly Egyptian President Gamal Abdul Nasser (Nutting 1972; Lacouture 1973), as well as Israeli leadership (Segev 2005). While there is a good deal of information available about the events leading up to the war, and a number of secondary sources that have used this data to make causal arguments, there is no single explanation that generates strong consensus. There are two reasons why scholars may not have settled on a dominant explanation. First, there is tendency to use the 1967 case to argue in favor of a favorite causal mechanism, psychological failure, escalatory brinkmanship or rational behavior (Mor 1991). Second, and perhaps more importantly, very few of the explanations offer a strategic assessment that details decision-making on both sides and the way that decisions interrelate.

We attempt to overcome these shortcomings in this case analysis by following the same format used in the Cuban Missile case. The study will start by providing some background information on the case that speaks to Israeli beliefs about its neighbors and

their intentions. Then we will cover the case details, including Egyptian actions along with Israeli responses to these Egyptian signals. We will also discuss the impact of external actors, particularly the U.S., on both sets of the decision-makers. The role of these third parties is a key component of the case and yields information that helps to explain why actors did or believed certain things at certain junctures. The case will then turn to an assessment of Israeli beliefs and biases, and Egypt's beliefs about Israel. Finally, we will compare the Bayesian and biased models as we did in the Cuban case, justify the coding, and evaluate the results.

Relations between Israel and its Arab neighbors had been turbulent since the state's founding by United Nations Charter in 1948. Israel's statehood immediately triggered an attack by neighboring Arab states, and Israel found itself surrounded by hostile forces on three sides. Israel was victorious, but that did not bring an end to small scuffles or major violent acts. Borders scuffles with Syria, Egypt and Jordan were regular occurrences, however, there was no further full-scale conflict until 1956 (Mor 2004, p. 315).

In 1956, Nasser announced the intention to nationalize the Suez Canal, and would likely close the Red Sea to Israeli shipping. Nasser's actions also concerned Britain and France, both of whom relied on open access to the canal for trade and supply. Egypt's attempt to nationalize the canal, therefore, motivated a marriage of convenience between the Israelis, British and French (Kissinger 1994, p. 540). With backing from the British and French, the Israelis attacked Egypt, and the combined force successfully captured the Sinai Peninsula and the Suez Canal. U.S. President Dwight Eisenhower was not pleased, in part because he had not been given advanced warning, and in part because the U.S.

was also dealing with the aftermath of the Hungarian uprising (Gaddis 2005, p. 70). The U.S. imposed a cease-fire whereby Britain and France withdrew from Egypt in 1956 and were replaced by troops from the United Nations Emergency Force (Oren 2002, p. 12). The Israelis withdrew from the Sinai in 1957.

From 1957 through 1967, there were few instances of violence between Israeli and Egyptian forces, owed in part to the UN troops stationed in the Sinai. Borders with Jordan and Syria, however, were not equally quiet. Palestinian forces regularly entered Israel through the Jordanian, and occasionally the Syrian, borders to carry out terrorist attacks (Segev 2005, p. 144). The Israelis responded with raids against Jordanian towns suspected of harboring or supporting Palestinian terrorists. The Israelis were convinced that a new Palestinian group Fatah, whether entering Israel from Jordan or Syria, were trained and supported by the Syrians (Segev 2005, p. 143).

Relations with Syria were particularly poor since the 1948 war. The Israelis and Syrians had fallen into a pattern of provocation and retaliation (Rabin 1979, p. 51). The 1948 war created de-militarized zones along the border that were off limits to both Syrians and Israelis. Both sides tried to extend their lands piecemeal. The Israelis would send tractors, often with bulletproof armor, to plow disputed territory and the Syrians would respond by firing on the tractor and the associated kibbutz. Likewise, the Syrians would bring flocks to graze in disputed territory and face Israeli retaliation. One particularly contentious issue for the Israelis was Syria's attempt to divert the Jordan River, thereby eliminating an important water supply (Segev 2005, p. 192). Israel often responded by bombing the construction equipment tasked with the river diversion. These

small border scuffles always had the possibility of escalating into war as both sides constantly tested each other's will.

There were 120 terrorist attacks in Israel in the 18 months leading up to the 1967 War, and during the first few months of 1967 there were 270 incidents along the Jordanian border (Oren 2002, p. 45). Two attacks in October 1966 were particularly important. The first was the bombing of an apartment building in Romema, a suburb of Jerusalem, on October 7 (Segev 2005, p. 147). The next major attack was on October 25, where a landmine derailed two engines of a cargo train traveling from Jerusalem to Tel Aviv (Segev 2005, p. 149). Finally, on November 11, a military car carrying three paratroopers struck a roadside landmine killing all three soldiers (Segev 2005, p. 149). Israeli policymakers, particularly Prime Minister Levi Eshkol, had exercised patience up until this point. The death of the soldiers, however, brought mounting pressure for military reprisal (Segev 2005, p. 152). King Hussein of Jordan tried to contain the situation by issuing an apologetic message to Eshkol on November 13, but the Prime Minister had already approved a raid on the town of Samau (Segev 2005, p. 154). The operation expanded beyond its original scope, destroying somewhere between 40 and 100 houses and brought Israeli troops into contact with Jordanian troops. The operation triggered riots in Jordan (Oren 2002, p. 36; Segev 2005, p. 182).

The regularity of the terror attacks coupled with the possibility of destabilizing Jordan with a military response, motivated Israeli policymakers to undertake major reviews of the significant threats from Jordan and Egypt (Segev 2005). A less formal review on Syria was conducted later. Since the Israelis believed that the terrorists were trained by neighboring states, they focused their assessments on state-level threats.

From November 1966 to January 1967, the Israelis set up two working groups, one focused on Jordan and the other focused on Egypt (Segev 2005, pp. 183-84). The assessment teams were staffed with personnel from Mossad, Israeli Defense Force (IDF) intelligence, and the Foreign Ministry. The threat assessment on Jordan noted its weakness, and largely focused on possible solutions to Palestinian sabotage (Segev 2005, p. 185). One possible approach was that the Israelis assemble saboteur teams, similar to those used by the Palestinians, and reciprocate violent acts. One problem with this approach, and any others that involved hostility against Jordan, was the possibility of destabilizing the regime (Segev 2005, p. 185). There was some debate whether the regime should be destabilized given the riots following the Samua operation. The group concluded that Egypt and Syria would intervene should King Hussein's government fall, and this might give Israel pretext to occupy the West Bank (Ha'aretz 1966). The problem, then, was what to do with the West Bank. The territory could be annexed, but the Israelis believed that the international community would reject such a move. Alternatively, setting up a weak state or protectorate with handpicked leaders and policed by the IDF was deemed as not viable. The assessment process concluded that a weak Jordan was the best possible scenario for Israel, in part to ensure that the IDF could operate in Jordan if necessary, and in part because King Hussein's policies had been beneficial for Israel (Segev 2005, pp. 186-86). King Hussein had encouraged Palestinians to migrate from the West Bank of the Jordan River to the East Bank, and roughly 100,000 had chosen to move. Since the Israelis were concerned about Palestinian population growth, particularly in the West Bank, King Hussein's policies matched Israeli interests. The group concluded that a continuation of the status quo was in Israel's best interests.

The assessment on Egypt, not surprisingly, had a different tone. There was no talk of undermining Nasser's rule or destabilizing Egypt, which would have been impractical. Instead, the discussions focused on Nasser's intentions and the possibility for improved diplomatic relations. The IDF assessment concluded that Egypt aspired to destroy Israel and there was no possibility for rapprochement (Segev 2005, p. 187). Mossad agreed that Egypt was interested in destroying Israel, but noted two possible avenues for diplomatic inroads. The first required dismantling Jordan and establishing a Palestinian state in its place (Segev 2005, pp. 187-88). They concluded, however, that this would indirectly strengthen Egypt and was counter to Israel's interest. The second option was the termination of the nuclear weapons project at the Dimona (Segev 2005, p. 188). The assessment reasoned that the Israeli nuclear program was far ahead of the Egyptian program. Coupled with IDF conventional military superiority, the Egyptians might find the concession sufficiently attractive.

The Foreign Ministry countered that nuclear weapons represented the best hope of ensuring Israel's safety, and in the absence of such weapons, it could not be content with the current security environment. The question was one of deterrence. Israel, for some time, believed that its greatest hope of survival lied in deterrence: first through the IDF and conventional military superiority and then through nuclear weapons (Segev 2005, p. 189). The deterrence strategy had been used to justify many previous operations in Jordan and Syria. Deterrence advocates argued that Israel must appear strong and resolved if it was to maintain the reputation necessary to deter future attacks (Rabin 1979, p. 81). This meant reciprocating acts of violence. The perceived deterrent benefits of nuclear weapons were difficult to abandon. The assessment concluded that Egypt wanted to destroy Israel

and there was very little the Israelis could do to alter that position or improve diplomatic relations (Segev 2005, p. 189).

This 1966-67 assessment, however, did not encapsulate the full range of beliefs about Egyptian intentions. There was consensus that Egypt wanted to destroy Israel, but IDF intelligence and senior decision-makers believed that a large-scale Egyptian attack was unlikely in the near term (Nutting 1972, p. 408). The IDF was superior to the Egyptian military, and Egypt was likely to lose a full-scale war. Senior Israeli decision-makers believed that Egypt shared this perspective, making large-scale Egyptian action unlikely until at least 1970. Though the Israelis believed that the Egyptians would not look to start a war, they thought it very possible that Egypt would mount an attack on the nuclear facility at Dimona or close the Straits of Tiran to Israeli shipping (Oren 2002, p. 63; Segev 2005, p. 229). These fears were sufficient to instill suspicion and justified careful monitoring of Egyptian forces. There were two examples in November 1966. The first was a confrontation with an Egyptian plane entering Israeli airspace. The Egyptian MiG was shot down. In the same month, Israeli intelligence also learned that Egypt was reinforcing its troop deployment in the Sinai. The Israelis called up reservists, but released them quickly. Israel was certainly concerned about their neighbor whom they believed harbored aggressive feelings and objectives.

The third major threat facing Israel, though not assessed formally in 1966, was Syria. The Israelis probably saw little reason for a formal review, since they believed that Syria would be the spearhead in any attack on Israel (Segev 2005, p. 203). The Israeli General Staff discussed the Syrian threat in January 1967, focusing on a few key issues. The first was the inability to contain Fatah, perceived to be a Syrian agent (Segev 2005,

p. 203). The reciprocal fighting along the border also caused concern. The regular border clashes led to the belief that Syria might drag Israel into an undesired war (Segev 2005, p. 203). Eshkol tried to focus discussions on defensive steps, but the military felt that there were no defensive steps sufficient to prevent attacks and kept raising the possibility of offensive action (Segev 2005, p. 206). The military proposed three plans for retaliating against Syrian provocations: an operation along the border, an operation that pushed further north, and an operation aimed at Damascus (Segev 2005, p. 206-7). The relationship between the Soviet Union and Syria also fueled concern. Should the Soviets wish to take action against Israel, the Syrians would be their proxy, and any Israeli action against Syria risked provoking a Soviet response (Segev 2005, p. 201).

The Israelis were just as suspicious of Syria as they were Egypt, if not more so. These suspicions, and the need to show resolve, helped trigger a crisis on April 7, 1967 (Segev 2005, p. 210). Eshkol, at Chief of Staff Yitzhak Rabin's suggestion, approved a plan to plow part of the demilitarized zone in the North. The Syrians opened fire on the armored tractor and the IDF returned fire. The conflict escalated and by the end of the day, the Israelis had shot down six Syrian MiG jets. Egypt took no action in support of Syria, which drew condemnation from the Arab world since Egypt and Syria had a defense pact since 1966 (Oren 2002, p. 47). Egypt had repeatedly asked Syria to moderate the border violence and terminate its support of Fatah lest they find themselves in a war with Israel before they were ready.

The Syrians disregarded the Egyptian request, and tensions escalated along the Israel-Syrian border. Between April 7 and Israeli Independence Day on May 15, Fatah carried out 14 attacks, and Syria regularly shelled Israeli settlements (Oren 2002, p. 48).

It was common for the Prime Minister and Military Chief of Staff to offer their thoughts about domestic and security policy in the month proceeding Independence Day. Chief of Staff Rabin was particularly bellicose in his statements, and the Israeli leadership tried to moderate them (Eban 1991, p. 353). Eshkol himself, however, suggested that the Syrians would be mistaken if they believed that there would be quiet on only one side of the border (Segev 2005, p. 215). Israeli statements, and possibly leaked military plans, began to turn some heads, particularly in Moscow (Oren 2002, p. 54). The Syrians had been concerned that Israel might attack for some time, but the Egyptians continued to ignore these warnings. The latest warnings of impending Israeli aggression were now coming from the Kremlin. The Soviets estimated that the Israelis had moved the bulk of their force, 14 divisions, up North (Oren 2002, p. 55). The Soviet ambassador made such charges, only to have them refuted by the Israeli Foreign Ministry. The Israelis offered to take the Soviet ambassador up north three times, but he chose against making the trip. The warning from Moscow alarmed Nasser, since Egypt could not afford to ignore further Israeli attacks against Syria without essentially destroying the defense pact (Draper 1968, p. 44).

Evidence leading up to the 1967 War suggests, however, that Nasser wanted to avoid war with Israel in the near-term (Dawn 1968, p. 202; Khouri 1968, p. 245; Stephens 1971, p. 466; James 2006, p. 105). Nasser believed that Egypt was not prepared for war, and was concerned that Egypt would lose territory and pride in such an encounter (Mor 1991, p. 364). This sentiment was echoed during January 1967 meetings of United Arab Command (UAC). Not all of the Egyptian leadership wholly agreed with this position. The head of the armed forces, Field Marshal ‘Ali ‘Ali ‘Amer, had filed a

report with UAC indicating that while they lacked the resources to liberate Palestine he believed that Egyptian numerical superiority could make up for Israel's technical superiority (Oren 2002, p. 20). 'Amer advocated expelling the UN force in advance of any operation (James 2006, p. 105). Nasser also wanted the UN force out of the Sinai, but did not think it was worth pushing until the military was better prepared to deal with Israel. There were, therefore, different perspectives among the Egyptian leaders. Nasser wished to avoid war with Israel in the near future (Dawn 1968, p. 202; Stephens 1971, p. 477), while others pressed for war sooner rather than later.

Irrespective of Nasser's desire to avoid war, he was unwilling to sit back if Israel was preparing a large-scale attack on Syria (Draper 1968, p. 44; Stephens 1971, p. 466). The Israelis had no immediate plans for such an attack, but the Soviet warnings and escalating tensions along the border worried Nasser. He took action on May 14, the day before Israeli Independence Day. The military called up a large force, marched them through Cairo and towards the Suez Canal. The mobilization was a public event intended to signal the Israelis that Egypt would not tolerate an attack on Syria. Many commanders and troops felt as though the mobilization was poorly organized, and many troops had no clear idea where they were going or why (Oren 2002, p. 65-66). The Egyptians mobilized according to a defensive plan called Operation Conqueror calling for troops to enter the Sinai and take position along three lines (James 2006, p. 105). While offensive operations were not ruled out, there were no orders issued for offensive action, and there was no discussion of offensive action until later in the crisis. This suggests that the Egyptians had no intention of attacking Israel with the initial mobilization, and he may have believed that the mobilization was unlikely to trigger war (Safran 1969, p. 268). Those close to

Nasser believed that he was trying to show support for Syria, bolster Egyptian pride, reaffirm his standing in the Arab world, and send a warning to Israel (Mor 1991, p. 365). Whether Nasser overestimated Egyptian military power and desired war (Nutting 1972, p. 410), or whether he was trying to bolster Egyptian position without firing a shot, the public and disorganized nature of mobilization would have hindered any attempt to successfully surprise Israel, making surprise a less attractive strategy at this point in the crisis.

The Israelis noticed the mobilization almost immediately (Rabin 1979, p. 68). The IDF intelligence chief contacted the Chief of Staff and the Prime Minister with details of the Egyptian troop movements through Cairo before the Independence Day festivities started on the evening of May 14. Neither Eshkol or Rabin were immediately concerned by the mobilization which may have been intended to counter Israel's annual Independence Day military display (Rabin 1979, p. 69). As the day wore on, and the troops continued moving towards the Suez, some Israeli decision-makers grew increasingly concerned. Rabin instructed the Chief of Central Command, seated a few rows behind him at the Independence Day festivities, to return to headquarters and monitor the Egyptian advance (Rabin 1979, p. 68). Both Eshkol and Rabin, however, clung to the belief that this Egyptian reinforcement was similar to the previous stillborn 1964 and 1966 mobilizations (Rabin 1979, p. 69; Segev 2005, p. 187).

Sentiment changed between May 15 and 16. On May 16, IDF intelligence warned that the Egyptians could be planning for an attack. This aggressive belief, which was not far removed from the assessment of Egyptian intentions a few months earlier, was brought on by better intelligence about the Egyptian mobilization. The initial IDF

assessment on the mobilization showed that the Egyptians were moving a single division consisting of 15,000 troops, which seemed like a reinforcement of the 30,000 troops already in the Sinai. This assessment was wrong. The Egyptians had sent the 2nd, 4th and 7th Infantry divisions across the canal and the 6th Armored division was on its way (Stephens 1971, 473; Oren 2002, p. 63). This brought the total mobilization to 45,000 infantry and 100 T-54 and T-55 tanks, along with 150 armored personal carriers and a range of artillery weapons. The eventual mobilization was three times larger than the Israelis initially estimated (Eban 1991, p. 370). There were also mixed signals of Egyptian intent. Rabin noted that troops were digging in as though they were taking defensive positions, while broadcasts from Cairo took a particularly bellicose tone in calling for the immediate destruction of Israel (Eban 1992, p. 353).

In addition to the revised estimates, May 16 was also the day that Egypt formally notified the UN that the emergency force was no longer welcome in the Sinai (James 2006, p. 108). Nasser initially requested that the troops move into Gaza and remain in Sharm esh-Sheikh, the latter would keep Egypt from seizing the Straits of Tiran, but UN Secretary General U Thant believed that the force should retain its current disposition in the Sinai or leave altogether (Stephens 1971, p. 471; McNamara 2000, p. 623). Nasser requested that the force be withdrawn and the positions handed over to the Egyptian army.

By May 17, Rabin and the IDF had changed their assessment of the mobilization (Rabin 1979, p. 70; Segev 2005, p. 228). The size of the mobilization, the UN withdrawal, and the bellicose language all suggested that Nasser was willing and possibly ready to go to war. Rabin requested the call-up of 18,000 reservists to begin defensive

preparations and Eshkol approved the order. Later that day, the Israelis faced another scare. For the first time, two Egyptian MiGs flew over the Dimona facility. The flight mission was limited to intelligence gathering, but it seemed to confirm earlier assessments that the Egyptians might target the facility (Oren 2002, p. 75). This did nothing to mollify Israeli fears, and it offered hawkish elements of the Israeli government the leverage needed to advocate immediate Israeli military action. The absence of military action would weaken Israel's deterrent credibility, thereby encouraging Egypt to strike first (Rabin 1979, p. 181; Stein 1988, p. 2). Forgoing the first strike advantage, the army argued, would mean a longer war with higher casualties. Eban, however, was not convinced that Nasser wanted war (Eban 1992, p. 360).

From May 17 to early June, the political situation deteriorated in Israel. Tom Segev (2005) provides voluminous detail about these intrigues, alliances, personalities, and compromises. Faced with possibility of an Egyptian attack, military leaders and government hawks assailed the leadership, and particularly Prime Minister Eshkol. The military argued that Israel needed to seize the initiative and waiting would severely weaken Israel's fighting ability. Waiting, they argued in politically charged language, was tantamount to inviting a second Holocaust. The Prime Minister was sensitive to the accusation that he might inadvertently damage Israel's defensive capabilities, but he also believed that U.S. support was vital for the long term survival of the country (Rabin 1979, p. 78). Beating the Egyptians was important, but it was equally important to maintain Israel's alliance with the Western powers. The U.S. had tacitly suggested that action against Syria would be acceptable, but had warned against going to war with Egypt (Rabin 1979, p. 77; Oren 2002, p. 77-78). If Israel was to go to war with Egypt, it

should only be in response to Egyptian attack or military provocation. Israel should not initiate such a war. Given the U.S. position, Eshkol believed it prudent to wait, despite the sentiments of the military leadership. General Staff meetings with the Prime Minister, who also served as Defense Minster, were often heated. In the later stages of the crisis, some military members hinted at the need to instill an emergency government or supplant civilian control (Segev 2005, pp. 292-294). The media and popular opinion echoed some of these concerns that Eshkol was not a suitable wartime president.

The Egyptians raised the stakes once more on May 23 (James 2006, p. 113). They announced that the Straits of Tiran were closed to Israeli shipping, an act that Israel had long warned would be justification for war (Stephens 1971, p. 474; Eban 1992, p. 366). As the noose tightened around the Israeli leadership further, the military again stepped up its insistence that they attack immediately. They claimed that the Israel's existence was now threatened for the first time in this crisis (Brecher 1980, p. 104). The U.S. responded to the announcement by asking for a 48-hour moratorium on any reaction (Rabin 1979, p. 77). Eshkol decided to send Foreign Minister Abba Eban to the U.S. in order to explain the Israeli position and get a first hand impression of U.S. sentiments. He also figured that the mission might offer more time for diplomacy (Eban 1995, pp. 375-395).

As soon as Eban arrived in Washington, he received a cable from Eshkol indicating the situation had grown even more dangerous. Between May 23 and 25, Israeli intelligence accurately reported that Egypt called up additional divisions. While the IDF had issued warnings before this latest mobilization, they now reasoned that the Egyptian force was large enough to mount an offensive and was outsized for defensive purposes (Rabin 1979, p. 85). Eban believed that Israeli intelligence and Cabinet members were

inflating the threat to Israel in order to pressure the U.S. into offering support (Eban 1992, p. 285). He resented the strategy, but delivered the message. U.S. security officials also believed that Israel was inflating the threat as CIA estimates maintained that the Egyptians were not planning an imminent attack (Rabin 1979, p. 88; Robarge 2007). While U.S. leaders also acknowledged the gravity of the threat, they reiterated that Israel should not be the state to initiate hostilities. Privately, U.S. assessments indicated that Israeli forces would easily defeat the Egyptians, whether or not they seized the initiative and used surprise (Bowen 2003, p. 60; Robarge 2007). This estimate drove U.S. decision-makers to urge patience and restraint. Eban also pushed for a response to the closing of the straits. He hoped that President Lyndon Johnson would agree to send the 6th Fleet to open the straits if necessary, but failed to get a commitment (McNamara 2000, p. 631). Instead, the U.S. discussed the possibility of an international force to reopen the straits to Israeli ships.

History shows that the Israeli intelligence assessments were better than those of the CIA (James 2006, p. 112). The initial Egyptian deployment in the Sinai, as disorganized as it was, followed the defensive plan called Operation Conqueror. The goal of Operation Conqueror was the remilitarization of the Sinai with rows of defensive fortifications capable of withstanding an IDF attack (Oren 2002, p. 65). As the crisis continued, Field Marshal ‘Amer began increasing the scope of military operations (Oren 2002, p. 92-93). Newly proposed operations, first Operation Lion aimed at cutting off Eilat from Israel, then Operation Dawn aimed at seizing the entire Negev, were offensive. Some questioned why Nasser did not try to counter ‘Amer’s ambitions, particularly when the scope of the operations seemed far removed from realistic political goals (James

2006, p. 114). There are three explanations for Nasser's behavior: the complicated relationship between 'Amer and Nasser, Nasser's conflicting feelings about war with Israel, and Nasser's belief that he could terminate the operation after the planning stage (Oren 2002, p. 92). Irrespective of Nasser's justification, Israeli intelligence accurately assessed Egypt's changing military posture, and the offensive nature of the most recent mobilization.

'Amer reported that the preparations for Operation Dawn were complete on May 25, however, Nasser was hesitant to approve the attack (Oren 2002, p. 94). He delayed attacking in order to gauge Soviet sentiment. Nasser was concerned that Egypt might draw the U.S. into a conflict if it attacked Israel, and wanted to avoid fighting both simultaneously. He dispatched a team to the Soviet Union to meet with officials (James 2006, p. 114). The Soviets had initially supported the remilitarization of the Sinai and expulsion of the UN force, but they were surprised when Nasser closed the straits (Oren 2002, p. 94-95). He had not warned the Soviets or consulted. Given the risks associated with closing the straits, the Soviets felt less apt to offer blanket support. While the Soviets encouraged the Egyptians to consolidate their gains and warned against further aggression, they also pledged that they could not stand by if the U.S. got involved. The Egyptian delegation, probably inaccurately, interpreted this as a sign of Soviet support and conveyed that message to Nasser (Oren 2002, p. 117). The Operation Dawn attack on Israel was planned for May 28, 4:00am.

The evening of May 27, the Soviet Ambassador to Egypt Dmitri Pojidaev told Nasser that the Soviets had received a warning from the U.S. indicating that intelligence revealed preparations for an Egyptian attack on Israel (Oren 2002, p. 120). The

ambassador told Nasser that such a move would give the U.S. a free hand to intervene. Nasser reassured the ambassador that there were no such plans, then went to Supreme headquarters to meet with ‘Amer. Nasser immediately canceled the operation. ‘Amer protested, but Nasser countered that the attack would draw international support for Israel and warned that the U.S. might send in the 6th Fleet (James 2006, p. 115). Later Nasser also cited security concerns. The U.S. warning to the Soviets had almost certainly come from the Israelis, meaning that the element of surprise was lost (Oren 2002).

By the time Eban returned from Washington, debates had grown even more heated, but Eban recalls a tranquil sense (Eban 1992, p. 396). This was partially a result of the May 25 IDF intelligence assessment indicating that an Egyptian attack was imminent (Rabin 1979, p. 85). The IDF used the revised intelligence to pressure Eshkol and the civilian leadership for action (Weizman 1976, p. 213). On May 27 and 28, the Israeli Cabinet was deadlocked: nine ministers supported waiting and nine supported immediate war (Rabin 1979, p. 91). Eshkol, while still advocating restraint and concerned about U.S. support, was bolstered by Eban’s return. The Foreign Minister reported that the U.S. would first try to assemble an international force to reopen the straits, but would use the 6th Fleet alone if necessary (Rabin 1979, p. 88). Eban, therefore, overstated the U.S. commitment, but in doing so, helped Eshkol maintain sufficient support among the civilian leadership (Segev 2005, p. 273). In light of U.S. support, the Cabinet unanimously agreed to wait two weeks in order to see if the straits would reopen (Rabin 1979, p. 91). Military leadership, and opposition leaders such as Moshe Dayan, were furious with the Cabinet and privately accused them of betraying the country (Rabin 1979, p. 92). While the media and military barrage against Eshkol escalated, the Prime

Minister believed that the waiting period he had secured would help maintain favorable relations with the U.S. even if Israel declared war later.

For a brief time, the Israeli cabinet agreed that they should wait rather than act. This unanimous sentiment was undone by Nasser's next escalatory step. Adding more fuel to the fire, Egypt and Jordan signed a defense pact on May 30. This dramatically changed the Israelis' security situation (Stein and Tanter 1980, p. 107). For the first time since 1948, Israel was completely surrounded by hostile states (Stein and Tanter 1980, p. 218). Any war with one state, would almost certainly involve war with all three (Eban 1992, p. 401; Hammel 1992, p. 139). The IDF now had to worry about an attack from the West Bank along with attacks from Egypt in Sinai and Syria in the Golan. Once again, the IDF requested immediate action.

While Israel was contemplating its response to the most recent provocation, events in Egypt seemed to have gone in a different direction. After canceling Operation Dawn, there is no evidence that Nasser reconsidered offensive action against Israel. Instead, Nasser stepped up diplomatic efforts to deescalate the crisis almost immediately after canceling Operation Dawn (Yost 1968, p. 315; Lacouture 1973, p. 305; Nutting 1972 p. 408). As early as May 25, Nasser considered sending the Straits issue to the International Court of Justice at the Hague and entertained a US suggestion that Egypt allow Israeli oil shipments through the straits (Stephens 1971, p. 489). The U.S. also raised the possibility of using ships from the 6th Fleet to escort Israeli ships not flying the Israeli flag, which Nasser also considered. The day after canceling Operation Dawn, U.S. representative Charles Yost arrived in Cairo on May 29 tasked with laying the groundwork for negotiations (Mor 1991, p. 369). The U.S. had hoped that Nasser might

be content with the remilitarization of the Sinai while agreeing to reopen the straits. Nasser told Yost that such an agreement was unacceptable since Egypt had merely overturned the injustices of the 1956 Sinai campaign (Stephens 1971, p. 479; James 2006, p. 116). Despite Nasser's seemingly intractable position, he did agree to negotiations and planned to send his Vice President to Washington on June 5 (Stephens 1971, p. 488).

Nasser's actions during this period are somewhat challenging to explain. Scholars suggest that Nasser was of two minds (Stephens 1971, p. 487; Oren 2002, p. 158; James 2006, p. 118): he believed that Israel was backed into a corner and would respond, but also believed that he might be able to achieve a bloodless victory by securing recent gains with negotiations. Nasser, however, was aware that the closed straits were unacceptable to Israel and represented a casus belli (Stephens 1971, p.479) It is, therefore, difficult to square the two positions in a satisfactory way. Nasser acted as though he was trying to consolidate the recent gains by diplomatic means (Khouri 1968, p. 247; Lacouture 1973, p. 305), despite the knowledge that Israel perceived those gains as unacceptable. Nasser publicly remarked that the crisis was subsiding and Egypt would not fire the first shot (Stephens 1971, p. 477). There was also sense that the crisis was subsiding in the military (Heikal 1968; James 2006, p. 116). In an interview with the British press on June 3, Nasser stated that the crisis was passing and war was unlikely (Stephens 1971, p. 480). There seems to be domestic tension in Egypt as 'Amer continued to push the military towards a war that Nasser was not particularly eager to fight. Israel had a very different sense of Egyptian intentions and the threat that Egypt posed in early June.

In Israel, political pressure mounted for Eshkol to take military action after the Egypt-Jordan pact was announced. The Prime Minister had held off requests to enlarge

the government or impose an emergency government, but the calls for a change in government could not be silenced any further. The military and government opposition parties demanded that Eshkol appoint a Defense Minster (Rabin 1979, p. 96; Segev 2005, p. 311). After much debate, and against his wishes, Eshkol appointed Moshe Dayan Minister of Defense on June 1. Dayan, a member of David Ben-Gurion's Rafi party, strongly supported action, but also recognized the importance of maintaining good relations with the U.S. (Dayan 1979).

The Cabinet members met with IDF leaders on June 2. The meeting began with briefings from the IDF intelligence chief General Aharon Yariv and Chief of Staff Rabin (Dayan 1976, p. 338). Yariv once again indicated that an Egyptian attack was imminent and stressed the need to act immediately (Dayan 1976, p. 343). His sense of concern was almost certainly heightened by the recently announced Egypt-Jordan defense pact, but he offered no new or current intelligence on military activity in support of his assessment (Segev 2005, p. 323). Rabin's briefing was similar in character. Rabin also indicated that action was needed. He noted that Egyptian positions were defensive, but that could change. He went further to elucidate a general warning, but he did not elaborate on the source of the warning or provide detailed evidence about changes in military posture that might signal an attack (Segev 2005, pp. 324-25). The ensuing question and answer period was not timid. Cabinet members asked a series of pressing questions, but no one challenged the origins or justifications for the underlying threat (Segev 2005, pp. 324-25). The Cabinet ministers, probably concerned about the recent Egypt-Jordan pact, ignored the vague origins of the threat and focused their questions on the logistics of an

Israeli attack, such as the time that northern settlements would be exposed to Syrian attack while the air force neutralized Egyptian targets.

Concerns about the U.S. position were redressed on June 2. Meir Amit, head of the Mossad, had flown to Washington to meet with contacts. He reported that the U.S. was not willing to commit to reopening the straits (Segev 2005). The international force they hoped to assemble had not materialized, and they were not prepared to use the 6th Fleet unilaterally (Rabin 1979, p. 95; Eban 1992, p. 405). This meant that Israel had waited for nothing. There was a second important piece of information gained from Amit's meetings with CIA and national security officials. Many insinuated that Israel had to defend itself however it saw fit, and it possible that they were concerned that U.S. demands on Israel had made the crisis more difficult to defuse (James 2006, p. 121). Eshkol interpreted this as tacit U.S. support for action. In light of recent developments, Eshkol and Eban decided that the time for restraint had come to an end (Kimche and Bawly 1968, p. 154; Eban 1992, p. 401). The majority of the Cabinet members were ready to go to war. On June 3, Eshkol approved Rabin's plan to attack on June 5 (Rabin 1979, p. 97). The attack caught Egypt by surprise, and the ensuing war was a tactical victory for the Israeli military (Hammel 1992, p. 171).

Before applying the Bayesian and biased models to the relevant sub-cases as we did for the Cuban Missile Crisis above, we have to examine Israeli beliefs and the possibility of entrenchment. The results here are somewhat similar to the Cuban case whereby the Israelis seem to maintain a rational belief set throughout most of the crisis. At the very least, it would be difficult to argue that the Israeli decision-makers exhibited anchoring or entrenchment early on in any convincing way. The Israelis began with the

prior belief that Egypt had aggressive intentions toward Israel, which was reflected by the intelligence assessment in 1966-67. Concerns about Egyptian aggression were partially offset by the belief that Egypt was unlikely to initiate a full-scale war until after 1970, due to Israeli military superiority, but the Israelis were very much concerned that Nasser might take some aggressive action short of war. The Israelis were worried that Nasser might move to close the Straits of Tiran or bomb the Dimona nuclear facility. It seems reasonable to say that Israel felt threatened by Egypt and monitored their behavior closely prior to the May 14 mobilization. Israel had monitored and reacted to similar Egyptian mobilizations by calling up troops in the past.

The May 14 mobilization was not, on it's own, sufficient to bring a major change in Israeli beliefs about Egyptian intentions. At first, the Israelis believed that the marches through Cairo towards the Sinai were held in response to the annual Israeli Independence Day military festivities. They believed that Egyptian action warranted close monitoring but were not prepared to consider the mobilization a major threat. While they did maintain their prior beliefs about Egypt, which could be reflective of entrenchment, the analysis was based on the available evidence such as force disposition, the public nature of the action and the knowledge that there were still UN forces in the Sinai. They maintained this belief until May 16.

There were two major events that caused concern for the Israelis on May 16. First, Israeli intelligence revised the initial troop estimate. The initial assessment suggested that the Egyptians were moving a single division toward the Sinai, which was not nearly enough to attack Israel. On May 16, that estimate had been accurately increased to four divisions. The larger force meant that Egypt could potentially contemplate offensive

action against Israel. The second event was Egypt's demand that the UN troops withdraw from their positions in the Sinai. This meant there would be no buffer to stop the Egyptians from attacking. On May 16, two of the three assumptions that the Israelis had used to downplay the risk were off the table. Given this new information, it is not surprising that the Israelis believed that an attack could be imminent. Whether the magnitude of the belief change was rational or irrational is difficult to tell, but there is no evidence to suggest that the Israelis ignored this new information or exhibited entrenched beliefs at the outset of this period.

The Israeli belief that formed on May 16 hardened throughout the crisis, but the information that reached Israeli decision-makers during much of that time makes it difficult to argue that this hardening was irrational in any way. After Nasser asked the UN to withdraw, the Egyptians sent MiG jets to fly over the Dimona facility on May 17. While the exercise was aimed at gathering intelligence it was a major cause for concern among the Israeli decision-makers. Nasser escalated the crisis further on May 23 by closing the straits, an action that the Israelis claimed justified war. Between May 23 and May 25, Israeli intelligence observed further military action and believed that more divisions were being called up. Though they did not know it at the time, these latest military moves were probably associated with 'Amer's planned surprise attack Operation Dawn. It is not clear whether Israeli intelligence observed the action and sounded the warning siren out of habit, or they actually had some intelligence on Operation Dawn. The Soviet warning to the Egyptians the night before the May 28 scheduled attack, by way of the U.S., probably indicates that the Israelis may have had some more specific information. There is no clear evidence to support this assertion in either direction. While

Nasser canceled the surprise Operation Dawn plan, he also chose to escalate the crisis one final time. On May 30, Egypt and Jordan signed the defense pact. Given this barrage of Egyptian activity, it seems perfectly reasonable that Israeli beliefs about Egyptians intentions hardened.

There is some evidence that Israeli beliefs become entrenched in the last stage of the pre-war crisis from May 31-June 5. On June 1, Israeli leadership and the military felt the pressure of Nasser's latest escalatory act, the Egypt-Jordan defense pact. In light of the fact that enemies with treaty obligations surrounded Israel on all three sides for the first time since 1948, it is not surprising that the IDF heightened their warning and decision-makers were receptive (Dayan 1976, p. 343). There were, however, no major troop movements in Jordan or Egypt after signing the pact on May 30. There were also more recent signals that Egypt was not planning an attack. The first sign was Nasser's interviews with the foreign press, mentioned briefly above, where he stated that the crisis would end without war. This signal, however, could be viewed as cheap talk since there is little cost to making such statements, and a deception campaign associated with surprise attack might explicitly call for such a statement. The second sign was the disposition of the senior Egyptian military leadership prior to June 5. On the evening of June 4, the Chief of the Ground Forces Command General 'Abd al-Mushi Kamil Murtagi was on vacation (Bar-Zohar 1970, p. 176), Field Marshal 'Amer was at an all-night party in Cairo, Commander of the Air Force General Sidqi Mahmud was attending his daughter's wedding, and subordinates were seen in Cairo playing tennis and partying (Churchill and Churchill 1976, p. 75). The Chief of the Air Force Major General 'Abd al-Hamid al-Dugheidi was dismayed that all of the commanders were away from their

stations (Oren 2002). This to, however, is not necessarily a reliable sign of Egyptian intentions.[27]

While these first two signs classify as cheap talk, the last two are better classified as costly signals. The first of these costly signals was Egypt's decision on June 2 to participate in diplomatic negotiations in Washington starting on June 5. While some diplomatic maneuvers are considered cheap talk, it is difficult to make such an argument in this case. Nasser's biggest concern throughout the crisis was that each escalatory might trigger the U.S. to join with Israel and take action against Egypt. By agreeing to negotiations, Nasser had essentially tied his hands. Any surprise attack against Israel after he had explicitly committed to negotiations would almost certainly anger U.S. policymakers, all but ensuring that they would side with Israel against a deceitful Egypt. Anthony Nutting (1972) argues that the announcement of upcoming negotiations emboldened the Israelis to attack, lest their free hand be tied by U.S.-Egypt negotiations. There is no evidence to support this claim or that Israeli decision-makers even considered the upcoming negotiations announced on June 2. Consideration of U.S. diplomatic efforts, and how it might impact decision-making in Egypt, was noticeably absent from the accounts of the Cabinet meeting on June 2, as well as the memoirs of key Israeli decision-makers such as Eban (1992, pp. 402-409), Rabin (1979, pp. 96-99), and Dayan (1976, pp.338-349). The second costly signal was the disposition of Egyptian forces when the Israelis attacked on June 5. In light of the possibility that Egypt might avoid war with Israel, the High Command issued an order that no action should be taken that would provide Israel and excuse to attack (Bar-Zohar 1970, p. 176). The majority of

[27] Note that since most officers were only a few hours from their command posts, this could also be part of a deception campaign aimed at getting Israel to let its guard down. There is no credible information to suggest that this was the Egyptian intention.

Egyptian jets were lined up tip-to-tip in plain sight near the Egyptian airport, making them an easy target for the Israeli Air Force (Nutting 1972, p. 416; Hammel 1992, pp. 168, 171). While the Egyptian army was hardening their defensive positions, they were doing so along forward lines meaning that a penetrating attack would leave the Sinai vulnerable (Oren 2002, p. 93). The poor state of Egyptian defense, and particularly their decision to gather the bulk of the Air Force in a single location, did not reflect a state ready for war. Taken together, the costly hand-tying diplomatic measures, and the poor logistical state of the army, both suggest that the Israelis may have irrationally overestimated the Egyptian threat after May 31. It is difficult to argue that their worries were wholly unjustified, but there were pieces of useful information that might have caused them to reevaluate the assessment that an Egyptian attack was imminent from May 31-June 5.

There is an additional aspect that is worth noting towards the end of the pre-war crisis period. Israel had delayed offensive action earlier in the crisis precisely because they were worried about U.S. response. Eshkol believed that U.S. support of Israel was crucial, and the Cabinet chose to wait when it appeared as though the U.S. might act to open the straits (Rabin 1979, p. 91). Israel went forward with the attack only after the Cabinet was informed that there was no consensus for a multilateral force, the U.S. would not act unilaterally, and Amit reported that President Johnson did not explicitly warn Israel not to attack. The U.S. had a major impact on Israel throughout the crisis, but Israel never considered that Egypt might have the same concerns (Kam 2004, p. 12). Israel had specifically decided not to act when they were actively holding discussions with the U.S.,

but assumed that these same constraints would not be sufficient to delay an Egyptian attack.

It would be remiss, however, to assert that the Egypt did not contribute to the Israelis beliefs about their intentions and the possibility of war. Outside of the Middle East, Nasser had repeatedly maintained that Egypt would not fire the first shot and issued statements to that effect to the U.S., the Soviets and the UN (Stephens 1971, p. 480). The rhetoric aimed at Arabs within the Middle East had a very different tone, urging Arab unity and declaring that they would soon attack and destroy Israel (Cohen 1988, p. 10; Eban 1992, p. 353; James 2006, p. 117). By June 2 when Eban had changed his vote in favor of war, he believed that the rhetoric had far outpaced reasonable limits (Eban 1992, p. 395). This certainly contributed to the hardened belief that war was imminent, but it does not necessarily justify Israeli beliefs during this period given other available information.

The Bayesian and biased model predictions are applied to the four sub-cases in Table 6.2. If the target, in this case Israel, starts with the belief that the sender is sufficiently threatening, then the models have divergent predictions. The theory usually predicts that one course of action is more likely than the other. It is important clarify the rationale for coding Israeli beliefs about Egypt as threatening. After all, Israel did not initially believe that the mobilization on May 14 was a sign of impending attack. Despite the belief that Egypt was not going to attack, Israel was deeply distrustful and suspicious of Egypt, which is evidenced by the 1966-67 intelligence assessment concluding that there was no serious possibility for improved diplomatic relations with Egypt and that relations would continue to be hostile. At the time of the assessment, Israeli decision-

makers believed that Egypt might look to close the straits or bomb the Dimona nuclear facility even if they did not want to start a full-scale war. Israel might not have been concerned that the initial mobilization was the start of the war, but they were sufficiently suspicious to be monitoring Egyptian activity closely both before and after May 14. For these reasons, Israeli beliefs about Egypt are coded as threatening at the outset of the crisis. Israel might not have believed an attack was imminent prior to the crisis, but they certainly did not believe that the Egyptians were docile or had benign intentions toward them.

The other important assumption to applying the biased model is that the sender, Egypt, believes that the target might be irrationally over-suspicious from entrenched beliefs. The target need not be entrenched, there only needs to be some probability. Given the historical animosity between the two countries, and Israel's security situation, it is difficult to imagine that Nasser did not recognize the possibility that Israel might be acutely sensitive to any provocations. In a speech after the crisis, Nasser noted that he expected Israel to be acutely sensitive to any provocations (Mor 1991). Analysts should be cautious about this statement since this took place after the crisis, but it is difficult to imagine that Egypt did not recognize the possibility that Israel might harbor irrational over-suspicion. Combined with Israel's prior belief that Egypt was threatening, this is the type of situation where the biased model predicts that public acts are more likely than expected in the Bayesian model.

Table 6.2: The 1967 War Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
Prior to May 14:					
Target's Belief			Aggressive; Full-scale War Unlikely; Closing Straits or Bombing Dimona Likely		
Evidence of Entrenched Beliefs			No		
Details of Entrenched Beliefs			1966-7 Assessment		
Sender Concern About Target Beliefs			High		
Proposition 2 (p>p*):					
Sender Use of Surprise[1]	Surprise or Violent Act Likely	Public Act Likely	Public Mobilization	0	1
Proposition 3 (p>p*):					
Target Response when Bayesian[2]	Mix	Wait	Wait; Monitor Closely	0	1
Target Response when Biased[2]	-	Mix	-	-	-

Continued…

Note: The table shows mostly divergent predictions. The biased model appears to have better explanatory power. See the footnotes or text for details on coding explanations.

1. If the Egyptians are perceived to be threatening, and they recognize the possibility that Israel might be irrationally over-suspicious, the biased model predicts that public challenges are more likely than the Bayesian model predicts. The Egyptians trigger the crisis publicly, meaning the biased model better captures Egyptian behavior.

2. If the Egyptians are perceived to be threatening, the Bayesian model predicts the target adopts a mixed strategy, the biased model with the Bayesian type predicts the target waits, and the biased model with the biased type predicts a mixed strategy. Since Israel does not appear biased at this point in the crisis, the Bayesian model should be compared to the biased model with the Bayesian type. Since Israel does not take major action after the mobilization, waiting in the biased model better captures behavior.

Table 6.2 (cont'd): The 1967 War Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
The Initial Mobilization, May 14-15					
Target's Belief			Aggressive; Full-scale War Unlikely; Closing Straits or Bombing Dimona Likely		
Evidence of Entrenched Beliefs			No		
Details of Entrenched Beliefs			Careful Monitoring		
Sender Concern About Target Beliefs			High		
Proposition 2 (p>p*):					
Sender Use of Surprise[3]	Surprise or Violent Act Likely	Public Act Likely	Request for UN Withdrawal	0	1
Proposition 3 (p>p*):					
Target Response when Bayesian[4]	Mix	Wait	Wait; Monitor Closely	0	1
Target Response when Biased[4]	-	Mix		-	-

Continued…

3. If the Egyptians are perceived to be threatening, and they recognize the possibility that Israel might be irrationally over-suspicious, the biased model predicts that public challenges are more likely than the Bayesian model predicts. The Egyptians publicly request removal of UN forces during this stage of the crisis, meaning the biased model better captures Egyptian behavior.

4. If the Egyptians are perceived to be threatening, the Bayesian model predicts the target adopts a mixed strategy, the biased model with the Bayesian type predicts the target waits, and the biased model with the biased type predicts a mixed strategy. Since Israel does not appear biased at this point in the crisis, the Bayesian model should be compared to the biased model with the Bayesian type. Since Israel does not take major action after the request, waiting in the biased model better captures behavior.

Table 6.2 (cont'd): The 1967 War Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
Revised Troop Estimates, Closing the Straits, Surprise Plans, and Jordan May 16-30					
Target's Belief			Aggressive; Attack Imminent		
Evidence of Entrenched Beliefs			No		
Details of Entrenched Beliefs			Shift in Belief of Attack		
Sender Concern About Target Beliefs			High		
Proposition 2 (p>p*):					
Sender Use of Surprise[5]	Surprise or Violent Act Likely	Public Act Likely	Preparation for Surprise Attack; Jordan Pact	1	0
Proposition 3 (p>p*):					
Target Response when Bayesian[6]	Mix	Wait	Wait For US Approval; Mobilize Troops	0	1
Target Response when Biased[6]	-	Mix		-	-

5. During this stage, the Egyptians prepare for a surprise attack against Israel, deviating from their prior strategy. If the Egyptians are perceived to be threatening, and they recognize the possibility that Israel might be irrationally over-suspicious, the biased model predicts that public challenges are more likely than the Bayesian model predicts. The Bayesian model, predicting a higher probability of attempted surprise, better captures Egyptian behavior.

6. If the Egyptians are perceived to be threatening, the Bayesian model predicts the target adopts a mixed strategy, the biased model with the Bayesian type predicts the target waits, and the biased model with the biased type predicts a mixed strategy. Since Israel does not appear biased at this point in the crisis, the Bayesian model should be compared to the biased model with the Bayesian type. Since Israel decides to wait for the U.S. and does not take major action despite the belief that Egypt is preparing for a surprise attack, waiting in the biased model better captures behavior.

Table 6.2 (cont'd): The 1967 War Model Comparison

	Predicted Bayesian Outcomes (z=0)	Predicted Biased Outcomes (z>0)	Observed Outcomes	Bayesian Score	Biased Score
Mixed Messages, Diplomacy, and Easy Targets, May 31-June 5					
Target's Belief			Aggressive; Attack Imminent		
Evidence of Entrenched Beliefs			Yes		
Details of Entrenched Beliefs			Ignores Upcoming Negotiations and Reduced Military Posture		
Sender Concern About Target Beliefs			High		
Proposition 2 (p>p*):					
Sender Use of Surprise[7]	Surprise or Violent Act Likely	Public Act Likely	Public, Though Noisy, Acts to Deescalate	0	1
Proposition 3 (p>p*):					
Target Response when Bayesian[8]	Mix	Wait	Attack	0	-
Target Response when Biased[8]	-	Mix		-	0
Total				1	6
Statistical Analysis on 8 Observations					
t-Statistic	-2.376				
P(T<=t): Two-Tail	0.049				
t Critical Value: Two-Tail	2.365				

7. If the Egyptians are perceived to be threatening, and they recognize the possibility that Israel might be irrationally over-suspicious, the biased model predicts that public challenges are more likely than the Bayesian model predicts. The Egyptians make public attempts to deescalate the crisis after they abandon their surprise plan (Operation Dawn). The attempts to publicly deescalate the crisis means the biased model better captures Egyptian behavior.

8. Neither model receives any points. The Bayesian model predicts a mixed strategy, while the biased model with the biased type (since the Israel is suffering from entrenched beliefs according to the argument above) predicts an identical mixed strategy. Neither model can claim to provide a better explanation than the other.

In the period prior to May 14, the Israelis view Egypt as threatening, and while they believed Egypt will not initiate a full-scale war, they are concerned that Nasser might close the straits or attack Dimona. While there is a chance that Israel exhibited some bias when Egypt mobilized, there is no clear evidence to suggest they did. Egypt opted to trigger the crisis by publicly mobilizing their force and marching troops through Cairo towards the Sinai. Since the public act is more likely in the biased model than the Bayesian model, the biased model seems to better capture this challenge decision. The Bayesian model predicts that the target adopts a mixed strategy. The biased model predicts that a biased type selects the same mixed strategy as in the Bayesian model, but the Bayesian type always waits. Israel chose not to take immediate action given the Egyptian mobilization, so the biased model with the Bayesian type outperforms the Bayesian model. In each instance the biased model receives a score of 1 and the Bayesian model a score of 0.

The model comparison results are very similar for the second sub-case for May 14-15. Again, Israel views Egypt as threatening, and while they believe Egypt will not initiate a full-scale war, they are concerned that Nasser take the limited measures discussed. The Israelis carefully monitor the Egyptian mobilization, but there is no evidence that the initial mobilization altered the prior beliefs significantly since the intelligence estimate indicated only one division was moving. During this period, the Egyptians chose to escalate the crisis publicly by demanding the withdrawal of the UN troops, rather than directly initiating conflict with Israel. Since public challenges are more likely in the biased than the Bayesian model, the biased model performs better. Again, the Bayesian model and the biased model with the biased type predict that target will

adopt the same mixed strategy. The biased model with the Bayesian type predicts that target will always wait, which better describes Israel's behavior during this time. In both instances the biased model receives a score of 1 and the Bayesian model a score of 0.

The third stage of the crisis, May 16-30, differs from the first two. After the Egyptians demand the UN withdrawal, the Israelis recognize that the mobilization is larger than they initially estimated. Israeli intelligence begins to predict an attack is imminent. The Egyptian fly over of Dimona on May 17 and the closing of the straits on May 23 further aggravated Israeli fears, which grew more acute as Egypt planned the surprise attack Operation Dawn. The preparations were complete on May 25, but Nasser canceled the operation on May 28. Despite canceling the surprise, the Egyptians escalated the crisis once again by announcing the defense pact with Jordan. In this period, the Bayesian model of sender behavior outperforms the biased model. Surprise continues to be more likely in the Bayesian model than the biased model, and since Egypt spends much of the period preparing for the surprise, the Bayesian model is coded 1 and the biased model 0. The biased model of target behavior, however, continues to offer a closer explanation of behavior than the Bayesian model. Despite the revised assessment that an attack was imminent, Israel chose to wait, in large part hoping to get support from the U.S. The Bayesian model and biased model with the biased type both predict the same mixed strategy. The biased model with the Bayesian type predicts that the target waits rather than takes action. Since there is no clear evidence that Israel has entrenched beliefs, especially since they receive repeated aggressive signals and have more accurate intelligence about the Egyptian mobilization than the U.S., the biased model has better explanatory power. The biased model is coded 1 and the Bayesian model 0.

The final stage of the crisis is May 31-June 5. This is the one point in the crisis where the Israelis may exhibit entrenched beliefs given some of the diplomatic and military signals sent by the Egyptians. The Egyptians continue to adopt public measures of crisis behavior during this period. They continue their bellicose language, but also agree to negotiations with the U.S. Any plans for surprise attack seem to have been abandoned on May 28. Since public acts are more likely in the biased rather than the Bayesian model, the biased model is coded 1. Since Israel is coded as having some level of entrenched beliefs during this stage of the crisis, the two models predict the target adopts the same mixed strategy. Since both predict that the target mixes their strategy, and the probability with which both mix their strategy is the same, neither one outperforms the other or wholly captures the Israel's behavior on June 5. Both models receive a score of 0.

When we sum up the raw numbers in this case, the Bayesian model scores a one and the biased model scores a six. In this case, where the models often predict different results since the target views the sender as threatening, the biased model seems to have significantly outperformed the Bayesian model. The results are statically significant, meaning there is little chance that difference in score was due to random numerical error. Both the one- and two-tailed tests show that the difference is statistically significant. The likelihood the biased and Bayesian results are different is 95.1% on the two-tailed test.

In this case, the biased model of decision-making faired very well relative to the Bayesian model. This suggests that the model has strong explanatory power in instances where the target is irrationally over-suspicious, or believes the sender is threatening, and

the sender knows that there is some probability that a target may be irrationally entrenched when faced with a crisis.

Conclusions

The purpose of this case analysis was to evaluate the causal theory developed in the prior chapters. The cases in this chapter illustrate how entrenched beliefs can impact the behavior of the sender and target when triggering and responding in crisis situations. The cases were selected, in part, because there was variation in the dependent variables, specifically the crisis triggers, the target's sense of surprise and the target's reaction. Each case examined the relevant deductive propositions and explicitly compared the expected results of the biased model against a standard Bayesian prediction. A summary of these results, reordered by proposition, is in Table 6.3.

The summary table below shows that the biased and Bayesian models performed equally well when their predictions converge in proposition 1. In instances where the actors should behave the same, irrespective of the potential biases, there is no differentiation between the two models. These convergent results seem to get support from the case history. For most of the observations in the table, propositions 2 and 3, the two models predict marginal or completely different outcomes. In both instances, the biased model appears to outperform the Bayesian rival.

For proposition 2, the Bayesian receives an aggregate score of two, scoring one point in each of the two cases. It is important to note, however, that the two models predict convergence in Cuba from October 16-22. Both models offer identical deterministic outcomes, which closely track the actual case history. The other Bayesian

point comes from the 1967 War during the period that Egypt prepares for the surprise attack Operation Dawn, which Nasser eventually cancels since the Israelis and the U.S. get wind of the proposed surprise. The raw tally indicates that the biased model outperforms the Bayesian with a score of five to two.

Table 6.3: Summary of Model Tests

	Bayesian Score	Biased Score
Proposition 1: Identical Pure Strategies		
Cuba: Prior to October 16		
Sender Use of Surprise	1	1
Target Response	1	1
Total	2	2
Porposition 2: Senders Use of Surprise		
Cuba: October 16-22	1	1
Cuba: Post October 22	0	1
1967: Prior to May 14	0	1
1967: May 14-15	0	1
1967: May 16-30	1	0
1967: May 31-June 5	0	1
Total	2	5
Proposition 3: Target Response		
Cuba: October 16-22	0	1
Cuba: Post October 22	0	0
1967: Prior to May 14	0	1
1967: May 14-15	0	1
1967: May 16-30	0	1
1967: May 31-June 5	0	0
Total	0	4

The results are similar for proposition 3. The biased model scores four points, and the Bayesian model scores zero. One of the problems evaluating this proposition is that the Bayesian model, and the biased model with a biased target type, predict identical mixed target strategies, making it difficult to argue that one model performs better than the other. The differentiating characteristic is that the target in the biased model would

only use the mixed strategy, and therefore act, when they had entrenched beliefs. If the target were the Bayesian type they would wait. Therefore, the Bayesian model predicts that a Bayesian would act probabilistically, whereas the biased model predicts that only a biased target would act probabilistically. This is an important difference, but one that does not really show up in the model scoring. In aggregate, the biased model outperforms the Bayesian model once again.

The final interesting aspect in the chapter was the approach to examining entrenched beliefs. Earlier in the text, and above in this chapter, we discussed the difficulty of establishing the existence of psychological biases. Simply claiming that bias exists is unconvincing, and the burden of proof should be on the one arguing that actors were biased in some way. We argued that one way to deal with entrenched beliefs was to trace the process over time, since entrenchment is explicitly about ignoring newer information relative to older beliefs or information. In strategic contexts, this is not always straightforward since new information may be supplied for deceptive purposes or may be cheap talk. It is also possible that new information is consistent with prior beliefs, thereby strengthening or reinforcing the Bayesian prior. This complicates the analysis significantly. In both case studies, however, there appears to have been new evidence that should have warranted a reexamining of prior assumptions at the very least, even if it did not warrant wholesale belief change or reinforced prior beliefs.

Chapter 7
Conclusions, Qualifications and Implications

Two general goals were outlined in the beginning of this project. The first was to improve our understanding of the relationship between foreign policy decision-making and psychological bias. The project examined these issues by formalizing some commonly addressed features in a novel and rigorous, developing a mathematical definition of a specific type of bias called anchoring bias or deeply entrenched beliefs[28] and integrating anchoring bias into a competitive context. When people anchor on their prior beliefs they irrationally discount new information relative to their preexisting beliefs. The mathematical formula asserted that biased actors anchor, or irrationally overweigh their prior beliefs, relative to perfect rational Bayesian actors. The mathematical formula that captured anchoring bias was sufficiently general that it could be used to represent entrenched beliefs for any one possible prior belief that an actor might have at the outset of an interaction. This general, formalized approach to entrenched beliefs was then incorporated into a deductive model with two players where one of them might have entrenched beliefs and the other player knew that their opponent might be biased. This rigorous approach helped to develop deductive expectations about the behavior of rational and biased actors,[29] and a comparison yielded some

[28] Note that when one overweighs or over-relies on their prior beliefs, they are exhibiting anchoring bias.
[29] Where biased actors are those with deeply entrenched beliefs.

counterintuitive propositions underappreciated in work on psychological bias in foreign-policy decision-making.

The second goal was developing a more robust theory of surprise[30] in international politics, and specifically the relationship between surprise and psychological bias. The project developed a series of deductive models for surprise that began with a simple form of imperfect information whereby a target was unsure if their opponent was preparing for a surprise attack. Incomplete information, where the target does not know about their opponent's preferences over outcomes, was then added to the basic model of surprise. Then the project compared a model of surprise with a perfectly Bayesian rational target to a model with a potentially biased target. This exercise helped to clarify the relationship between surprise and entrenched beliefs, providing set of propositions that were empirically examined later in the project quantitatively and qualitatively. Both empirical approaches showed that the model with a potentially biased actor appears to have strong explanatory power relative to a model with perfectly rational actors.

The primary conclusions stated and examined below were derived using a series of formal models, which were broadly outlined in the two paragraphs above. While some of the limitations of the models will be discussed in greater detail, it is important to state some of the explicit assumptions needed to derive the primary results. The series of games all involve two players with divergent preferences acting in a sequential fashion. The sequential movement is often referred to as an extensive form game. The games, as is common in extensive form games, are one-shot plays where there is no repeated interaction after the game concludes. All of the models address surprise and take a

[30] Surprise is an outcome in the model that results when a sender challenges privately (attempts surprise), and the target chooses to wait rather than take defensive action.

common structural form. As discussed above, the modeling section begins with a basic model of surprise driven by imperfect information. Throughout the project, surprise is treated as a strategic outcome rather than a reaction to unexpected events. An extension of the model then adds incomplete information so that the actors' beliefs influence equilibrium outcomes. Once beliefs are incorporated, then the belief bias, anchoring, can be examined.

In the rest of this chapter, we examine the theoretical and empirical conclusions, discuss some of the limitations of the project, and then discuss some of the policy implications arising from this analysis.

Conclusions

Bias in Decision- Versus Game-Theoretic Settings

Work in foreign policy, surprise and intelligence often implies, or argues outright, that psychological biases increase the likelihood of suffering policy failures. This argument is found in the literature focusing on psychological biases as the cause of analytical error (Heuer 1999), poor policy outcomes (Levy 1984; Snyder 1984; Jervis 1976), or surprise (Betts 1978; Kam 1988). Psychological biases and wishful thinking are supposed to blind decision-makers when crafting policy, making it more likely that the policy is ill suited to the current environment. This extends into intelligence literature where psychological biases make analysts and policymakers misinterpret signals and miss cues about changing conditions or impending attacks (Betts 1978; Levite 1987; Kam 1988). In short, the conventional tenet about psychology in foreign policy dictates that biases contribute to poor policy outcomes.

There are, however, a few important deviations from these arguments. The most notable is Thomas Schelling's (1960) work on the rationality of irrationality, which was addressed in some detail in the introductory chapter. Schelling argued that actors may have incentives to appear irrational, or more accurately act in apparently irrational ways, in order to constrict their opponent's behavior. In some ways, this project actually provides a truer deductive proof for the "rationality of irrationality" since Schelling's actor remained instrumentally rational throughout the game.[31] Schelling argued that actors who pre-committed themselves to a certain action, particularly one that appeared irrational, could derive benefits by making their opponents adopt conciliatory strategies if the commitment were credible. Schelling's argument, however, does not explicitly incorporate an irrational actor. Schelling showed that actors might benefit by committing themselves before a game began, and offered logical proofs to show that seemingly irrational commitments could be credible under certain conditions. Schelling advocated two strategies discussed in the introduction, burning bridges and randomization. Both allow actors to make credible commitments when commitment would be otherwise irrational and unsustainable. The next section examines ways that this project extends Schelling's work.

Schelling is not the only scholar to note that foreign policy decision-makers and analysts may benefit from psychological biases. Writing about intelligence, Heuer (1981) argues that biases may make actors impervious to attempted deception. This reduces the incentive to use deception and reduces the likelihood that it is effective when used. The

[31] This is discussed by Zagare and Kilgour (2000) who define instrumental rationality as an economist's definition of rationality dictating that actors can compare any two possible outcomes to determine a relative preference and that the preference ordering be transitive (i.e. if an actor prefers option *a* to *b* and option *b* to *c*, they cannot prefer option *c* to *a*).

common theme linking Schelling's and Heuer's argument is the presence of a strategic or game-theoretic environment rather than a decision-theoretic environment. Decision theory focuses on the optimality of a single actor's decision based upon factors such as preferences and the probability over outcomes. Some of the work on bias in foreign policy implicitly takes a decision-theoretic perspective whereby a unitary actor, or bureaucratic group, evaluate the available options in an attempt to adopt the optimal policy. In this context, psychological biases may distort probability assessments leading the decision-makers to adopt a sub-optimal policy outcome by mistake. This decision-theoretic context, however, ignores the strategic contexts of international politics whereby actors usually condition their polices on the behaviors, preferences or beliefs of other actors.

While biases may distort probability assessments and lead to sub-optimal policies in decision-theoretic contexts, an assumption used when formalizing entrenched beliefs here, it does not necessarily mean that these same biases are sub-optimal in strategic contexts where outcomes are contingent on the behavior of others. The tendency to think about psychological bias in decision-theoretic contexts assumes away a vital competitive or interactive aspect in international politics and produces a hole in our understanding of foreign policy decision-making. Just as Schelling and Heuer Jr. rely on strategic contexts, this project provides a deductive proof that shows how bias may have some unexpected effects in strategic contexts under certain conditions. We turn to those unexpected effects and conditions next.

Some Counterintuitive Effects of Bias

In the deductive game-theoretic models developed here, anchoring bias is neutral when it results in irrational under-suspicion[32] and beneficial when it results in irrational over-suspicion.[33] This is a very different result from the conventional tenet based on decision-theoretic logic arguing that biases contribute to poor policies or surprise. The game-theoretic conclusions substantively differ from the conventional decision-theoretic argument, because the former explicitly takes the opponent's optimal strategy into account whereas the latter does not. An actor potentially suffering from entrenched beliefs may find they better off than a perfectly rational actor, if their opponent knows that the actor might be biased. All the opponent needs to know is that there is a probability that the actor has entrenched beliefs, they do not need to know whether the actor is biased or not. The possibility that the actor has entrenched beliefs may be sufficient to alter the opponent's optimal strategy under certain conditions. The results and necessary conditions are briefly repeated below, we continue to use the convention of calling the potentially biased player the "actor" and the other player the "opponent."

If the opponent has no prior knowledge about the actor's biases, meaning they believe the actor is always rational, the actor's bias has no impact on the equilibrium results. Since the opponent adopts a strategy that makes the actor mathematically indifferent between their choices, it does not matter whether the actor's bias leads to inaccurate assessments or poor decision-making. The expected payoff for the biased actor does not change since their opponent's rationally optimal strategy cancels out the effect

[32] Irrational under-suspicion occurs when an actor has a prior belief that their opponent is non-threatening and anchors on that belief.
[33] Irrational over-suspicion occurs when an actor has a prior belief that their opponent is threatening and anchors on that belief.

of bias. In a game-theoretic construct, bias only matters when an opponent has some prior knowledge about an actor's anchoring biases or potential anchoring biases.

When an opponent knows that the actor may anchor on prior beliefs, the opponent may have incentive to change their strategy. The opponent does not change their strategy when the actor believes that their opponent is non-threatening.[34] An actor with entrenched beliefs and the initial belief that their opponent is non-threatening could be irrationally under-suspicious. One might imagine that the initiating opponent might be able to take advantage of that under-suspicion, but the deductive logic shows that the opponent already exploits the actor's belief that the opponent is non-threatening by attempting surprise as much as possible. They are unable to exploit the irrational under-suspicion any further. In other words, the opponent is already taking advantage of every opportunity to exploit the actor, and there is nothing more they can do even if the target has entrenched beliefs that would result in under-suspicion. In this instance, the biased actor is no better or worse off than if they were perfectly rational. Bias is neutral or irrelevant in its affect on the target's success.

When entrenched beliefs lead an actor to be irrationally over-suspicious, which is to say they start with the prior belief that the opponent is threatening, the actor derives a benefit from their bias. An opponent aware that an actor may be irrationally over-suspicious is best off changing their strategy, adopting a less provocative plan. In this instance, a target that is irrationally over-suspicious is more likely to take defensive action, thereby undermining the opponent's strategy. The opponent's best response is to act conservatively as long as they know that there is a chance that the actor is biased. In

[34] Note that non-threatening is label used to describe a situation where the prior belief that the opponent is aggressive is below the endogenously determined critical value ($p<p^*$).

the models developed above, the opponent, or sender, attempts surprise less often when the actor may be biased in order to ensure that the actor does not thwart too many surprises. The net effect is that the actor's bias is often beneficial under these conditions.

This project reaches some of the same conclusions that Schelling did, but it also deviates from and extends it in a few important ways: 1) Schelling's actor feigned irrationality for rational purposes, but actors with over-suspicion from naturally occurring anchoring bias can derive similar strategic benefits if an opponent recognizes the possibility of bias; 2) while Schelling showed that seemingly irrational commitments could be made credible by randomization or bridge burning, this project shows that the mere *possibility* of irrational over-suspicion is sufficient to capitalize on the benefits irrationality; 3) when there is a possibility that actors are actually irrational, there may be limits to the benefit of irrationality (which are erased when the probability of anchoring gets too high); 4) irrational under-suspicion does not have any meaningful impact on the target's utility or likelihood of being of being surprised; and 5) Schelling talked about deriving rational benefits from feigning irrationality whereas the analysis here showed that rational types[35] may achieve slightly higher utility than irrational types, but irrational types were less likely to suffer surprise.

Since decision-theoretic approaches to foreign policy do not consider how one actor's biases impact the behavior of others, it is not surprising that the conclusions from decision-theoretic and game-theoretic contexts differ. This project clearly echoes some of the key sentiments that Schelling made, but also offers some additional insights. Some of these are derived from the extension of irrationality to the topic of surprise. Given that the

[35] In the model developed in chapter 4, the actor could be a rational type or a biased type.

models specifically address the issue of surprise, it makes sense to reexamine the relationship between surprise and psychological bias next.

Surprise, Necessary Conditions and Bias

Scholars have come up with an array of reasons to explain why states suffer intelligence failures or get surprised. Some of the most commonly cited reasons include noisy environments that make it difficult to decipher real information from noise (Wohlstetter 1962), explicit attempts to deceive (Whaley 1969), organizational impediments (Betts 1978; George 1980) and psychological impediments (Betts 1978; George 1980; Heuer 1999). The earliest of these commonly cited explanations (Wohlstetter 1962) explicitly considered surprise and intelligence failure in game-theoretic or strategic contexts. These later explanations focus more decision-theoretic environments, or at least environments where the strategic constraints come from internal bureaucratic sources rather than external sources.

The models of strategic surprise developed here show that psychological or organizational impediments are not necessary conditions for surprise or intelligence failure. There were three models of surprise developed above. The basic model incorporated a simple form of imperfect information where a target did not know if an initiator was preparing for surprise or not. Under certain preference assumptions, specifically when the initiators payoff for mixing between attempted surprise and maintaining the status quo is preferable to the payoff for a public standoff, surprise is expected some of the time. In a perfectly rational setting, the initiator mixes between attempting surprise and preserving the status quo, which we call adopting the ambiguous

strategy. Since the target cannot differentiate the two actions in the ambiguous strategy, the target responds by mixing acting and waiting. In the rational model with a simple information constraint, the status quo, false alarms, thwarted surprises, and successful surprises are all possible outcomes. Bias is neither necessary nor sufficient to explain surprise or intelligence failure.

This same result is true when an additional information constraint, incomplete information about the initiators preferences over outcomes, is added to the model. Once again, surprise is a probabilistic outcome of a perfectly rational interaction. When a target believes that an initiator is non-threatening, the model predicts status quo and successful surprise as possible outcomes. If the target initially believes that the initiator is threatening, they will mix between waiting and acting, while the initiator may use public challenges, attempted surprise or preserve the status quo. Once again, psychological bias or organizational impediments are neither necessary nor sufficient to explain surprise.

In the deductive analysis, irrational under-suspicion, the entrenched belief that a target is non-threatening, did not increase the likelihood of being surprised, because the initiator attempted surprise as much as possible without prompting the target actor to take action. This required playing a strategy that made a Bayesian actor indifferent, and did not allow the initiator to capitalize on the actor's irrational under-suspicion. A rational actor that believes an opponent is non-threatening will not take defensive action since they do not believe it is necessary, and the opponent specifically acts to cultivate this belief. The irrational under-suspicion resulting from anchoring bias was irrelevant in the deductive outcomes.

Irrational over-suspicion decreased the likelihood that targets would be surprised, and since the initiator had less incentive to attempt surprise, they adopted a less provocative strategy. Ironically, when initiators adopt less provocative strategies, it is perfectly rational actors that are more likely to get surprised and biased types that are more likely to adopt an optimal response. In the next section we discuss the performance of the biased model and the associated propositions in the two empirical tests.

Empirics: Comparing The Bayesian and Biased Models

The empirical analyses suggest that the model incorporating entrenched beliefs may capture state behavior reasonably well. Empirical examination of unobservable phenomena like psychological biases, particularly in the uncontrolled context of international crises, can be difficult to say the least. Both the quantitative and qualitative tests were constructed carefully in order to isolate relevant variables and case events.

The quantitative study relied on the International Crisis Behavior dataset, using existing dataset variables to test a proposition dealing with the initiator's strategy. Proposition 2 in chapter 4 stated that initiators are less likely to attempt deception against potentially entrenched targets than they were against Bayesian targets. The quantitative analysis relied on instrumental variables, variables that could be representative of the primary variables, to test the proposition about the initiator's behavior. A variable for protracted rivalry was used to identify the possible presence of irrationally entrenched beliefs, relying on the argument others have made that biases like entrenchment are common in these rivalries (Bar-Tal 2000; Thies 2001; Mor 2004; Hassner 2007). Rivals also have more experience with one another in conflict situations, and it is important to

control for prior experience if one is to separate irrational entrenchment from Bayesian rational entrenchment. A set of rational belief variables were designed to capture the past experience of crisis actors and were incorporated in the quantitative analysis. The dependant variable for the study was the initiators decision to use violence to trigger the crisis. A logit analysis showed that initiators were more likely to use violence to trigger crises as the probability that prior crises turned violent increased. Consistent with the predictions of the biased model, when controlling the violence in past crises, initiators were less likely to trigger a crisis violently against protracted rivals. While the use of instrumental variables makes this an imperfect test, it does provide illustrative support for the proposition.

The results of qualitative analysis also offered support for the biased model of surprise developed here. There are two case studies: the Cuban Missile Crisis in 1962 and the Arab-Israeli War in 1967. In the Cuban Missile Crisis case, the Soviet Union is treated as the initiator and the U.S. is a target that starts with belief that the Soviet aid to Cuba is non-threatening. Over the course of the crisis, U.S. decision-makers come to believe that Soviet intentions in Cuba are aggressive and eventually wind up anchoring on beliefs about nuclear weapons in Cuba and Soviet intentions. In the Arab-Israeli War case, Egypt is treated as the initiator and Israel as the target. Israel starts the crisis with the belief that Egypt harbors aggressive intentions toward Israel, and may even carry out a limited offensive move, but they do not believe a full attack is likely. As they gather information about Egyptian plans, they come to accurately believe that an attack is imminent, but do not take immediate action in the hopes that the U.S. might defuse the

crisis. In the later stages of the crisis, Israel anchors on the belief that an Egyptian attack is imminent despite some signals that warranted reconsideration of this belief.

The two cases are used to explicitly compare the predictions from the standard Bayesian model and the model where the target is potentially biased. The qualitative analysis broke the two cases into decision-making sub-cases, and identified the predictions each model offered for each sub-case. The events and pertinent beliefs in each crisis phase or sub-case were coded and explained. Then the predicted results from the two models were compared against the sub-case events. Both the Bayesian and potentially biased models performed well in the instances where the results of the models converge, specifically instances where the target believed the initiator was non-threatening. In all three pertinent sub-case observations, both models offer deterministic predictions that are correct. When the predictions of the models diverge, the potentially biased model outperformed the Bayesian model. Summary results in the conclusion of chapter 6 showed how the two models faired based on the major deductive propositions. A tally of proposition 2, focusing on the sender's reduced likelihood of using surprise given a potentially biased target, showed that the potentially biased model scored five of a possible six and the Bayesian model scored two. The results for proposition 3, the target's response and crisis outcome, favored the potentially biased model four to zero out of six possible observations.

In sum, the empirical analyses in this project do not provide confirmatory evidence for the logically deduced propositions, but they do suggest that the major deductive propositions may be empirically valid. This section started by discussing the

difficulty associated with empirically testing the deductive propositions, and below we will discuss some of the empirical and theoretical limitations associated with the project.

Qualifications

Before discussing some of the policy implications of this work, it is important to confront some of the theoretical and empirical limitations of this research. In this section, we briefly discuss the major qualifications to the conclusions presented above.

Difficulty of Studying Biases

One recurrent theme in both the theoretical and empirical sections was the difficulty associated with studying psychological bias in uncontrolled settings like international crises. Most psychological studies of bias take place in controlled or laboratory setting such that subjects can be manipulated in specific ways to test the impact of different treatments (Tetlock 1998a). This is impossible in historical study, and scholars are only left with actions, events, correspondences and the like to assess an actor's rationality at the time they made a decision. This means that explanations of events that rely on psychological biases may appear compelling, but alternative explanations that do not rely on biases may be easier to prove.

This adds an additional constraint, one that is warranted, on those advocating psychology-based explanations of behavior. Scholars cannot just declare that actors suffered from some sort of psychological bias, they must painstakingly illustrate, since proof is almost impossible, the existence and impact of the bias under consideration. Simply observing a situation where actors did not use all of the information available,

give it equal weight, or rely on a new piece of information that turns to be accurate, is not sufficient to claim that actors exhibited a psychological bias like anchoring bias. Chapter 3 explicitly compared perfect Bayesian belief updating to irrational belief anchoring. A perfectly rational Bayesian actor may discount or marginalize new information if their prior belief is sufficiently strong. The marginalizing or discounting of new information in this context is perfectly rational. Studies that focus on psychological bias have to be careful to compare rational information discounting or limited updating to the irrational activity they claim to observe. This was a key aspect of the empirical analysis in chapter 6, which explicitly compared the Bayesian and biased models in light of the case information.

Related to the unobservable nature of psychological biases, is the number of different biases that may weigh on any given actor during a set of events. Since psychological biases are difficult to observe, it is also difficult to tell which of the many biases an actor is suffering from. We turn to this qualification next.

Anchoring as the Primary Bias

Psychologists have identified many different types of cognitive and decision-making biases. Anchoring bias or entrenched beliefs, the primary focus of this paper, is only one of many different biases that might impact an individual or group. This means that decision-making may be impacted by many different biases at once. An individual may anchor on prior beliefs, suffer from ambiguity aversion whereby they irrationally overvalue guaranteed outcomes to lotteries, and may also suffer from a gambler's fallacy whereby they erroneously believe that random events are somehow correlated, all at the

same time. Anchoring may or may not be the strongest bias in any given situation. The number of biases and the combinations of biases make it very difficult to talk about bias or psychology in foreign policy decision-making without offering numerous qualifications. Further, there are group decision-making biases, like groupthink where all the members of a group gravitate towards the same belief.

While it is important to recognize that there are multiple possible biases that might be at work in a given situation, it would be intractable and unreasonable to incorporate every type of bias in a deductive model. Aside from making the model difficult to solve, the relative importance of different biases might well alter deductive results in nontrivial ways, making empirical verification reliant on being able to calculate the strength of every bias relevant in a decision. The exercise of incorporating all of these factors into a formal model is unlikely to yield robust meaningful results that are testable. It is possible that biases have different effects and that a model incorporating ambiguity aversion or gamblers fallacy might yield different results. This project focused on anchoring bias since it is commonly employed in foreign policy and intelligence literature (Jervis 1976; Betts 1978; Betts 1982; George 1980; Lebow 1981; Levy 1984; Snyder 1984; Kam 1988; Reiter 1996; Heuer 1999; Bar-Tal 2000) particularly when isolating reasons for failed policy and surprise. Another approach might call for modeling the effect of individual biases in the group settings common in foreign policy decision-making. This, however, would require softening the unitary actor assumption used throughout the project. We turn to this limitation next.

Unitary Actor Assumptions

This project followed in the footsteps of the initial realists and rational choice scholars that treated the state as a unitary or unified actor (Schelling 1960; Waltz 1979). Each player in the formal models developed here operates as a unitary actor, thereby assuming away organizational and bureaucratic dynamics that actually exist in reality (Allison 1971, 1999). The reason why many rational choice scholars of international relations have relied on unitary actor assumptions is that it simplifies the analysis of interstate interaction. Even more recent rational choice work that integrates domestic politics into models of international interaction (Fearon 1994; Milner 1997) does so in very stylized or simplistic ways. This work on the interaction between domestic and international politics has led to some important conclusions, but it often involves simplifying complex processes to make the analysis tractable.

One disjoint that arises from the use of the unitary actor assumption is that the entire state is treated like an individual suffering from anchoring bias, which is usually considered an individual rather than a group bias. This project, therefore, assumed that a group, in this case the state, suffered from an individual anchoring bias on a given issue. While this disjoint is real and worth mentioning, many others have applied individual psychological biases to state behavior (Jervis 1976; Betts 1978; Lebow 1981; Levy 1984; Snyder 1984; Reiter 1996). This carries over into surprise, whereby individuals within the state may experience or react to surprise differently. The use of surprise for this study is the final qualification discussed in this section.

Application to Surprise

The project addressed foreign policy decision-making, but the formal models specifically addressed surprise, or more accurately crisis situations where one actor could attempt surprise against another. Surprise is an anomalous concept in the context of international politics and could refer to emotional reactions, but here it refers to an outcome. Surprise outcomes occur when targets do not take proper defensive or preemptive action to avoid or thwart the surprise. Like psychological bias, surprise can be difficult to study, but for different reasons. One can often observe a successful surprise attack easier than they can observe a psychological bias, but unsuccessful surprises, or canceled surprises, might be more difficult or impossible to identify.

The basic imperfect and incomplete models of surprise (prior to incorporating anchoring bias) developed for this project are applicable to any situation where one party has the opportunity to attempt surprise, and the opposing party can take action to thwart the surprise. That means the model deals with surprise as a general concept, and the models can be applied to instances of surprise other than attacks. One possible extension in the field of foreign policy could be economic action. Despite the substantive focus on surprise attack, the models are sufficiently general to apply broadly to foreign policy where there are two actors with competing preferences, they move sequentially, one cannot clearly observe the past actions of the other, and there is only play of the game. The approach to incorporating and assessing entrenched beliefs could also be transported to other models in international politics such as the basic deterrence model (Morrow 1994a) or a crisis bargaining model (Fearon 1995). These are possible areas for future extension.

Implications

The final section of this chapter briefly summarizes some of the main policy implications from this research. Unlike many works on intelligence or surprise, this project does not offer specific organizational remedies. The basic surprise model actually shows that external conditions make surprise possible even when an intelligence organization works perfectly. The policy implications derive directly from the theoretical and empirical analysis and address issues that decision-makers should keep in mind when creating policy.

Retrospective Analyses of Policy Failures

There are many retrospective studies of intelligence failure from both academic scholars and government commissions.[36] This research suggests that scholars and policymakers should approach these retrospective studies with a degree of caution. Many of these projects identify an intelligence failure and reverse engineer causes using detailed analysis of information, events and decisions. The problem with these studies is that they often search for extenuating circumstances rather than stressing the basic logic of surprise. Under certain strategic conditions with information constraints, surprise is a perfectly rational and plausible outcome without major organizational, psychological or legal constraints stemming from domestic political processes and organization. It is not

[36] There are many to note here. The Reference section includes many of these examples. A few include *Analysis of NFAC's Performance On Iran's Domestic Crisis, Mid- 1977-7 November 1978* (1978), *The Tet Offensive: Intelligence Failure in War* (Wirtz 1994), *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States* (2004) and *Report to the President: Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction* (2005).

surprising that many of these studies focus on organizational, psychological and legal aspects of intelligence and policymaking since this is the area that policymakers have the most control over. These types of recommendations can serve as prescriptions for policymakers. The irony is that basic competitive forces or environmental constraints are the primary necessary and sufficient conditions for surprise, and approaching the problem of surprise as one that can be fixed by domestic prescriptions is unlikely to meet such high expectations.[37]

The second reason to treat many retrospective studies of intelligence failure with some skepticism is the reliance on psychological explanations. In many of these studies, psychological variables emerge as causal factors either explicitly (Betts 1978; Wirtz 1994; Heuer 1999) or implicitly (9/11 Commission 2004). Studies that focus on psychology explicitly, often assume that relevant decision-makers suffered from some bias or some set of biases because bias is a pervasive cognitive feature. In other cases, one might reason that a certain decision must have been motivated by bias since it appeared irrational otherwise. This is the type of thinking that this study has tried to move away from, and in so doing, has shown how some of our long-held assumptions about psychological biases in foreign policy decision-making may be wrong.

Bias as a Stabilizing Force

Both the deductive theoretical proof and the quantitative analysis suggest that psychological bias, specifically irrationally entrenched beliefs, can act as a stabilizing force. While irrational under-suspicion had no impact on the deductive equilibria, the

[37] There are instances where imperfect domestic systems can be beneficial, such as Schelling's randomization or domestic constraints in international negotiations.

mere possibility of irrational over-suspicion did alter the initiator's optimal strategy. Initiators found that they were better off adopting strategies with a lower likelihood of attempted surprise then they would try if the target were perfectly rational. The target's possibility of irrational over-suspicion reduced the likelihood that surprise attempts would be successful since the biased target would be more likely to take defensive action than a Bayesian target. Initiators respond to the threat of over-suspicion by adopting a signaling strategy that is more conservative or less provocative.

Since the mere possibility of irrational over-suspicion is sufficient to constrain an initiator's aggressive behavior,[38] the deductive proof suggests that the same psychological biases that are often used as reasons for policy failure may inadvertently contribute to policy success by acting as a stabilizing mechanism. The quantitative analysis across a large dataset of international crises revealed that initiators were less likely to use violent crisis triggers when the target was a protracted rival, controlling for violence in previous crises. Controlling for past experience, the animosity or over-suspicion associated with protracted rivalry did have stabilizing impact by reducing the likelihood that crises were triggered violently. It is worth noting that previous violent experience did increase the likelihood that crises would be triggered violently, which may also provide some support for spiral conflict models where one actor's provocative action encourages another. This suggests that previous violence in itself correlated with future violence, but beliefs associated with protracted rivalry may act as a stabilizing force.

According to both the theoretical and quantitative analysis, then, psychological biases are not necessarily cognitive deficiencies that should be eliminated whenever

[38] The degree to which it is constrained is dependent on the exogenous factors like preferences and prior beliefs.

possible. In some instances, it constrains an opponent's behavior and helps policymakers justify taking necessary defensive action in ambiguous circumstances.

Taking Others' Biases Into Account

This is certainly not the first piece of scholarship in international politics to warn about the negative aspects or risks associated with aggressive strategies that carry the possibility of provoking unwanted responses. The security dilemma (Herz 1950; Jervis 1978) argues that armament policies intended to make one state more secure might inadvertently provoke a rival into arming themselves since they feel more vulnerable as a result of the initial armament policies. Similar research has also compared and contrasted the deterrent model of international action, whereby actors must take credibly signal resolve to dissuade the opposition from undesirable action, against the spiral model, in which signals or actions necessary for successful deterrence actually create a spiral of provocative action (Jervis 1976). A similar theme is present here. The difference between this project and prior work warning about the dangers of provoking a rival is that the primary causal variable in the prior work stressed the underlying nature of the game or interaction. A typical example of a spiral is an arms race whereby both actors might benefit from cooperation, but neither can afford to fall behind. Where actors' behavior and the structure of the game drive the traditional spiral, this project reinforces the concern that actors should also be concerned about their opponent's biases. The mere possibility of over-suspicion is enough to motivate unwanted action that could create a situation analogous to the spiral model.

One of the deductive conclusions is that an actor may be able to use their own psychological biases to their benefit by dissuading certain behavior. It is important, however, not to simply treat psychological biases as shield to hide behind. This research shows that it is equally important to take other's biases into account as well. The crisis initiator in the deductive models did not adopt a more conservative challenge strategy out the goodness of their theoretical heart, they did it because it offered the greatest utility given the possibility that their target had irrationally entrenched beliefs. This means that less provocative strategies, given the possibility of a biased opponent, are strategically optimal. Policymakers, particularly in foreign policy, would be wise to remember that others have biases and less provocative strategies may be strategically optimal at times, despite the desire to signal resolve or attempt a *fait accompli*.

The Pitfall of Perfect Rationality

The final recommendation is a warning: beware the pitfall of perfect rationality. The pitfall here does not refer to the benefits of psychological biases despite the fact that biases may be strategically optimal. The pitfall refers to the desire, even obsession, to eliminate psychological biases from the decision-making process. Research on psychology and decision-making in international politics sometimes concludes either that bias resulting in poor outcomes is inevitable since it is impossible to eliminate (Betts 1978), or stresses the need to minimize or eliminate biases so as to lessen their impact (George 1980; Heuer 1999). Neither of these approaches is necessarily correct. In the first case, psychological biases do not always correlate with poor decisions or outcomes, and may be beneficial in numerous circumstances. In the second case, the attempt to

minimize biases may send the inaccurate signal that one is perfectly rational. The attempt to appear perfectly rational eliminates the possible strategic benefits of bias since an actor might believe their opponent is perfectly rational. The only way to derive strategic benefits from bias is expose one's possible biases to an opponent.

Chasing after perfect rationality creates an unrealistic goal for analysts and policymakers. It may also be strategically sub-optimal. Instead of running away from biases, policymakers would be better off understanding how they impact their decisions, and this should be an area that scholars continue to focus on.

References

Allison, Graham and Philip Zelikow. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. 2nd ed. 1999. New York, NY: Longman.

Altfeld, Michael F. and Bruce Bueno de Mesquita. 1979. "Choosing Sides in War," *International Studies Quarterly* 23(1): 87-112.

Axelrod, Robert. 1979. "Coping With Deception." *Applied Game Theory*. Brams et al. eds. Wuerzberg, Germany: Physica-Verlag.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York, NY: Basic Books.

Babcock, Linda and George Lowenstein. 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Bias." *Journal of Economic Perspectives* 11: 109-126

Barberisa, Nicholas, Andrei Shleiferb, and Robert Vishny. 1998. A Model of Investor Sentiment." *Journal of Financial Economics* 49(3): 307-343.

Bar-Joseph, Uri. 2003. "Intelligence Failure and the Need for Cognitive Closure: The Case of Yom Kippur." *Paradoxes of Strategic Intelligence: Essays in Honor of Michael I. Handel*. Betts and Mahnken eds. London, UK: Frank Cass: 166-189.

Baron, Jonathan. 2000. *Thinking and Deciding,* 3rd ed. New York, NY: Cambridge University Press.

Bar-Tal, Daniel. 2000. "From Intractable Conflict Through Conflict Resolution to Reconciliation: Psychological Analysis." *Political Psychology* 21(2): 351-365.

Bar-Zohar, Michael. 1970. *Embassies in Crisis: Diplomats and Demogogues Behind the Six Day War*. Englewood, NJ: Prentice-Hall.

Beschloss, Michael R. 1991. *The Crisis Years: Kennedy and Khrushchev 1960-1963.* New York, NY: Edward Burlingame Books.

Bennett, Andrew and Alexander L. George. 1997. "Process Tracing in Case Study Research." MacArthur Foundation Workshop on Case Study Methods October 17-19.

Bennett, Andrew and Colin Elman. 2006. "Qualitative Research: Recent Developments in Case Study Analysis. *Annual Review of Political Science* 9: 455-476.

Betts, Richard K. 1978. "Analysis, War, and Decision: Why Intelligence Failures Are Inevitable." *World Politics* 31(1): 61-89.

Betts, Richard K. 1982. *Surprise Attack: Lessons for Defence Planning*. Washington DC: Brookings Institution Press.

Blau, George. 1955. *German Campaign in Russia*. Washington, DC: Center of Military History.

Blight, James G., Bruce J. Allyn and David A. Welch. 2002. *Cuba on the Brink: Castro, the Missile Crisis, and the Soviet Collapse*. Lanham, MA: Rowman & Littlefield.

Bowen, Jeremy. 2003: *Six Day: How the 1967 War Shaped the Middle East*. London, UK: Simon and Schuster.

Brady, H. E. 2003. *Models of Causal Inference: Going Beyond the Neymen-Rubin-Holland Theory*. Presented at the Annual Meeting of the Midwest Political Science Association.

Brecher, Michael. 1980. *Decisions in Crisis: Israel, 1967 and 1973*. Los Angeles, CA, University of California Press.

Bueno de Mesquita, Bruce and William H. Riker. 1982. "An Assessment of the Merits of Selective Nuclear Proliferation." *Journal of Conflict Resolution* 26(2): 283-306.

Bueno de Mesquita, Bruce and James D. Morrow. 1999. "Sorting Through the Wealth of Notions." *International Security* 24(2): 56.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, James D. Morrow. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93(4): 791-807.

Butler, Christopher K. 2007. "Prospect Theory and Coercive Bargaining." *Journal of Conflict Resolution* 51(2): 227-250.

Carell, Paul. *Hitler's War on Russia*. London, UK: George G. Harrap & Co. Ltd.

Camerer, Colin and George Lowenstein. 2004. "Behavioral Economic: Past, Present and Future." *Advances in Behavioral Economics*. Camerer et al. eds. Princeton, NJ: Princeton University Press.

Campbell, Donald T. and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Dallas, TX: Houghton Mifflin.

"CIA Special Intelligence Estimate: The Military Buildup un Cuba, September 19, 1962." in *The Cuban Missiles Crisis, 1962: A National Security Archive Documents Reader*. Lawrence Chang and Peter Kornbluh eds. New York, NY: The New Press.

Clark, Alan. 1965. *Barbarossa: The Russian-German Conflict, 1941-45*. New York, NY: Perennial.

Cohen, Raymond. 1988. "Intercultural communication between Israel and Egypt: Deterrence Failure Before the Six Day War." *Review of International Studies* 14(1): 1-16.

Compte, Olivier and Andrew Postlewaite. 2004. "Confidence-Enhanced Performance." *American Economic Review* 94(5): 1536-1557.

Daniel, Kent, David Hirshleifer, Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under-and Overreactions." *The Journal of Finance* 53(6): 1839-1886.

Daugherty, William J. 2001. "Behind the Intelligence Failure in Iran." *International Journal of Intelligence and Counterintelligence* 14(4): 449-484.

Dawn, C.E. 1968. "The Egyptian Remilitarization of the Sinai: May 1967." *Journal of Contemporary History* 3(3): 201-224.

Dayan, Moshe. 1976. *Moshe Dayan: The Story of My Life*. New York, NY: William Morow and Company.

De Bondt, Warner F.M. and Richard Thaler. 1985. "Does the Stock Market Overreact?" *Journal of Finance* 40(3): 793-805.

De Bondt, Warner F.M. and Richard Thaler. 1987. "Further Evidence on Investor Overreaction and Stock Market Seasonality." *Journal of Finance* 42(3): 557-581.

Downs, George W. et al. 1996. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50(3): 379-406.

Draper, Theodore. 1968. *Israel and World Politics*. New York, NY: Viking.

Eban, Abba. 1992. Personal Witness: Israel Through My Eyes. New York, NY: G.P. Putnam's Sons.

Ellsberg, Daniel. 1962. *Risk, Ambiguity and Decision*. New York, NY: Routledge.

Estes, William. 1964. "Probability Learning." In A. W. Melton (ed.), *Categories of Human Learning*. New York, NY: Academic Press

Fearon, James. 1994. “Domestic Political Audiences and the Escalation of International Disputes,” *American Political Science Review* 88(3): 577-592.

Fearon, James. 1995. “Rationalist Explanations for War,” *International Organization* 49(3): 379-414.

Fearon, James D. 1998. “Bargaining, Enforcement and International Cooperation,” *International Organization* 52(2): 269-306.

Feaver, Peter D. 1993. “Proliferation Optimism and Theories of Nuclear Operations.” *Security Studies*. 2(3&4): 159-191.

Feaver, Peter D. and Emerson M. S. Niou. 1996. “Managing Nuclear Proliferation: Condemn, Strike, or Assist?” *International Studies Quarterly* 40(2): 209-233.

Fischhoff, Baruch and Ruth Beyth. 1975. “‘I knew it would happen’: Remembered probabilities of once-future things.” *Organizational Behavior and Human Performance* 13(1): 1-16.

Freedman, Lawrence, and Ephraim Karsh. 1993. *The Gulf Conflict*. London, UK: Faber and Faber.

Freedman, Lawrence. 2000. *Kennedy’s Wars: Berlin, Cuba, Laos and Vietnam.* Oxford, UK: Oxford University Press.

Fuller, Sally Riggs, and Ramon J. Aldag. “Challenging the Mindguards: Moving Small Group Analysis beyond Groupthink.” *Beyond Groupthink: Political Group Dynamics and Foreign Policy-making,* Paul t’Hart, Eric K. Stern and Bengt Sundelius eds. Ann Arbor, MI: University of Michigan Press.

Fursenko, Aleksandr and Timothy Naftali. 1997. *“One Hell of a Gamble”: The Secret History of the Cuban Missile Crisis*. New York, NY: W. W. Norton and Company.

Fursenko, Aleksandr and Timothy Naftali. 2006. *Khrushchev's Cold War: The Inside Story of an American Adversary*. New York, NY: W.W. Norton.

Gaddis, John Lewis. 2005. *The Cold War: A New History*. New York, NY: Penguin Press.

Gartner, Scott Sigmund. 1997. *Strategic Assessment in War*. New Haven, CT: Yale University Press.

Gelpi, Christopher. 2002. *The Power of Legitimacy: Assessing the Role of Norms in Crisis Bargaining*. Princeton, NJ: Princeton University Press.

George, Alexander L. 1980. *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice*. Boulder, CO: Westview Press.

Gerber, Alan and Donald Green. 1999. "Misperceptions About Perceptual Bias." *The Annul Review of Political Science* 2: 189-210.

Gerring, John. 2001. *Social Science Methodology: A Critical Framework*. Cambridge, UK: Cambridge University Press.

Gibbons, Robert. 1992. *Game Theory for Applied Economists*. Princeton, NJ: Princeton University Press.

Glantz, David M. 1990. *The Role of Intelligence in Soviet Military Strategy in World War II*. Novato, CA: Providio Press.

Glantz, David M. and Jonathon House. 1995. *When Titan Clashed: How the Red Army Stopped Hitler*. Lawrence: University of Kansas Press.

Glaser, Charles. 1997. "The Security Dilemma Revisited," *World Politics* 50(1): 171-201.

Goertz, Gary and Paul F. Diehl. 1993. "Enduring Rivalries: Theoretical Constructs and Empirical Patterns." *International Studies Quarterly* 37(3): 147-171.

Goldgeier, James M. and Philip E. Tetlock. 2001. "Psychology in International Relations Theory." *The Annual Review of Political Science*, 4: 67-92.

Gooch, John and Amos Perlmutter. 1982. "Introduction." *Military Surprise and Strategic Deception*. John Gooch and Amos Perlmutter, eds. London, UK: Frank Cass.

Gorodetsky, Gabriel. 1999. *Grand Delusion: Stalin and the German Invasion of Russia*. New Haven, CT: Yale University Press.

Grinblatt, Mark, Rinald W. Masulis, and Sheridan Titman. 1984. "The Valuation Effects of Stock Splits and Stock Dividends." Journal of Financial Economics 13(4).

Groth, John C., Wilbur G. Lewellen, Gary G. Schlarbaum, and Ronald C. Lease. 1979. "An Analysis of Brokerage House Securities Recommendations." *Financial Analysts Journal* 35(1): 32-40.

Haas, Mark L. 2001. "Prospect Theory and the Cuban Missile Crisis." *International Studies Quarterly* 45: 241-270.

Halperin, Morton H. 1972. "The Decision to Deploy the ABM: Bureaucratic and Domestic Politics in the Johnson Administration." *World Politics* 25(1): 62-95.

Hammel, Eric. 1992. *Six Days In June: How Israel Won the 1967 Arab-Israeli War*. New York, NY: Charles Scribner's Sons.

Handel, Michael I. 1989. "Intelligence and the Problem of Strategic Surprise." *War Strategy and Intelligence*. Michael I. Handel ed. London, UK: Frank Cass and Co. Ltd.

Hassner, Ron E. 2007. "The Path to Intractability: Time and the Entrenchment to Territorial Disputes." *International Security* 31(3): 107-138.

Hastings, Max. 1987. *The Korean War*. New York, NY: Simon and Schuster.

Heikal, Mohammed Hssanein. 1973. *The Cairo Documents*. New York, NY: Doubleday.

Herek, Gregory M., Irving L. Janis, and Paul Huth. 1987. "Decision Making During International Crises." *Journal of Conflict Resolution* 31(2): 203-226.

Herman, Michael. 1996. *Intelligence Power in Peace and War*. Cambridge UK: Cambridge University Press.

Hermann, Margaret G. and Charles F. Hermann. 1989. "Who Makes Foreign Policy Decisions and How: An Empirical Inquiry." *International Studies Quarterly* 33(4): 361-387.

Hermann, Charles F. 1990. "Changing Course: When Governments Choose to Redirect Foreign Policy." *International Studies Quarterly* 34(3): 3-21.

Herz, John H. 1950. "Idealist Internationalism and the Security Dilemma." *World Politics* 2(2): 157-180.

Heuer Jr. Richards J. 1981. "Strategic Deception and Counterdeception: A Cognitive Process Approach." *International Studies Quarterly* 25(2): 294-327.

Heuer Jr. Richards J. 1999. *Psychology of Intelligence Analysis*. Washington DC: Center for the Study of Intelligence.

Higgins, Trumbull. 1966. *Hitler and Russia*. New York, NY: Macmillan Company.

Hopple, Gerald W. 1984. "Intelligence and Warning: Implications and Lessons of the Falkland Islands War." *World Politics*, 36(3): 339-361.

Huth, Paul and Bruce Russett. 1984. "What Makes Deterrence Work? Cases from 1900 to 1980." *World Politics* 36(4): 496-526.

Huth, Paul. 1999. "Deterrence and International Conflict," *Annual Review of Political Science* 2: 25-48.

James, Laura M. 2006. *Nasser at War: Arab Images of the Enemy*. Houndmills, UK: Palgrave.

Janis, Irving L. 1972. *Victims of Groupthink: A Psychological Study of Foreign Policy Decisions and Fiascoes*. Boston, MA: Houghton, Mifflin.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48(1): 65-91.

Jervis, Robert. 1970. *The Logic of Images in International Relations*. New York, NY: Columbia University Press.

Jervis, Robert. 1976. Perception and Misperception in International Politics. Princeton, NJ: Princeton University Press.

Jervis, Robert. 1979. *Analysis of NFAC's Performance On Iran's Domestic Crisis, Mid-1977-7 November 1978*.

Kahneman, Daniel and Amos Tversky. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185: 1124-1131.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47(2): 263-292.

Kahneman, Daniel, Paul Slovic, Amos Tversky. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.

Kam, Ephraim. 1988. *Surprise Attack: The Victim's Perspective*. Cambridge, MA: Harvard University Press.

Kennedy, Peter. 2003. *A Guide to Econometrics*, 5th ed. Cambridge, MA: MIT Press.

Keohane, Robert. 1984. *After Hegemony*. Princeton, NJ: Princeton University Press.

Khouri, Fred J. 1968. *The Arab-Israeli Dilemma*. Syracuse, NY: Syracuse University Press.

Khrushchev, Nikita. 1970. *Khrushchev Remembers*. Translated by Strobe Talbott. Boston, MA: Little, Brown and Company.

Khrushchev, Nikita. 1974. *Khrushchev Remembers: The Last Testament*. Translated by Strobe Talbott. Boston, MA: Little, Brown and Company.

Khrushchev, Sergei N. 2000. *Nikita Khrushchev: And the Creation of a Superpower*. University Park, PA: Pennsylvania State University Press.

Kilgour, D. Marc and Frank C. Zagare. 1991. "Credibility, Uncertainty, and Deterrence." *American Journal of Political Science* 35(2): 305-334.

Kilgour, D. Marc and Steven J. Brams. 1992. "Putting the Other Side 'On Notice' Can Induce Compliance in Arms Control." *Journal of Conflict Resolution* 36(3): 395-414.

Kimche, David and Dan Bawly. 1968. *The Sandstorm*: *The Arab-Israeli War of June 1967: prelude and Aftermath*. New York, NY: Stein and Day.

Kindleberger, Charles, P. 1970. *Power and Money: The Economics of International Politics and the Politics of International Economics*. New York, NY: Basic Books.

King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Kissinger, Henry. 1994. *Diplomacy*. New York, NY: Simon and Schuster.

Krasner, Stephen D. 1976. "State Power and the Structure of International Trade," *World Politics* 28(3): 317-347.

Kydd, Andrew. 1997. "Game Theory and the Spiral Model." *World Politics* 49(3): 371-400.

Lacouture, Jean. 1973. *Nasser: A Biography*. New York, NY: Knopf.

Langlois, Jean-Pierre P. 1991. "Rational Deterrence and Crisis Stability." *American Journal of Political Science* 35(4): 801-832.

Leaming, Barbara. 2006. *Jack Kennedy: The Education of a Statesman*. New York, NY: W.W. Norton.

Lebow, Richard Ned. 1981. *Between Peace and War*. Baltimore, MD: Johns Hopkins University Press.

Legro, Jeffrey W. 1995. *Cooperation Under Fire: Anglo-German Restraint During World War II*. Ithaca and London: Cornell University Press.

Leng, Russell J. 1983. "When Will They Ever Learn?" *Journal of Conflict Resolution* 27(3): 379-419.

Levite, Ariel J. 1987. *Intelligence and Strategic Surprise*. New York, NY: Columbia University Press.

Levitt, Barbara and James G. March. 1988. "Organizational Learning." *Annual Review of Sociology* 14: 320.

Levy, Jack S. 1984. "The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis." *International Studies Quarterly* 28(2): 219-238.

Levy, Jack S. 1994. "Learning and Foreign Policy: Sweeping a Conceptual Minefield." *International Organization* 48(2): 279-312.

Lowenthal, Mark M. 2003. *Intelligence: From Secrets to Policy*. 2nd ed. Washington DC: CQ Press.

Luttwack, Edward N. 1987. *Strategy: The Logic of War and Peace.* Cambridge, UK: Belknap Press.

Maoz, Zeev. 2006. "Intelligence Failures: An Analytic Framework." Paper presented at the Annual Meeting of the American Political Science Association, September 2006, Philadelphia, PA.

Martin, Lisa K. 1999. "The Contributions of Rational Choice: A Defense of Pluralism." *International Security* 24(2): 74.

May, Ernest R. and Philip D. Zelikow. 2001. *The Kennedy Tapes: Inside the White House During the Cuban Missile Crisis*. New York, NY: W. W. Norton and Company.

McDermott, Rose. 1998. *Risk-taking in International Politics: Prospect Theory in American*. Ann Arbor, MI: University of Michigan Press.

McDermott, Rose and Jacek Kugler. 2001. "Comparing Rational Choice and Prospect Theory Analyses: The U.S. Decision to Launch Operation Desert Storm, January 1991." *Journal of Strategic Studies* 24(3): 49-85.

McDermott, Rose. 2004. "Prospect Theory in Political Science: Gains and Losses From the First Decade." *Political Psychology* 25(2): 289-312.

McNamara, Robert. 2000, "Britain, Nasser and the Outbreak of the Six Day War." *Journal of Contemporary History* 35(4): 619-639.

Mertens, J. F. and S. Zamir. 1985. "Formulation of Bayesian Analysis for Games With Incomplete Information." *International Journal of Game Theory* 14: 1–29.

Milner, Helen. 1997. *Interests, Institutions, and Information*. Princeton, NJ: Princeton University Press.

Mor, Ben D. 1991. “Nasser’s Decision-Making in the 1967 Middle East Crisis: A Rational-Choice Explanation.” *Journal of Peace Research* 28(4): 359-375.

Mor, Ben D. 2004. “Strategic Beliefs and the Formation of Enduring International Rivalries: Israel's National Security Conception, 1948-56.” *International Relations* 18(3): 309-330.

Morgenthau, Hans J. 1948. *Politics Among Nations*, Brief Edition (1993). McGraw-Hill.

Morrow, James D. 1991. “Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances,” *American Journal of Political Science* 35(4): 904-33.

Morrow, James. 1994a. *Game Theory for Political Scientists*. Princeton, NJ: Princeton University Press.

Morrow, James D. 1994b. “Modeling the Forms of International Cooperation: Distribution Versus Information,” *International Organization* 48(3): 387-423.

Myerson, Roger B. 1991. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

Nalebuff, Barry. 1991. “Rational Deterrence in an Imperfect World.” *World Politics* 43(3): 313-335.

Neustadt, Richard E. and Ernest R. May. 1986. *Thinking in Time: The Uses of History for Decision Makers*. New York, NY: The Free Press.

Niou, Emerson M. S. and Peter C. Ordeshook. 1990. “Stability in Anarchic International Systems.” *The American Political Science Review* 84(4): 1207-1234.

Niou, Emerson M. S. and Peter C. Ordeshook. 1999. “Return of the Luddites.” *International Security* 24(2): 84.

Nisbett, Richard and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Human Judgment*. Englewood, NJ: Prentice-Hall.

Noble, Harold Joyce. 1975. *Embassies at War*. Seattle, WA: University of Washington Press.

Nutting, Anthony. 1972. *Nasser*. New York, NY: Dutton.

O’Brien, Michael. 2005. *John F. Kennedy: A Biography*. New York, NY: Thomas Dune Books.

Odean, Terrance. 1999. "Do Investors Trade Too Much?" *American Economic Review* 89: 1279-1298.

Oren, Michael B. 2002. *Sox Day of War: June 1967 and the Making of the Modern Middle East*. New York, NY: Ballantine Books.

Pierson, Paul and Theda Skocpol. 2002. "Historical Institutionalism in Contemporary Political Science." *Political Science: The State of the Discipline*. Washington, DC: American Political Science Association.

Paige, Glenn D. 1968. *The Korean Decision, June 24-30, 1950*. New York, NY: Free Press.

Parker, Charles F. and Eric K. Stern. 2005. "Bolt from the Blue or Avoidable Failure? Revisiting September 11 and the Origins of Strategic Surprise." *Foreign Policy Analysis* 1(3): 301-331.

Pillar, Paul. 2004. "Counterterrorism after Al Qaeda." *The Washington Quarterly* 27(3): 101-113.

Powell, Robert. 1991. "Absolute and Relative Gains in International Relations Theory." *American Political Science Review* 85(4): 1303-1320.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *American Political Science Review* 87(1): 115-132.

Rabin, Yitzhak. 1979. *The Rabin Memoirs*. Boston, MA: Little, Brown.

Rabin, Matthew and Joel Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *The Quarterly Journal of Economics* 114(1): 37-82.

Reiter, Dan. 1995. Exploding the Powder Keg Myth: Preemptive Wars Almost Never Happen." *International Security* 20(2): 5-34.

Reiter, Dan. 1996. *Crucible of Beliefs: Learning, Alliances and World Wars.* Ithaca, NY: Cornell University Press.

Report to the President: Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction. 2005. Available at http://www.wmd.gov/report/report.html.

Ripley, B. 1995. "Cognition, Culture, and Bureaucratic Politics." *Foreign Policy Analysis: Continuity and Change in Its Second Generation*, Laura Neack, Jeanne A. K. Hey and Patrick J. Haney eds. New York, NY: Prentice Hall,

Robarge, David S. 2007. “CIA Analysis of the 1967 Arab-Israeli War: Getting It Right.” *Studies in Intelligence*, Center for the Study of Intelligence 49(1).

Rostow, W. W. “Memorandum to the President, Assessing Soviet Military Aid to Cuba, September 3, 1992.” in *The Cuban Missiles Crisis, 1962: A National Security Archive Documents Reader*. Lawrence Chang and Peter Kornbluh eds. New York, NY: The New Press.

Ruggie, John Gerard. 1998. “What Makes the World Hang Together?” *International Organization* 52(4): 855-886.

Safran, Nadav. 1969. *From War to War: The Arab-Israeli Confrontation, 1948-1967*. Indianapolis, IN: Bobbs-Merrill.

Sakovics, Jozsef. 2001. “Games of Incomplete Information Without Common Knowledge Priors.” *Theory and Decision* 50: 347-366.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Schultz, Kenneth A. 1998. “Domestic Opposition and Signaling in International Crises.” *The American Political Science Review* 92 (3): 829-844.

Schultz, Kenneth. 2001. *Democracy and Coercive Diplomacy*. Cambridge, UK: Cambridge University Press.

Scott, Len and Steve Smith. 1994. “Lessons of October: Historians, Political Scientists, Policy-Makers and the Cuban Missile Crisis.” *International Studies Quarterly* 70(4): 659-684.

Scott, Len. 2007. *The Cuban Missile Crisis: And the Treat of Nuclear War*. London, UK: Continuum.

Segev, Tom. 2005. *1967: Israel, the War, and the Year That Transformed the Middle East.* New York, NY: Metropolitan Books.

Smith, Steve. 1985. “Policy Preferences and Bureaucratic Position: The Case of the American Hostage Rescue Mission.” *International Affairs* 61(1): 9-25.

Smith, Vernon L., Gerry L. Suchanek and Arlington W. Williams. 1988. “Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets.” *Econometrica* 56(5): 1119-1151.

Smith, Alastair and Allan C. Stam. 2004. “Bargaining and the Nature of War.” *Journal of Conflict Resolution* 48(6): 783-813.

Snidal, Duncan. 1985. "The Limits of Hegemonic Stability Theory." *International Organization* 39(4): 579-614.

Snyder, Jack. 1984. *The Ideology of the Offensive: Military Decisionmaking and the Disasters of 1914*. Ithaca, NY: Cornell University Press.

Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87(3): 355-374.

Stein, Janice Gross and Raymond Tanter. 1980. *Rational Decision-Making*. Columbus, OH: Ohio State University Press.

Stein, Janice Gross. 1988. "Building Politics into Psychology: The Misperception of Threat." *Political Psychology* 9: 245-271.

Stein, Janice Gross. 1991. "Deterrence and Reassurance." *Behavior, Society, Nuclear War*. Volume 2. Tetlock et al. eds. New York, NY: Oxford University Press.

Steinberg, Paul F. 2007. "Causal Assessment in Small-N Policy Studies." *The Policy Studies Journal* 35(2): 181-204.

Stephens, Robert. 1971. *Nasser: Apolitical Biography*. New York, NY: Simon and Schuster.

Stern, Eric and Bertjan Verbeek. "Whither the Study of Governmental Politics in Foreign Policymaking? A Symposium." *International Studies Review* 42(2): 205-255.

Stern, Sheldon M. 2005. *The Week the World Stood Still: Inside the Secret Cuban Missile Crisis*. Stanford, CA: Stanford University Press.

Stueck, William. 2001. *Rethinking the Korean War: A New Diplomatic and Strategic History*. Princeton, NJ: Princeton University Press.

Suedfeld, Peter and Philip Tetlock. 1977. "Integrative Complexity of Communications in International Crises." *Journal of Conflict Resolution* 21(1): 169-184.

Svenson, Ola. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica*, 47: 143-48.

Taliaferro, Jeffrey W. 2004. "Power Politics and the Balance of Risk: Hypotheses on Great Power Intervention in the Periphery." *Political Psychology* 25(2): 177-211.

Taubman, William. 2003. *Khrushchev: The Man and His Era*. New York, NY: W.W. Norton.

Tetlock, Philip E. 1998a. "Social Psychology and World Politics." *Handbook of Social Psychology*. Gilbert et al. eds. New York, NY: McGraw-Hill.

Tetlock, Philip E. 1998b. "Close-Call Counterfactuals and Belief System Defenses: I Was Not Almost Wrong But I Was Almost Right." *Journal of Personality and Social Psychology*, 75: 639-652.

Tetlock, Philip E. 1999. "Theory-Driven Reasoning About Possible Pasts and Probable Futures: Are We Prisoners of Our Preconceptions?" *American Journal of Political Science*, 43: 335-366.

The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States. 2004. New York, NY: W.W. Norton & Co.

Thies, Cameron G. 2001. "A Social Psychology Approach to Enduring Rivalries." *Political Psychology* 22(4): 693725.

Vandenbroucke, Lucien S. 1984. "Anatomy of a Failure: The Decision to Land at the Bay of Pigs." *Political Science Quarterly* 99(3): 471-491.

Vasquez, John A. 1993. *The War Puzzle*. Cambridge, UK: Cambridge University Press.

Vertzberger, Yaacov. 1984. "Bureaucratic-Organizational Politics and Information-Processing in a Developing State." *International Studies Quarterly* 28(1): 69-85

Vertzberger, Yaacov. 1986. Foreign Policy Decision Makers as Practical-Intuitive Historians: Applied History and its Shortcomings." *International Studies Quarterly* 30(3): 223-247.

Vertzberger, Yaacov I. 1997. "Collective Risk-taking: The Decision-making Group." *Beyond Groupthink: Political Group Dynamics and Foreign Policy-making,* Paul t'Hart, Eric K. Stern and Bengt Sundelius eds. Ann Arbor, MI: University of Michigan Press.

Wagner, Harrison. 1991. "Nuclear Deterrence, Counterforce Strategies, and the Incentive to Strike First." *American Political Science Review,* 85(3): 727-749.

Walt, Stephen M. 1999. "Rigor or Rigor Mortis? Rational Choice and Security Studies." *International Security* 23(4): 5-48.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.

Weathersby, Kathryn. 1999. "Should We Fear This?: Stalin and the Korean War." Paper presented at the Stalin and the Cold War conference, Yale University.

Wendt, Alexander. 1992. "Anarchy is What States Make of It," *International Organization* 46(2): 391-425.

Weizman, Ezer. 1976. *On Eagles' Wings: The Personal Story of The Leading Commander of the Israeli Air Force*. New York, NY: Macmillan.

Welch, David A. 1998. "A Positive Science of Bureaucratic Politics?" *Mershon International Studies Review* 42(2): 210-216.

Whaley, Barton. 1969. *Stratagem: Deception and Surprise in War*. Unpublished manuscript. Cambridge, MA.

Whaley, Barton. 1982a. "Toward a General Theory of Deception." *Military Deception and Strategic Surprise*. John Gooch and Amos Perlmutter eds. London, UK: Frank Cass and Co. Ltd.

Whaley, Barton. 1982b. "Covert Rearmament in German 1919-1939: Deception and Misperception." *Military Deception and Strategic Surprise*. John Gooch and Amos Perlmutter eds. London, UK: Frank Cass and Co. Ltd.

Wohlstetter, Roberta. 1962. *Pearl Harbor: Warning and Decision*. Stanford, CA: Stanford University Press.

Wirtz. James J. 1991. "Intelligence to Please? The Order of Battle Controversy during the Vietnam War." *Political Science Quarterly* 106(2): 239-263.

Wirtz, James J. 1994. *The Tet Offensive: Intelligence Failure in War*. Ithaca, NY: Cornell University Press.

Wittman, Donald. 1989. "Arms Control Verification and Other Games Involving Imperfect Detection." *American Political Science Review* 83(3): 923-945.

Wonnacottt, Thomas H. and Wonnacott, Ronald J. 1990. *Introductory Statistics For Business and Economics*, 4th ed. New York, NY: John Wiley and Sons.

Yost, C. 1968. "The Arab-Israeli War: How It Began." *Foreign Affairs* 26(2): 304-320.

Zagare, Frank C. 1999. "All Mortis, No Rigor." *International Security* 24(2): 107.

Zagare, Frank C. and D. Marc Kilgour. 2000. *Perfect Deterrence.* Cambridge, UK: Cambridge University Press.