Part I

Coercion and Credibility

1

Introduction

Preparation for war does not make war inevitable. On the contrary, prudent preparation for war, accompanied by a wise policy, provides a guarantee that war will not break out except for the gravest of reasons.

Count Sergei I. Witte

Military power is what gets one's voice heard in world affairs. Creating and maintaining armed forces is among the costliest undertakings for a nation short of their employment in hostilities. Even a casual glance at history reveals that whatever their defensive role is, armed forces are often used to menace others. More often than not, they are used indirectly, as an implicit or explicit presence in the background of negotiations, rather than directly in fighting.[1] States frequently find themselves on the opposite sides of disputes, and in their attempts to wrangle concessions out of each other they sometimes resort to military threats. The threat to use force can be verbal without any overt preparation to do so, or physical with all the measures – putting forces on alert, recalling reservists, mobilizing, dispatching the navy, deploying troops – required for its actual use. These physical measures, which I collectively refer to as *military moves*, do not have to be accompanied by an explicit warning. They are so menacing that the threat of hostile intent is implicit in their use. Sometimes these moves are nothing but necessary steps on the road to war. But more often, they are intended as a warning that war may come unless the adversary accedes to one's demands. War, with its enormous costs, pain, and risks, is not something to be contemplated lightly. But there are things worse than war and common sense dictates what history reveals: even state leaders who are averse to war can deliberately risk it to convince others to bend to their wishes.

It is the function of military moves as instruments to induce desired behavior in others, rather than their proper application in the deadly arts of destruction, that interests me. This is a book on military coercion. It is a

[1] Goldhamer (1979, 9); Karsten et al. (1984, 3–5); Naroll et al. (1974, 1–2); Schelling (1966); Blechman and Kaplan (1978); Young (1968).

study of how military threats can be employed in the pursuit of political goals. For a military threat to succeed as a coercive device, it has to accomplish two objectives: (a) it has to persuade the opponent that one is sufficiently likely to resort to violence if one's demands are not met, and (b) it has to render fighting sufficiently unpleasant for the opponent relative to the concessions demanded. What makes military threats effective? Why might they fail even if they are believable? Why would an actor forego the possibilities of militarized diplomacy and opt for war instead? How are military threats different from other instruments of coercion? These are all questions I address in this book. Although my interest is primarily theoretical, I will draw upon numerous historical cases to motivate the research and illustrate the logic of its findings.

The fundamental result is that military threats can be very effective tools of coercion. They can establish intent to wage war and can communicate that fact to the opponent in a way that he will believe it. Military threats can even reduce the likelihood that the confrontation will end in war, relative to other coercive instruments. Unfortunately, these threats also tend to be expensive, especially if their intent is to coerce the opponent rather than wage war. Whereas this may discourage their use and thereby reduce the chances of a militarized dispute, it may also convince leaders that it is easier to settle the matter by force instead of trying to coerce the opponent with threats. This makes war more likely and underscores the need to distinguish between military moves that are a prelude to war and those that are designed to influence the opponent's behavior. These, as Count Witte observed, are not quite the same even though they may take similar outward appearances.[2]

The findings have implications for international relations theory and policy. On the theoretical side, the results contradict a long tradition of arguing that nations with more powerful militaries tend to get their way more often than others but at the cost of having to risk war more often too. This may be so for non-military instruments but not for military threats. Through the judicious use of military threats, powerful states can secure better peaceful outcomes and lower the risk of war. Their task can be made more difficult if they misperceive the magnitude of the stakes for their opponent. Their overconfidence may prove to be their undoing if they fail to muster the resources necessary to coerce a determined adversary. However, even if they are pessimistic, their actions may make war more likely because they mistakenly believe that it would take too much effort to coerce the opponent

[2] Ironically, Witte made this remark about the preparations right before the outbreak of the Russo-Japanese War of 1904–05 (Harcave, 1990, 308–09).

and opt for war instead. In fact, the finding that the overall danger of war is mediated through the distribution of interests can help explain why attempts to link it directly to the distribution of power have generally failed. The likelihood of war depends on the extent to which one is prepared to use military threats to deter challenges to peace and compel concessions without fighting. The price of peace may be military establishments that are both costly and unused. These armed forces are not useless, for their employment is indirect but nevertheless crucial.

I am more reluctant to draw conclusions with policy implications because no one is more acutely aware of the shortcomings of my theories than I am. However, even I cannot resist a couple of observations. Despite the attractiveness of the military instrument as a tool for coercion, one cannot have militarized coercion on the cheap. Gunboat diplomacy is unlikely to work unless it represents firepower that can make a difference in an actual engagement. In other words, military threats cannot be token in character if they are to succeed. They are not a cheap way for the powerful to throw their weight around. In fact, wealthier and more powerful nations may have to engage in relatively more aggressive behavior in order to make their threats stick. They may have to mobilize overkill capability compared to the issues at stake. Shooting flies with an elephant gun may well be the prudent thing for them to do.

The argument in this book depends on a series of theoretical models which all share the same basic assumptions. In this, they all stand or fall together, so it may be worthwhile to provide some justification for the choices I have made. I assume that a conflict of interest exists between two unitary rational actors who confront each other once to resolve it. A number of important assumptions are already buried in this simple statement.

I assume that the two actors are unitary and rational; that is, they behave as individuals with well-defined preferences. By "well-defined" preferences I mean that the actors can rank-order all the various possible outcomes of their interaction in a logically coherent way. More importantly, they can rank-order risky alternatives. For instance, suppose an actor is confronted with an ultimatum from his opponent and, for simplicity, suppose he has three options at his disposal: agree to the terms, launch a preemptive attack, or let the ultimatum expire to see if his opponent will attack. To decide on the best course of action, the actor must evaluate the likely consequences of the various options at his disposal. Capitulation to the opponent's demands avoids war but (presumably) imposes political and economic costs by forcing the actor to agree to unpalatable conditions. Launching a preemptive strike means going to war, with all the attendant risks and costs. There is no guarantee of victory but there is a chance to

avoid the bad outcome. The third option is to let the ultimatum deadline lapse in the hope that the opponent will not attack. Unlike the outright capitulation, there may be a chance to avoid the bad outcome but at the risk of a war. Unlike launching a preemptive strike, there is a chance to avoid war but at the risk of foregoing whatever advantages such an attack would confer.

Each of these options has its own costs and benefits and each involves some trade-offs. We say that preferences are rational when they are logically consistent. For instance, it cannot be the case that the actor expresses a preference for adopting a wait-and-see stance to preemptive attack and preemptive attack to outright capitulation and then also be the case that he expresses a preference for outright capitulation over adopting a wait-and-see stance.[3] Throughout this book, we shall remain agnostic as to where these fundamental preferences come from. We shall take them as given and fixed.

This last assumption is actually less demanding than one might suppose. For instance, it does not imply that actors will not change their minds about what they want to do in a given situation when they obtain new information. To see that, suppose that we begin with the above rank-ordering which implies that waiting is the most preferred course of action. Suppose then that the actor receives information that if he lets the ultimatum deadline lapse, his opponent is almost certain to attack. As a result, he launches a preemptive strike. One might think that this indicates that the actor's preferences have changed, which would imply that taking them as fixed would be a serious problem. However, this is not so: all it means is that our original specification of preferences is not quite right, for it misses an important bit that determines the trade-off between waiting and preempting. In this instance, the actor prefers to wait if there is some reasonable chance that his opponent will not attack (because this would avoid war) but prefers to attack himself if war seems unavoidable. His estimate of the probability that his opponent will attack if concessions are not forthcoming is part of the expected consequences of the actions and as such must be included in the preferences. The correct way to specify the preferences, then, would be to give a full account of the contingencies.

One possibility is that the actor prefers to wait if there is at least a 50% chance that the opponent will not attack and prefers to preempt otherwise,

[3] In technical terms, preferences must be complete (i.e., actors must be able to consider all possible outcomes) and transitive (i.e., they do not admit logical contradictions like the one in the text). There are some more subtle requirements when it comes to rank-ordering risky choices. See von Neumann and Morgenstern (1947) for the classic treatment.

with both of these being preferable to capitulation. Now, the reception of new information that causes that actor to revise upward his estimate of the probability of war if he waits may well cause him to choose to preempt even though he would have chosen to wait in the absence of this information. Loosely speaking, his preference for preemption over waiting has changed. Strictly speaking, this is not the case: the *choice of action* changed because his estimate of its likely *outcome* changed because of the new information. But notice that the preference ordering is:

- If the probability that the opponent will attack when the ultimatum deadline expires without response is more than 50%, preemption is preferable to waiting and waiting is preferable to capitulation;
- If the probability that the opponent will attack when the ultimatum deadline expires without response is less than 50%, waiting is preferable to preemption and preemption is preferable to capitulation.

The actor's initial estimate was that there was less than a 50% chance of an attack if he waited, which meant that he would choose to wait. However, in light of the new information received, he has revised his estimate of that probability upward, and now chooses to preempt. Observe that his fundamental preferences have remained fixed even though his choice of action has changed. In other words, what the actor learns during the crisis can affect his behavior even though his fundamental preferences stay the same. In fact, this entire book is about how actors can alter the behavior of their opponent by manipulating information and the strategic environment.

In addition to having well-defined preferences, rational actors must pursue their goals to the best of their ability given the information they have and the constraints they must operate under. It is often supposed that rationality requires full information and the evaluation of all possible alternatives. That is not the case. As we shall see, the dynamics of military threats are highly contingent on uncertainty, both about the opponent's intentions and the outcomes of risky choices. It is true that to evaluate the best course of action, the actors will have to compare all the alternatives available to them, but as analysts we have already simplified the world by limiting the actors' choices. It is very likely that in reality the actors are similarly constrained to just a handful of options and they do not consider all possible options. In that sense, the model's limitations are perhaps more realistic than one might suppose. It is a fascinating puzzle to see how actors frame the problem and decide what actions are simply not to be considered. In this book, we abstract away from that and assume that they have arrived at a particular frame of reference. Whether this simplification is distorting or

not depends on how many relevant choices it leaves out, something that we would have to investigate in future work.[4]

There are very good reasons to assume that actors pursue their goals as best as they can. If this were not so, then behavior becomes unintelligible. We all tend to assume that on the average actors pursue their goals the best they can given the resources and information they possess and the constraints they must labor under. We then form some assumptions about their preferences over these goals, and the assumption that they pursue these goals enables us to form expectations about their behavior. Some interactions are so routine and involve preferences so stable across the population that we do not even have to think about it: our behavior is guided by rules of thumb rather than conscious decisions. We tend to avoid bumping into other pedestrians on the sidewalk because we know this sort of thing will be unpleasant for both, but we do not really do it consciously. More importantly, we assume that this preference is shared, which is why we do not expect to be bumped into as well.

More care has to be taken in situations that are riskier. For instance, we also tend to assume that a driver would rather not hit us when we cross the street. However, for him to act on this preference, he must be able to see us in time to react and be able to avoid us when he reacts. When we cross, we take these factors into account by asking how far the car is, whether the driver is likely to see us, and whether he will be able to avoid hitting us. Most of us are quite risk-averse when crossing the street, but even in situations when we run the auto gauntlet – as some of us who grew up in large cities with near constant traffic know – we assume that the drivers do not want to hit us. We then form an expectation about their behavior (that they will do what they can to avoid hitting us) and then we decide whether to cross and when to do so. Of course, we know that drivers know that we do not want to be hit either. The danger is that they might assume that we will jump aside to avoid the accident precisely when we are assuming that they will swerve for the same purpose. Which one of us has not deliberately

[4] Karsten et al. (1984, 8–10) discuss the shift of cost–benefit calculations during crisis and conclude that "situations that are characterized by … lags in identifying the national interest indeed pose several problems for any assumptions made prior to the crisis concerning the nature of a rational response." Their argument is that "the attribution of rationality to the decisionmaking process presumes that the parties on each side of the threat possess full information" and that "during periods of high tension, decisionmakers tend to adopt simplified cognitive structures; goals are reduced, and the range of perceived alternatives shrinks." In other words, rationality supposedly requires full information about all possibilities. As we have seen, rationality does not require full information and we will not be considering all possibilities. The very simplicity of the formal model – something that critics often fault it for – is in fact its strength here.

turned his head away from the driver to demonstrate that he "cannot see him"? The implication is that if we cannot see the car, we cannot act on our preference to avoid being hit. This leaves the driver with the remaining option to swerve and we "win" the unequal confrontation. The point is that all this behavior is predicated on the assumption that choices are, to a large extent, predictable because they connect rationally to preferences. The fact that people are sometimes hit is not remarkable. The fact that they are so seldom hit given traffic density is.

It is possible that actors make mistakes because of faulty interpretation of information, or wrong decision-making under stress, or incorrect implementation of correct decisions. However, it is difficult for me to believe that mistakes are systematic. It seems much more fruitful to treat them as deviations from the optimal course of action that *may* occur but in a more or less random fashion. If mistakes were systematic, one would have to wonder why actors do not correct them. Actors do make mistakes, true, but they also learn from these mistakes. Whether this results in them making fresh mistakes – as the famous quip that the generals are always preparing to re-fight the last war suggests – is a depressing (but unlikely) possibility. If actors do not act in their own interest, then we cannot hope to understand their behavior, much less form expectations about it. Anything is possible in such a world and therefore nothing is comprehensible. Every action can be "explained" by assuming that actors are deluded or inept or both. If this were true and actions were divorced from preferences, then it is a mystery why decision-makers spend so much time trying to divine the intent of their opponents and search their actions for meaning.

Although the concept of rationality used here is rather thin, the assumption that the players are unitary actors is more problematic. States are not individuals, they are collectives that comprise groups that themselves may be composed of other groups, all the way down to the individual. One may think of domestic politics as a way of aggregating these individual preferences in some sort of collective preference. Different political systems enfranchise different segments of the population in various ways. In the end, however, all that matters from our perspective is how these individual (rational) preferences translate into state preferences. It is well known that there is no way to guarantee that the preferences of a collective will be rational even if the collective itself comprises rational individuals. No way, that is, except taking one of these individuals' preference as the one for the collective (Arrow, 1970).

A complete theory of crisis behavior would have to take domestic politics into account. It would have to show how the (possibly competing) interests of various groups within the state coalesce to determine state behavior.

I will not do so in this book for two reasons. First, the underlying logic that will arise between two unitary actors will also be present in the more complicated interaction albeit at the lower level of aggregation. Whether it translates into analogous behavior at the state level remains to be seen but the fundamental problem will remain, whether or not the solution is the same. Second, when it comes to the types of disputes that may end in war – the crises where military threats are employed – decision-making is usually restricted to a small group of people at the highest level. Collective irrationality is less likely to arise in smaller groups, especially when their members are not too dissimilar in their preferences, which tends to be the case at the highest level of political power. Whether this assumption is too distorting remains to be seen but, as before, the individuals would have to confront the basic issues that arise from the unitary actor interaction regardless.

Throughout the book I will consider two-actor interactions only. This allows me to abstract away from many important considerations that would doubtless affect behavior in the real world. For instance, in the real world, decision-makers are likely to take into account the expected behavior of their allies, of potential other belligerents, or of non-aligned states that may be carefully monitoring the interaction. Limiting the model to only two actors serves to illuminate the features of military threats that have bearing on the puzzle of credible communication. This may not be the only concern policy-makers have in their confrontation but it will be among the most important ones. Hence, a thorough investigation of this isolated role of threats is a necessary first step toward a theory of their use.

The restriction of attention to a single encounter is made to remove any considerations for consequences of one's actions beyond the current crisis. Reputational concerns and long-term repercussions can enter this model only as part of the payoff specification. In other words, while it is possible to incorporate them, I will only do so by assuming that they can be reduced into the payoffs. A richer theory would model future interactions to see how the consequences one expects to follow tomorrow affect behavior today.

A more intriguing problem with the single-encounter assumption is its implications for equilibrium behavior. In this book, the analysis boils down to finding the optimal course of action in a crisis for each of two opponents. What we are looking for are strategies for the actors that are mutual best responses: neither actor has incentives to choose a different strategy given what his opponent is doing. The strategies then form an equilibrium because neither actor would want to deviate from his strategy. This approach depends on actors making accurate conjectures about the strategies of the other players. It is known that common knowledge of rationality is

not sufficient to guarantee that conjectures about behavior will be correct.[5] The upshot is that we do not know that rational players would necessarily choose actions that are prescribed by the equilibrium strategies. One common justification for expecting them to is that actors learn to play the game through repeating the interaction and successively refining their conjectures (Binmore, 2007). In our context, actors who have more experience with each other because they encounter the same game repeatedly will be more likely to behave how equilibrium logic predicts they should (provided our model is capturing the essence of the interaction). By assuming a single encounter, we effectively destroy the possibility for learning. If actors confront an entirely unfamiliar environment, then their behavior may deviate significantly from the equilibrium prescription.

I have several responses to this problem. First, as I explained above, the fact that I do not model repeated interactions does not mean that one cannot think of the model as representing one encounter among several similar ones. The model will make incorrect equilibrium prescriptions if it does not specify the actors' incentives properly, but that has nothing to do with their ability to play the game. Second, in high-stakes encounters where military threats are possible, decision-makers have very strong incentives to analyze their options much more carefully than we normally would in everyday life. It is more likely that they arrive at the optimal course of action and expect their opponents to do so, which means they should be able to make conjectures that are more likely to be correct on the average. Third, even in single-shot encounters of this type, decision-makers are likely to bring their prior experience and their knowledge of the opponent's past behavior into their analysis. Moreover, at this level decision-makers are often socialized through years of experience within relevant bureaucracies or decision-making groups which are likely to have imparted a code of behavior which is derived from the experience of the organization: corporate learning, if you will, that extends beyond the individual. In other words, decision-makers may be able to do quite well even in situations they have not faced before provided these situations are not totally unique and the decision-makers' backgrounds (or their advisors' backgrounds) include socialization within organizations that have longer memories and experience.[6] Finally, even if one does not buy into any of the defenses above, I am prepared to concede that this problem may limit the predictive power of game-theoretic models. However, in no way does

[5] Pearce (1984); Bernheim (1984); Brandenburger (1992).

[6] See Farkas (1998) for an argument that, if stretched a bit, may be used to support this line of reasoning.