CHAPTER 10

Commitments Based on Face

Face is like honor in that it helps one make a credible commitment. The commitment is reinforced by setting up a situation where backing down would mean losing face before some important audience, such as one's adversary, allies, or domestic groups. This chapter discusses how face can be manipulated to achieve credibility, parallel to the earlier treatment of honor.

The basic idea is that to bolster credibility, one establishes a connection between keeping the current commitment and similar decisions in the future. Often the connection is made by setting up focal symbolism. An example is the Korean tree-cutting incident of August 1976 (Head, Short, and McFarlane 1978; Kirkbride 1989). In the demilitarized zone separating the North and South, a poplar tree was blocking the view from the U.S. observation post, and a party of American soldiers set out to trim it. North Korean soldiers surrounded them and clubbed two Americans to death. Angry diplomatic exchanges followed, and finally, in "Operation Paul Bunyan," the United States dispatched a party of tree trimmers to do the job. Armaments were forbidden in the demilitarized zone, so they were accompanied by 64 Tae Kwon Do experts. As they were cutting the tree down, three B-52s from Guam flew along the border. The North Koreans let the operation proceed.

The idea that B-52s would bomb North Korea over a tree trimming seems far-fetched. Of course, they were not there for their military function. In part, they were a symbolic message, a warning, using a metonymy from the scenario of going to war. More relevant to this chapter, however, the B-52s set up focal symbolism. They were a symbolic precedent that put U.S. credibility in future crises at stake, since backing down now would influence expectations of what the United States would do in more important games of Chicken. The use of bombers strengthened the analogy between the incident and more serious confrontations that might involve strategic war to make the symbolic precedent more important.

Setting up a focal symbol is one way to invest face, and another is to hurl an insult. The insult would not be prototypical by the definition, since it would not be aimed at diminishing the other's face. The face in jeopardy would be the insulter's own. Iklé (1964) states that in early disarmament negotiations, the Soviet delegates often adopted an abusive tone to signal their unwillingness to give in on a particular point. Strong words suggest a strongly held position that the insulter is expected to stick to. Like the symbolism around the tree trimming, insulting is a form of bridge burning. One sets up the situation so that one cannot afford to back down.

The prelude to the 1991 Gulf War put Bush and Saddam in a Chicken contest, each trying to convince the other that he would carry through with the resolute move. Bush put his own face increasingly at stake through the use of stronger and stronger language. When he said early on that Iraq should leave Kuwait, he was investing a small amount of face. When he announced that the annexation of Kuwait "will not stand," the explicitness of his words made it harder to back away from them. When he compared Saddam to Hitler, he committed himself strongly. Insults like these are diplomatic moves, on the continuum with expressions of concern, cancellations of official visits, and solemn declarations. They lie at an extreme position of self-commitment.

The chapter gives a model of how insults could function strategically, in a "war of face." Since backing down from them would be costly, they prove one's willingness to fight for a prize. There are other ways to show one's resolve, such as engaging in a costly arms race or starting a violent conflict, and the chapter asks whether insult trading in the war of face is a relatively cost-efficient way to prove resolve. From an outside viewpoint, the answer seems obvious: real wars cost human lives, arms races cost money, but insults are just about face. From the decision maker's viewpoint, however, all three commodities are valuable. When costs are measured in a common medium, it will turn out that wars of face are no more or less cost efficient than most other ways of conducting the contest, but if the rules could be modified in certain ways, they could be made more cost efficient.

The War of Face

The *war of face* is a game that shows the skeletal structure of self-commitment through face. Two players compete for credibility. Instead of the back-and-forth escalation of the Gulf confrontation, each player chooses one move from the diplomatic continuum of self-commitment. They act simultaneously, then observe each other's move. The one who has chosen the stronger move takes the prize, while the other backs away from the commitment and thereby loses the face it committed. The amount of face committed and the value of the prize are measured in a common medium of money, and losing a certain amount of face is translated into a loss of money.

Players know their own value for the prize but have only a probability distribution over the other's. It will turn out that at equilibrium, there is a one-to-one relationship between the player's value for the prize and the amount bid, so the bid credibly indicates the value. When the bids are revealed, it is assumed that the lower valuer backs down and lets the other take the prize. Starting a conflict for the prize would be a waste at that point, since it is assumed that expectations of who would back down in that conflict have been set by the revelation of who values the prize more. This assumption will be amended later.

To bid, each chooses a nonnegative number, and, as noted, the prize goes to the higher bidder. The game has an unusual feature: only the loser pays. That is, if you made the lower bid, you receive nothing and pay the amount you bid, while the higher bidder pays nothing and gets the prize.[1] This loser-pays rule is the opposite of a normal auction where it is the winner who pays, but it fits the notion of investing face. When the United States prevailed in the tree-trimming incident, it lost no face, but the North Korean government, by backing down, lost face commensurate with the strong language it had used during the crisis. That is the feature of wars of face that the loser-pays rule means to capture.

The game is as follows.

STAGE 1: Each player learns its value for the prize, which is chosen in [0,1), uniformly and independently of the other's value.[2]

STAGE 2: Each simultaneously chooses a nonnegative number as a bid for the prize. Bidding "infinity" is not allowed.

PAYOFFS: The player making the higher bid is the winner, pays nothing, and receives the prize; the player making the lower bid pays the bid and gets nothing. Ties are resolved by choosing the winner equiprobably.

How much should each bid? Each player's strategy will be a function giving the amount to bid depending on the player's value for the prize, and an equi-

1. The notion of only the loser paying seems to have been discussed first by Riley and Samuelson (1981), who called it the "sad loser auction." Fearon (1994) gives a version of the game where, like here, the payment is interpreted as the reputational cost of backing down. The present model is dual to Crawford's bargaining model (1982), where the players choose how much to demand and the costs of backing down are then drawn randomly. Here the demands are given as the whole pie, and the players choose their costs of backing down.

2. A prize value of 1 is not allowed, since it would lead to an infinite bid.

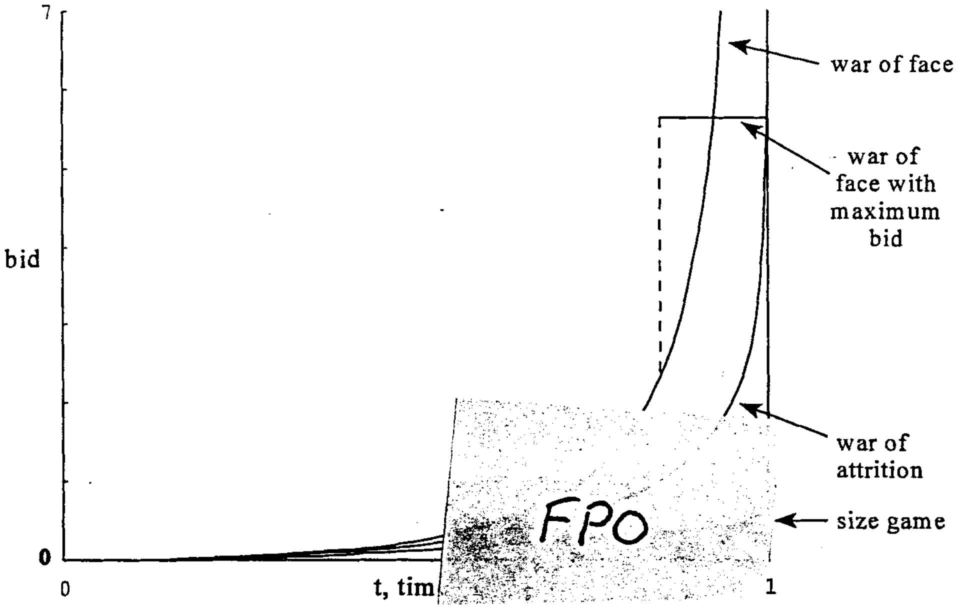

Fig. 24. Equilibrium bid as a function of value for the prize for four bidding games

librium will be a pair of such strategies. There is a unique symmetrical equilibrium,[3] it turns out, and it tells a player whose value is v to bid $\frac{1}{2} v^2/(1 - v)$. This function is shown in figure 24. For $v = \$.10$, which is a low value for the prize, the bid is half a cent, but at $v = \$.90$, one should bid \$4.05, several times what the prize is worth. This seems unreasonably high, but someone with a value at the 90th percentile is confident of winning and not having to pay anything. The optimal bid goes to infinity as v approaches its maximum of 1. In fact, it rises so quickly that the mean bid is infinity. This does not mean the auction is infinitely costly to a player, since the higher the bid, the less likely it is to have to be paid. The loser's expected bid, and therefore the expected amount paid, is \$.33.

Nicolson's comment (chap. 9) that the art of diplomacy is politeness seems contradicted by the pre–Gulf War exchange. The explanation may be that it was a contest where each was a high value type, so they bid very high to achieve self-commitment. The United States had much at stake compared to past military

3. A symmetrical equilibrium is one where both players use the same strategy. The game also has a family of asymmetrical equilibria where the two players use different functions of their values, and it has extreme equilibria where one player bids a very steep function of its value and the other bids nothing at all (appendix C).

actions, like those in Panama, Grenada, and Libya. It saw a threat to its oil supply and to its credibility at a defining moment in international politics. If B-52s were appropriate for the tree trimming, it would have wanted a very strong signal to show its motivation in the Gulf. Bush sent troops, recruited allies, and procured supporting resolutions from the UN Security Council, but still his concern was "getting the message through." Pronouncements that Saddam was worse than Hitler may have been issued because of the high demands of bidding in face. Of course, it is not just one's own high motivation that generates a high bidding—one also must be confident that one is more motivated than the other. This probably fit the Gulf confrontation, as each side probably saw its reasons as more justified and pressing.

Simple Signaling Contests

The war of face can be understood by comparing it with other games where players make simultaneous bids to prove their value for a prize. The games in table 2 share its features, in that they have two bidders whose values for the prize are uniformly and independently drawn from [0,1). The players bid simultaneously, and the higher bidder wins. However, the rules about who pays differ for each game. The winner might pay either the winner's bid, the loser's bid, or nothing, and the loser might pay one of these values. Three choices for each player give the nine different game rules shown.

The games are listed in order of their "volatility," the average bid made at the symmetrical equilibrium. The first three games possess no equilibrium but

TABLE 2. The Nine Simple Signaling Contests, with Statistics for Their Symmetrical Nash Equilibria, Assuming a Uniform Distribution on [0,1) for Each Player's Privately Known Value for the Prize (listed in decreasing order of mean bid)

Game	Winner Pays/ Loser Pays	Equilibrium: Bid for Prize Value v	Bid for $v = .9$	Mean Bid	Expected Net Payoff
War of words	Nothing/nothing	. . . No equilibrium . . .			
Reverse sealed-bid auction	Nothing/W's bid	. . . No equilibrium . . .			
Reverse size game	L's bid/W's bid	. . . No equilibrium . . .			
War of face	Nothing/L's bid	$\frac{1}{2} v^2/(1 - v)$	4.05	∞	1/6
War of attrition	L's bid/L's bid	$-v - \log_e (1 - v)$	1.40	1/2	1/6
Second-price auction	L's bid/nothing	v	.90	1/2	1/6
Bidding with a public bad	W's bid/W's bid	v	.90	1/2	$-1/3$
Sealed-bid auction	W's bid/nothing	$\frac{1}{2} v$	.45	1/4	1/6
Size game	W's bid/L's bid	$\frac{1}{2} v^2$	.41	1/6	1/6

appear at the top since the best response to any bid, no matter how high, is to bid higher. First comes the *war of words:* whoever names the higher number wins, and no one pays anything. In the next two games, each player bids with the other's money and it is not surprising that they have no equilibria. They are the *reverse auction* and the *reverse size game.* Their names come from two more sensible games that appear lower down.

The equilibria of the other games are not always unique, but all have only one symmetrical equilibrium. The *war of face* is next, and with only the lower bidder paying, it puts barely enough accountability on the bids to keep them finite. After that come various well-known games. In Maynard-Smith's *war of attrition* (1974), two animals wait by a watering hole, and one can drink as soon as the other leaves. This is equivalent to an auction where each animal's bid is the maximum time it is willing to wait, and both pay the bid of the loser. Next is the *second-price auction,* where the winner pays the loser's bid. An equilibrium strategy—in fact, a strategy that dominates all others—is to bid exactly the value v that you hold for the prize. This was the reason for Vickrey's interest in the game (1962): that its rules draw out the players' true values. Next is the *game of bidding with a public bad,* where each player pays the winner's bid.[4] It is disastrous; each player has an expectation $-1/3$. Since players bid their true values and pay the winner's bid, the winner gains exactly zero, and the loser has a negative payoff. No one would enter it voluntarily. If they could decide beforehand to throw away the prize instead of holding the auction, they would lose nothing.

The regular *sealed-bid auction* comes next in size of equilibrium bid. The equilibrium is that players bid half their values. It makes sense that they would bid below their values: they can assure themselves at least zero by bidding zero, and they will only bid something in hopes of a profit. Last comes the *size game,* so-named by Maynard-Smith and Brown (1986) for a contest between two animals who want to be physically small for the purpose of surviving on the available food but bigger than their rival. In the war of attrition, one animal stopped waiting as soon as the other left, but an animal cannot observe its rival's size and shrink to an optimum response. The fact that each is stuck with its chosen size is a motive to bid lower than in the war of attrition. The game is often called the "all-pay auction," and Hirshleifer and Riley (1992) discuss it as the "secret arms race," drawing an analogy with the United States and Germany during World War II, where each chose an amount to spend on developing an atomic bomb

4. The name would apply as well to any game where the players pay the same amount—one where they pay the sum of their bids, for example.

without knowing the other's amount. It is a continuous and simultaneous version of the dollar auction (O'Neill 1987).

Is the War of Face an Efficient Way to Settle a Conflict?

The conflicts suggested by these games are truly costly, sometimes in dollars and sometimes in human lives. The analysis sidesteps the question of whether it is better to lose face in a war of face or lose lives in a literal war of attrition and asks which rules are least costly. It takes the decision maker's viewpoint in which the various commodities used to bid are valuable, and it compares the cost in the common medium. The last column of the table gives the mean net payoff, that is, what a player would expect to gain averaged over all prize values of the player and the adversary. All but one give the same value, ⅙. The games in the table have the feature that the players compete for the prize and in doing so reveal their true values for it. The player with the higher value wins, and this promotes efficiency, but the test to determine the higher valuer is a costly one since it involves paying bids. All of them have a lower expectation than random allocation as will be shown. The bidding competition eats up the benefit from correct selection of the winner, and more.

One can compare the table's rules with some other allocation rules that are not auctions. One alternative is to simply give the prize to the player who wants it more. A player's expectation would be ⅓, as shown in appendix C. It cannot be beaten by any other rule, but it can be implemented only if the judge knows which player has the higher value, and often there is no way to know this. A second possibility would assign the prize randomly to one of the players. Each would have expectation ¼ (since a given player gets the prize with probability ½ and has a mean value for it of ½). This is still better than the simple signaling contests in the table.

Almost all the solvable bidding games have the same expectation, ⅙. This is not a coincidence; it is an instance of a broad fact discovered in the study of auction design, the *revenue-equivalence theorem* (Myerson 1981). It is really a group of related results, using slightly different conditions on the rules of the auction and the information available but coming to a common conclusion that any auction rule in the class extracts the same expected payment from participants. One version (Milgrom and Weber 1982) applies to auctions where the bidders have privately known values for the prize, independently determined according to a single distribution. It states that if two auction mechanisms have equilibria where the prize goes to the player with the highest value, and where a player who values it to the minimum degree of zero gains zero, then the expected payments to the auctioneer are the same at these two equilibria. "Expected payment to the auctioneer" in a literal auction is a simple negative linear function of expected net payoff to a player in our conflict games (since the prize always goes to the higher valuing player). It follows that the expected net payoffs to our players are constant over the allocation mechanisms that satisfy the theorem's conditions.[5]

A good way to understand the revenue-equivalence theorem is to look at an auction rule that violates its conditions. Bidding with a public bad, in the table, evidently does not satisfy the conditions because a player expects $-1/3$, rather than $1/6$. In this game, a player can hold zero value for the prize but still suffer a loss, so it violates a condition of the revenue-equivalence theorem. This fact explains the lower expectation.

Variants of the War of Face

Those who developed the strategic theory of auctions usually took the viewpoint of the auctioneer and were interested in maximizing revenue. They saw the revenue-equivalence theorem as an impediment to taking in more money and often looked for special rules that exempted an auction method from the theorem. In the context of international conflict, the payments go into no one's pocket. They are a waste, so we want to *minimize* "payments to the auctioneer." Some variants of the war of face will be outlined, which are chosen to imitate aspects of real conflicts and possibly reduce the players' costs.[6]

War-of-face variants with expectation different from $1/6$ can be generated by violating different conditions of the revenue-equivalence theorem. In the *war of face with a maximum bid,* the rule is that neither can bid higher than M. This would happen if insults could not go beyond a certain point of nastiness, or if, in some sense, M represented all the face a party has to invest. One can calculate the symmetrical equilibrium, and figure 24 shows it for the case of $M = 6$. A player with v from 0 to $M/(M + 1)$ bids just as it would in the original game, and a player with v from $M/(M + 1)$ to 1 bids the maximum M. Compared to the original game, some players are lowering their bid because they have to, and others are raising their bid to the maximum. The intuition behind the latter be-

5. The equivalence among mechanisms is true not only for the expected gains to players before they know their values for the prize but also afterward.

6. The different signaling contests will have different expectations if the players' distributions are not independent. Milgrom and Weber (1982) provide some results for the first and second prize auction, and Krishna and Morgan (1994) treat the war of attrition and the size game.

havior is that they are taking a greater risk in hopes of tying with the types of players who are clustered at the maximum, who would otherwise beat them. This equilibrium is not constrained by the revenue-equivalence theorem since the prize is assigned randomly when both bid M, and so the higher valuing player may not get it. Would the expected payoff be higher, lower, or equal to ⅙? It might rise since the higher players are restrained from high bids; on the other hand, it might fall because the middle group is bidding higher than before and because of the inefficiency of the player with a lower value sometimes getting the prize. Overall, it turns out, the first effect outweighs the latter two—a player's expected payoff is higher with a ceiling on the amount of face bid. For a large M, the game is close to the original war of face, and the expectation is only slightly higher than ⅙, but as M decreases to zero, the expectation rises to ¼ as the rule goes effectively to a random allocation.

In the *war of face with unreliable commitments,* each player i tries to commit face but succeeds only with probability $p_i < 1$, as if sometimes an insult does not come off. Failing to commit, which has probability $1 - p_i$, is equivalent to bidding zero. This version violates the theorem's condition that the prize always goes to the higher valuer. At the equilibrium, bids submitted are exactly as in the original game. This makes sense, since the exact value of a nonzero bid makes a difference only when both are successfully committed—in other words, only when chance has them play the original game. The game is more efficient than the original: as both p_i go to 0, the mean payoff goes from ⅙ up to ¼, as the game moves to the random allocation rule.

The original war of face misses one feature of the 1990–91 Gulf crisis. In the game, whoever bids lower backs down, but in the Gulf neither one backed down. The *war of face with the option of not backing down* is close to a game discussed by Fearon (1994). It postulates that after a very strong commitment it is not worthwhile to back down, even if the other has bid higher than you. Suppose the cost of a conflict is C to both, so a side who bids above C and loses will be motivated to let the conflict start, rather than pay the bid. In a conflict, neither player gets the prize. Making a bid in this game is not just sending a credible signal of your value but burning a bridge. Again, the revenue-equivalence theorem does not apply, since the higher valuer sometimes misses the prize. The equilibria are simple and essentially all alike: a player makes a bid according to the original formula, as long as the bid is below C. If the formula calls for a bid higher than C, the player makes some bid or other higher than C. The exact value does not matter since it will not be paid, and in fact the symmetrical equilibrium strategies of the original war of face would do as an equilibrium here too.

The resulting expected net payoff is just below ⅙ for high values of C but goes to zero as C is reduced. Unlike the previous modifications, this is a change that hurts the parties' welfare.

The *war of face with misjudged signals* is also suggested by events around the Gulf War. After an Iraqi threat to retaliate against sanctions, according to the *New York Times* (September 1990), U.S. officials "took pains to note that the threats were conveyed within an established context of intimidating public displays and comments emanating from Baghdad. . . . 'We get a daily diatribe from Saddam and this is another one of them.'" In the model, this could be interpreted as the United States believing that with such talk as standard, Iraq is not really staking much face, and its words should not be seen as high bid. The sender and the recipient might judge the strength of the signal differently. One can imagine a situation in which the strength of i's insult from i's viewpoint is multiplied by some randomly determined constant K_i when j hears it, and j's insult is multiplied by K_j in i's ears. If a side judges its own insult as sent to be stronger than the other's as received, it does not back down. This opens the possibility that neither party backs down or both do. In a stalemate, neither gets the prize and both are charged C, or if they both back down, the winner is chosen at random. Since it is assumed they are unaware of the biases, they see the game as the original one and make bids according to the original equilibrium. A simulation gives some values of a player's expectation as a function of C, assuming that K_i and K_j are independently uniformly distributed on the interval (0,2): for $C = 0$, it is .163; for $C = 1$, it is .075; and for $C = 5$, it is $-.277$. This variant is far worse than the original. Noise in the communication channel is severely harmful.

Dangers in the Diplomacy of Insults

The pre–Gulf War confrontation seems close to a war of face, especially considering the stance of President Bush. In an age of pervasive world news, it is tempting for a president to use this tone—Bush was able to rally domestic support for his policies, appear resolute to allies, and engage in the war of face with Saddam all at once. However, it should be clear that the diplomacy of insults was a failure. Saddam did not withdraw from Kuwait, and Bush, having overcommitted his own face, paid a price at the polls when he did not pursue the war to its end.

Why did the tactic not work? The last variant of the game illustrates a significant problem with insults as commitment devices: that they can get distorted when they cross cultures. Relevant to the Gulf War, much has been written on how the style of insulting in Arabic cultures differs from that of European cultures (e.g., Parkinson 1989; Cohen 1987, 1990; Abu-Zahra 1970), but neither leader seemed to understand the difference well enough to use insults for diplomacy. Some Arabists in the United States were worried about Bush's "worse-than-Hitler" talk, but Bush held little consultation with specialists within his administration. White House aides gave the questionable explanation that the crisis was a global, not a regional, crisis (Goshko 1990).

There is a further danger. In a game of Chicken, each has a motivation to convince the other that it disbelieves its claimed degree of commitment. When a rival dismisses one's insults, the natural response is to raise their intensity. This action-reaction process could run out of control.

Also, each side might misunderstand the other's view of the game. Saddam Hussein's speeches and interviews suggested that Bush's tone was not a tactic to establish self-commitment in the current crisis but a continuation of how Western powers had always treated Arab and Islamic nations and small states.[7] Manuel Noriega had been treated this way, dismissed by Bush as an "amateur" just before the invasion of Panama. Saddam seemed to be interpreting the insults as having the prototypical purpose, that of diminishing his face.

Musing on the ironies of language, Kenneth Boulding liked to tell an upside-down version of the Tower of Babel. When the builders set to work, everyone spoke different tongues. They would gesture for help and point to what needed to be done, and the building rose toward heaven. To thwart their ambition, God gave them all one language. They fell to bickering, and the work was abandoned. Not all communication is helpful. Insults can have a strategic purpose in theory, but using them that way is unreliable. The models bring out the assumptions that must be satisfied for this tactic to work, and they are far from real politics. Nicolson stated that insults are unusual in diplomacy, and this norm has evolved for a good reason.

7. Examples of U.S. insults toward Presidents Nasser of Egypt and Mossadegh of Iran are given by Neff (1981, 258–62), and Parker (1993, 105).

Honor, Symbols, and War
Barry O'Neill
http://www.press.umich.edu/titleDetailDesc.do?id=14453
The University of Michigan Press