ECONOMIC RESEARCH REPORTS

GAME THEORY AND EMOTIONS

by Steven J. Brams

RR # 95-23 July 1995

C. V. STARR CENTER
FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

Abstract

Emotions are feelings like anger, jealousy, and love that arise spontaneously. Although they would not appear to be products of rational calculation, that view has recently been challenged by analysts from several different disciplines. A game-theoretic model is developed to analyze structural features of *frustration*, which is engendered by a player's being in an unsatisfying situation and feeling an inability to escape it because of a lack of control.

Of the 78 distinct 2 x 2 strict ordinal games of conflict, 57 are "conflict games" that contain no mutually best outcome for the players. Of these, 12 are "frustration games," in which the choice of a dominant strategy by one player inflicts the two worst outcomes on the other (frustrated) player; six are "self-frustration games," in which it is the player with the dominant strategy who is frustrated by the best response of the other player. Altogether, there are 17 different games of frustration or self-frustration (one is common to both classes), which is 30% of all the conflict games.

In four of the 12 frustration games, the frustrated player can, if it has "threat power," gain some relief, which is illustrated by Aristophanes' play, *Lysistrata*, in which the frustrated women induced the men to stop fighting by abstaining from sex. In the six self-frustration games, the player with the dominant strategy can always induce a better outcome, called a "nonmyopic equilibrium," based on the "theory of moves." Shakespeare's *Macbeth* illustrates how a self-frustrated Lady Macbeth incited her husband to kill King Duncan by choosing her dominated strategy. In both cases, the frustrated and self-frustrated players, who start out at inferior outcomes, move initially to still worse outcomes—and explode in anger—to effect better outcomes. Conditions are given for the rationality of such moves.

Game Theory and Emotions[1]

To say that the bulk of social activity consists of playing games does not necessarily mean that it is mostly "fun" or that the parties are not seriously engaged in the relationship (Berne, 1964, p. 17).

1. Introduction

Emotions like anger, jealousy, and love are spontaneous feelings that practically all of us experience at one time or another. Although there may be good reasons for us to be angry, jealous, or fall in love, these feelings, especially when they overtake us suddenly, seem not to be the product of rational calculation. Rather, they overpower us, so to speak, which seems the antithesis of the careful means-ends analysis that we normally associate with rational choice. Indeed, Gerlernter (1994, p. 29) argues that "you cannot choose your emotions. Emotions choose themselves," which suggests that emotions have no rational basis.[2]

In a spate of recent books and articles, economists and game theorists such as Hirshleifer (1987, 1994), Frank (1988), Gilboa and Schmeidler (1988), Geanakoplos, Pearce, and Stacchetti (1989), Rabin (1993), Howard *et al.* (1993), and Howard (1994), philosophers such as de Sousa (1987) and Solomon (1993), a political scientist (Elster, 1994), a neurologist (Damasio, 1994), and sociologists and other social scientists ("Emotions and Rational Choice," 1993) have argued that expressing emotions is wholly compatible with acting rationally. Our goals, according to some of these analysts, are not always the narrow ones postulated by economic theory but broader, if

---

[1]I thank Eva Brams and Ben D. Mor for valuable comments on an earlier draft, and the C. V. Starr Center for Applied Economics at New York University for its support,

[2]Gerlenter (1994, pp. 27-28) defines an emotion to be "a mental state with physical correlates; it is a *felt* state of mind, where 'felt' means that signals reach the brain that are interpreted as bodily sensations," creating an "affect link" to thought.

somewhat inchoate and ineffable, ends (e.g., about living meaningful lives). Moreover, we may be passionate in our pursuit of these ends, which may undermine the achievement of more commonplace goals, such as ensuring our personal safety or economic security.

How can we explain these passions, and the sudden welling up of emotions that give rise to them? In this paper I focus on the emotion of frustration, which arises from being "prevent[ed] from accomplishing a purpose or fulfilling a desire" (*American Heritage Dictionary*, 1980). I hypothesize that *people become frustrated when they are in an unsatisfying situation and feel an inability to escape it because of a lack of control.*

Put another way, when a person cannot help himself or herself but must depend on the actions of others who are not helpful, this person will tend to react emotionally, suddenly becoming aroused. Feeling hemmed in, without room to maneuver, he or she explodes in anger—the most common behavioral response to frustration—to try to escape. In fact, a person's anger may well be aggravated if it is not impersonal forces, but a "freely acting agent" (Frijda, 1986, p. 198), that provokes him or her.

Psychological theories of frustration, as well as other emotions, say little if anything about the *structural features* of a situation that incite people to become emotional. By contrast, game theory provides powerful tools for identifying situations in which people experience a lack of control—and those in which they can regain it.[3]

By definition, a game is an interdependent decision situation, with the outcome depending on what all so-called players do. Although measures of

---

[3]Here I am exclusively concerned with emotions that arise in interpersonal situations rather than "existential" (Lazarus and Lazarus, 1994, pp. 41-66) or "self-conscious" (Lewis, 1995) emotions, like anxiety, guilt, or shame. The latter, which concern the way we see ourselves, may be quite independent of the actions of others in a game.

the degree of interdependence of player choices in a game have been proposed (Thibaut and Kelly, 1958; Kroll, 1993), they are quantitative and cannot be applied to the strict ordinal games that I analyze here, in which players are assumed only to be able to rank outcomes from best to worst but not attach numerical values, or cardinal utilities, to them.

I limit the analysis to strict ordinal games in which there are just two players, each of whom has two strategies, which defines a 2 x 2 normal-form, or matrix, game. Although there are 78 distinct 2 x 2 strict ordinal games (Rapoport and Guyer, 1966), 21 are games in which there is a mutually best outcome, which are unlikely to create frustration for the players.

Of the remaining 57 games, which I call *conflict games*, I distinguish two classes, the first of which subsumes 12 "frustration" games and the second of which subsumes six "self-frustration" games. Altogether, there are 17 different games (one game is common to both classes), which is 30% of all conflict games.

In frustration games, a player's lack of control takes the form of an "advantaged" player's having a dominant strategy that inflicts the two worst outcomes on the "frustrated" player. In the self-frustration games, it is the "self-frustrated" player who has the dominant strategy; the advantaged player does not, but his best response to the self-frustrated player's dominant strategy induces his next-worst outcome.[4]

I consider whether, based on the "theory of moves" (Brams, 1993, 1994b, 1995a, 1995b; Brams and Mattli, 1993), there is any escape from an inferior outcome for the frustrated or self-frustrated player. It turns out that

---

[4]How gender is used in the subsequent analysis will be explained in section 2.

if the frustrated player has "threat power," then she can induce a better outcome for herself with a "deterrent threat" in four of the 12 frustration games.

The situation is better for the self-frustrated player in the six self-frustration games. Although threat power is effective in these games, she can induce a "nonmyopic equilibrium" better for herself, and sometimes for the other player as well, without resort to such power. However, implementing this superior outcome without threat power requires that players think ahead about their moves and countermoves, based on rules of play (to be described) different from those postulated by standard game theory.

To illustrate the role that emotions play in the frustration and self-frustration games, I turn to two works of literature (Brams, 1994a), each of which has strong political overtones. *Lysistrata,* Aristophanes' most popular play, describes how women in ancient Greece, immensely frustrated when their men went off to war—leaving them lonely and with little support for their children—were moved to use abstinence from sex as a weapon to induce the men to stop fighting. In *Macbeth*, Shakespeare describes how Lady Macbeth, seething at Macbeth's initial hesitation to seek his phrophesied royal destiny, flew into a rage, belittling Macbeth's manhood, to incite Macbeth into killing King Duncan.

In both games, the frustrated and self-frustrated players, stuck at inferior outcomes, initially moved to still worse outcomes, exploding in anger to try to effect an ultimately better outcome. By linking the theory of moves to emotions expressed along the path to equilibria, one can identify conditions, in the dynamic play of a game, when it is rational to express emotions like frustration.

I emphasize the *process* by which a stable outcome is achieved, rather than the outcome itself, because once a game has stabilized at an equilibrium outcome, emotions are less likely to be expressed. Although emotions may be feigned, I take them to be genuine in the analysis that follows, which is plausible in games in which players have little recourse but to escalate a conflict in order to try to escape their frustration.

2. The Frustration Game

The "generic" Frustration Game, which is shown at the top of Figure 1,

---

Figure 1 about here

---

is a 2 x 2 game in which the row player, Advantaged (A), has two strategies, s_1 and s_2, and the column player, Frustrated (F), has two strategies, t_1 and t_2. The payoffs to the players at the resulting four possible outcomes are given by ordered pairs (x_{ij}, y_{ij}), where x_{ij} is the payoff to A, and y_{ij} the payoff to F, when A chooses strategy s_i and F chooses strategy t_j ($i, j = 1$ or 2).

"Advantaged" does not mean that A has any special prerogatives or privileges; rather, he has a dominant strategy.[5] Associated with his dominant strategy is a Nash equilibrium, which in a 2 x 2 strict ordinal game is unique (Hamilton and Slutsky, 1993, p. 50, Lemma 1).[6]

F, who may or may not have a dominant strategy, is "frustrated": her two worst outcomes are associated with A's dominant strategy, which he

---

[5]Henceforth I assume that A is male and F (and later SF in the Self-Frustration Game) is female.

[6]I could have called the row player simply "Dominant," but this would not have distinguished him from the column player, who sometimes has a dominant strategy, too. Moreover, in the Self-Frustration Game in section 5, it is column, the self-frustrated player, who has the dominant strategy—row does not—so calling the row player in the Frustration Game "Dominant" would create confusion.

would presumably choose.[7] By contrast, I will show shortly that, with one exception, A never suffers his two worst outcomes when F chooses a dominant strategy (if she has one).

More formally, the Frustration Game satisfies the following three conditions:

1. *Dominance for A.* Without loss of generality, assume that s_1 is A's dominant strategy, so

$$x_{11} > x_{21}; \quad x_{12} > x_{22}. \tag{1}$$

That is, whatever strategy F chooses (t_1 or t_2), A prefers his payoffs associated with s_1 to those associated with s_2, making s_1 A's unconditionally better strategy.

2. *Nash equilibrium.* Without loss of generality, assume that (x_{11}, y_{11}) is the unique Nash equilibrium that is associated with A's dominant strategy of s_1, so

$$y_{11} > y_{12}. \tag{2}$$

That is, F prefers her payoff at (x_{11}, y_{11}) to that which she would receive at (x_{12}, y_{12}) if she switched from t_1 to t_2. Similarly, A prefers (x_{11}, y_{11}) to (x_{21}, y_{21})—and so would not switch from s_1 to s_2—because of the dominance of s_1, assumed in inequalities (1).[8]

---

[7]I will discuss shortly under what circumstances A would *not* choose his dominant strategy, enabling F to escape her lack of control in this game.

[8]Technically, a Nash equilibrium is not an outcome like (x_{11}, y_{11}) but the strategies of the players that lead to this outcome. For convenience, however, I will identify equilibria—both Nash and later nonmyopic—by the outcome, or payoffs at this outcome, rather than by the strategy pair that produces this outcome.

3. *Lack of control.* Let 4 indicate the best payoff to a player, 3 the next best, 2 the next worst, and 1 the worst. Thus, the higher the number, the greater the payoff. But because these payoffs are *ordinal*, they indicate only an ordering of outcomes from best to worst, not the degree to which a player prefers one outcome over another.

F's lack of control, and consequent frustration, stems from her two worst outcomes (1 and 2) being associated with A's dominant strategy of s_1. From inequality (2) it follows that

$$y_{11} = 2; \quad y_{21} = 1.$$

Depending on whether $y_{21} = 3$ and $y_{22} = 4$, or $y_{21} = 4$ and $y_{22} = 3$, F may have two different complete orderings of his payoffs.

On the other hand, the dominance of s_1 for A, as given by inequalities (1) that define a partial ordering of his payoffs, admits six different complete orderings:

$$x_{11} > x_{21} > x_{12} > x_{22}; \qquad x_{12} > x_{22} > x_{11} > x_{21};$$
$$x_{11} > x_{12} > x_{22} > x_{21}; \qquad x_{12} > x_{11} > x_{21} > x_{22};$$
$$x_{11} > x_{12} > x_{21} > x_{22}; \qquad x_{12} > x_{11} > x_{22} > x_{11}.$$

The two orderings of F, and the six orderings of A, define a total of 12 different strict ordinal games subsumed by the "generic" Frustration Game, which are divided into two classes in Figure 1.[9]

I summarize properties of these games with three propositions, the last of which distinguishes between the two classes of games:

---

[9]The *generic* Frustration Game is the game that satisfies the three conditions, whereas the numbered games are specific instances of the generic game. The numbers shown above each numbered game are those given in the classification schemes of Brams (1994b) and, in parentheses, Rapoport and Guyer (1966).

Proposition 1. *In six of the 12 specific games subsumed by the generic Frustration Game, F has a dominant strategy (indicated by vertical arrows for games 5, 6, 10, 11, 32, and 35 in Figure 1), but only in game 32 (Prisoners' Dilemma) does F's choice of this strategy lead to A's two worst outcomes (1 and 2).*

Thus, the apparent frustration that F experiences when A chooses his dominant strategy cannot be duplicated by F when she chooses her dominant strategy (if she has one), except in Prisoners' Dilemma.

Proposition 2. *In the 12 specific games subsumed by the generic Frustration Game, A's ranking of his payoff at the Nash equilibrium is better than F's ranking in 10 games and the same as F's in two games (games 28 and 32).*

Of the 10 games in which A's ranking is better, six are games in which A obtains his best payoff (4) and four are games in which he obtains his next-best payoff (3). In the two games in which the players tie, they both obtain their next-worst payoffs (2).

So far the picture looks pretty grim for F: she always obtains only her next-worst payoff (2) at the unique Nash equilibrium in the 12 games, whereas A does better than F—at least in terms of comparative rankings (I do not assume interpersonal comparison of utilities)—in 83% of the games. Because A obtains a higher-ranked payoff in these games, F might not only be frustrated but also envious of A. But I leave for another paper a game-theoretic explication of envy as an emotion.

Fortunately for F, there is a way out of her plight in four of the 12 games if she possesses "threat power" of the "deterrent" kind (Brams and Hessel, 1984; Brams, 1994, ch. 5).

Proposition 3. *The exercise of deterrent threat power by F in the class II games in Figure 1 induces an outcome [i.e., either (3,3) or (3,4)] for F at least as good as and sometimes better than that which A obtains at this outcome (in terms of comparative rankings).*

In a 2 x 2 strict ordinal game, a necessary condition for a player to have a *deterrent threat* is that he or she can threaten the choice of a strategy that leads to the other player's two worst outcomes. For example, in game 22 in Figure 1, F can threaten the choice of t_2, which inflicts upon A a payoff of either 1 or 2; if F carried out her threat, A would presumably choose s_1, which yields (2,1), which is called the *breakdown outcome*, rather than choose s_2, which yields (1,4). If there is another outcome, associated with F's other strategy (t_1), that is better for both players (i.e., Pareto-superior) than the breakdown outcome, then this is the *threat outcome* F can induce if she has "threat power."

Threat power is the ability of the threatener (F in this case) to withstand the Pareto-inferior breakdown outcome [(2,1) in game 22]—if she is forced to carry out her threat—and thereby induce the other player (A) to accept a Pareto-superior threat outcome.[10] In fact, both (4,2) and (3,3),

[10]A threat may be communicated before the players choose strategies, or it may be communicated if players are at a state that F considers unsatisfactory. In game 22, for example, assume that the players are at (4,2). F's *threat* is to switch to t_2 if A is unavailing and continues to choose his first strategy. This would cause the game to move from (4,2) to the Pareto-inferior breakdown state of (2,1)—at which, presumably, the emotional temperature of the conflict would rise—from which F can induce (3,3) with threat power (for reasons given next in the text).

associated with t_1, are Pareto-superior to (2,1). But because F prefers (3,3), this is the threat outcome she can induce with threat power, which would supersede the Nash equilibrium of (4,2).

A similar analysis shows (3,4) to be the threat outcome that F can induce in game 35, with (2,1) again the breakdown outcome. In games 28 and 32, the breakdown outcome is (2,2), which is also the Nash equilibrium. Consequently, it is in *both* players' interest that F be able to induce, with threat power, (3,4) in game 28 and (3,3) in game 32.

In fact, if A has threat power, he, too, can induce (3,4) or (3,3), respectively, in these games by threatening, with the choice of s_1, F's two worst outcomes. Clearly, each player's exercise of threat power in games 28 and 32 would have a salutary effect, leading to an outcome preferred by both players. By contrast, in games 22 and 35 there is a conflict of interest between the players, with A's preferring to induce (4,2), which he can do with a "compellent threat" if he has threat power, and F's preferring to induce (3,3) or (3,4) with a deterrent threat.[11]

The possession of threat power enables F to relieve her frustration in all the class II games, making these games, in a sense, *controllable* (recall

---

[11]A *compellent threat* is one that involves not just the threat of choice, but the actual choice, of a strategy. In making this choice, the threatener compels the threatened player to choose between a Pareto-inferior breakdown outcome and a Pareto-superior threat outcome. For example, in game 22, by choosing s_1 and refusing to budge from it, A can force F to choose between (4,2) and (2,1); if A has threat power, F will choose t_1, which in this game simply reinforces the choice of the Nash equilibrium. In both this game and game 35, threat power is *effective* in the sense that the player who possesses it does better than if the other player possessed it. By comparison, in games 28 and 32, threat power is *irrelevant*, because the outcome induced is the same whichever player possesses it. (However, it is not irrelevant in the sense that one player's possession of it enables both players to escape the Pareto-inferior Nash equilibrium in these games.) The distinction between "compellence" and "deterrence" was first made, informally, in Schelling (1966). Brams and Hessel (1984) formalized this distinction in terms of threat power; Aumann and Kurz (1977) incorporate the power to hurt others into a different game-theoretic model.

that it is a *lack* of control that leads to frustration). By comparison, F's lack of control in the class I games may turn her frustration into *despair*, or a feeling of helplessness. Of course, A, too, will be frustrated by the choice of (2,2) in games 28 and 32; but he, also, can help the players escape this Pareto-inferior outcome if he is able to exercise threat power. Like F's deterrent threat to choose t_1, A's threat involves threatening to choose s_1—and inflicting on F one of her two worst outcomes—in order to induce the Pareto-superior (3,4) outcome in game 28 and (3,3) in game 32.

The frustration of players caught in real Prisoners' Dilemmas (game 32) is well-known. What is less appreciated is that a deterrent threat on the part of one or both players in this game, as well as game 28, offers both an escape from (2,2), but it must be *credible*: the threatener must be able and willing to carry it out.

Such a threat will not be credible unless the game is likely to be repeated, because a threat that leads to a breakdown outcome is, by definition, irrational to carry out in any single play. But if the game is a continuing one, it may be rational to suffer this outcome in early play in order to build up one's reputation for "toughness" in the future. With one's reputation established, the need to carry out threats will be obviated, rendering the earlier suffering at breakdown outcomes worthwhile.

3. Overcoming Frustration with a Credible Threat: The Case of *Lysistrata*[12]

Aristophanes' comic yet biting anti-war play, *Lysistrata* (411 B.C.E), was written in one of Athens's darkest hours, after the total destruction of

[12] This section is adapted, with significant changes, from a student paper of a former NYU undergraduate, Elana Zaas.

her expeditionary force in Sicily, the threat of invasion from Sparta and Syracuse, and the defections of some of her allies. It is a "dream about peace" by a playwright who believed "any peace must be satisfactory to both sides, . . . and the women of both sides have to cooperate in bringing it about" (Sommerstein, pp. 177-178).

While *Lysistrata* is part of the Old Comedy of Athens, the dream of Aristophanes is not pure fantasy, tied to real problems. For one thing, the calculations of its characters are unremittingly strategic, as in real life. For another, powerful emotions, especially frustration, fuel the central action of the play, which gives *Lysistrata* some of the weight of the Greek tragedies of Aeschylus, Euripedes, and Sophocles, with which, perhaps, people can more readily identify.

The play begins when Lysistrata encounters her neighbor, Calonice, who recognizes that Lysistrata is "bothered." In fact, Lysistrata admits to being "furious," having "called a meeting [of women] to discuss a very important matter [while] our husbands . . . are all still fast asleep" (p. 180). She has hatched a plan, after "many sleepless nights," to "save Greece" from a never-ending battle between Athens and Sparta, which had left the lives of women of both city-states lonesome and desolate, and their male children prospective soldiers who might be killed like their husbands.

But Calonice mocks her:

> The women!—what could they ever do that was any use? Sitting at home putting flowers in their hair, putting on cosmetics and saffron gowns and Cimberian see-through shifts, with slippers on our feet? (p. 181).

Yet these are exactly the weapons that Lysistrata intends for the women to use in order to get the men to "no longer lift up their spears against one another . . . nor take up their shields . . . or their swords" (p. 182).

Lysistrata is frustrated, however, because "the ones I was most counting on" haven't shown up (p. 182). Eventually they drift in, and Lysistrata lashes out at their collective plight, with more than a touch of acerbic wit:

> The fathers of your children—don't you miss them when they're away at the war? I know not one of you has a husband at home . . . . There isn't anyone even to have an affair with—not a sausage! (p. 184).

When, after first hesitating, Lysistrata blurts out her plan—"we must give up—sex"—to stop the carnage of war, the stage directions say there are "strong murmers of disapproval, shaking of heads, etc. Several of the company begin to walk off" (p. 184). Indeed, there is a chorus of "I won't do it" from several women, including Calonice, who says she'll "walk through the fire, or anything—but give up sex, never! Lysistrata, darling, there's just nothing like it" (p. 185).

Next there is a discussion of how women, by threatening abstinence, can use their feminine wiles and charm to induce the men to give up fighting. The women, however, worry that the men might divorce them or, worse, "drag us into the bedroom by force." But Lysistrata points out that, even in this case, the women can be

> as damned unresponsive as possible. There's no pleasure in it if they have to use force and give pain. They'll give up trying soon enough. And no man is ever happy if he can't please his woman (p. 186).

After still more cajoling from Lysistrata, the women take an oath to refrain from sex "with my boyfriend or husband . . . wearing my best make-up and my most seductive dresses to inflame my husband's ardour" (p. 188).

The women then retire to the Acropolis and close and bar the gates to the men, vowing not to open the gates "except on our terms" (p. 189). In the confrontation that follows, the men are heard to say such things as,

> "Euripides was right! 'There is no beast so shameless as a woman'!" (p. 195); and "Back to your weaving, woman, or you'll have a headache for a month" (p. 201),

but their derision is to no avail.

Lysistrata remonstrates to the men that the women will "unravel this war, if you'll let us. Send ambassadors first to Sparta" (p. 204), she says, to negotiate a peace. She adds poignancy to her argument by describing the difficulty women face, which men do not, when the men go off to war:

> Even if we've got husbands, we're war widows just the same. And never mind us—think of the unmarried ones, getting on in years and with never a hope—that's what really pains me . . . . A man comes home—he may be old and grey—but he can get himself a young wife in no time. But a woman's not in bloom for long, and if she doesn't succeed quickly, there's no one will marry her (p. 205).

The second act of Lysistrata occurs five days later, and Lysistrata herself is wavering: she wanders "restless to and fro" from "sex-starvation" (p. 210). Moreover, some women are rengeging on their vows. But Lysistrata resolves to continue and prevails over most women, even persuading a woman named Myrrhine, who sees her own unwashed and unfed baby, to rebut the pleas of her husband for companionship by saying, caustically, to him: "I pity him [the baby] all right. His father hasn't looked after him very well" (p. 217).

Ambassadors from Sparta, whose women have followed the Athenian lead and abstained from sex, meet the Athenian negotiators. The men on both sides are distraught, but they take places on either side of Lysistrata, who makes a triumphant entrance ("no need to summon her") "magnificently arrayed" (p. 226).

Lysistrata begins her soliloquy with the brilliantly disarming statement: "I am a woman, but I am not brainless" (p. 227). Laying guilt on both Athens and Sparta, she helps them reconcile their differences, evoking from one impatient negotiator exactly the connection she wanted to make: "Peace! Peace! Bed! Bed!" (p. 229). The play ends with rejoicing and dancing at a banquet: the women have won.

How did they do it? To begin with, "winning" is not really an accurate characterization of the outcome, because the men and women have strong common interests, as depicted in the game in Figure 2a (game 35 in Figure

---

Figure 2 about here

---

1). The women, led by Lysistrata, can either refrain (R) from sex or not refrain ($\bar{R}$), and the men can continue to fight (F) or not fight ($\bar{F}$), giving rise to the four outcomes. Starting in the upper left cell of the matrix, and moving clockwise, I rank them from best (4) to worst (1) for (women, men) as follows:

RF—Frustration: (1,2). The women suffer greatly (1), because their strategy of abstention fails to stop the fighting and, in addition, they are deprived of sex, which makes the men unhappy (2), too.

R$\bar{F}$—Partial success for women: (3,1). Abstention results in an end to fighting, which is good for the women (3)—but not as good as if they had continued sex—whereas the men are extremely upset (1) because, despite their not fighting, sex is withheld.

$\bar{R}$F—Success for men: (2,4). The women are unhappy (2), even though they have sex, because the fighting continues, whereas the men not only enjoy sex but are also able to fight without repercussions (4).

$\overline{RF}$—Resolution: (4,3). The women succeed completely (4), resuming their sex after the fighting ends, and the men benefit from sex but must desist from fighting (3).

While the benefits of fighting for the men may not be readily apparent today, the perspective I offer is that of the characters in the play.

The reason that game 35 in Figure 2a looks different from the game in Figure 1 is that I have interchanged the row and column players, and their strategies, in Figure 2a. Structurally, however, these games are the same, with the women having a dominant strategy of $\bar{R}$ and the men having a dominant strategy of F, which yields (2,4) in the Figure 2a game.

Frustrated at this outcome, and being able by themselves only to move to the breakdown (and worse) outcome of (1,2), the women rebel. They not only threaten the men with R but also carry out this threat, which leaves the men the choice of their two worst outcomes. Between (1,2) and (3,1), they would choose (1,2), except that the threat outcome of (4,3) that the women offer entices them. With apparently few regrets, they choose it.

What is most relevant, from the viewpoint of this paper, is the emotional head of steam Lysistrata worked up to get the men to capitulate. First, she realized that her threat had to be carried out to be real. Using an artful combination of logic and zeal, she rallied the women, and later the men, to her side. Furthermore, she expressed her heartfelt anger with clarity and force, as I have already noted, and often with ribaldry, sometimes referring explicitly to male and female sex organs.

What invests this comedy with a sharp edge is the serious nature of the choices the characters face. Both the men and the women are torn by conflicting feelings and do not make their choices frivolously, justifying a game-theoretic view of their situation. At the same time, the characters do not confine themselves purely to an intellectual plane, which gives the women's threats, especially, credibility. Indeed, the characters are variously infused with feelings of sadness, despair, frustration, anger, and betrayal (e.g., when some women desert Lysistrata). In the end, though, even the men seem pleased by the reconciliation between Athens and Sparta.

4. The Self-Frustration Game

It is one thing for F to experience a lack of control when it is A's dominant strategy that leads to her (i.e., F's) two worst outcomes. But what if it is F, not A, who has the dominant strategy, and—anticipating F's choice of this strategy—A's best response is to choose the one of his nondominant strategies that leads to F's two worst outcomes?

Then F's grief at A's choice seems as much attributable to her as to A: it is F's preferences—and dominant strategy—that induce A to frustrate her, rather than A's preferences and dominant strategy. (Recall in the Frustration Game that it is A who always has a dominant strategy.) If F is at least partially to blame for her grief in this situation, is escape from frustrating herself any easier than escape from being frustrated by A?

In order to answer this question, I define a new generic game, which I call the Self-Frustration Game (see Figure 3). In this game, the row player,

---

Figure 3 about here

Advantaged (A), is, as before, male. The column player, whom I call Self-Frustrated (SF), is assumed to be female. The Self-Frustration Game satisfies the following four conditions:

1. *Dominance for SF.* Without loss of generality, assume that t_1 is SF's dominant strategy, so

$$y_{11} > y_{12}; \quad y_{21} > y_{22}. \tag{3}$$

That is, whatever strategy A chooses (s_1 or s_2), SF prefers his payoffs associated with t_1 to those associated with t_2, making t_1 SF's unconditionally better strategy.

2. *Nash equilibrium.* Without loss of generality, assume that (x_{11}, y_{11}) is the unique Nash equilibrium that is associated with SF's dominant strategy of t_1, so

$$x_{11} > x_{21}. \tag{4}$$

That is, A prefers her payoff at (x_{11}, y_{11}) to that which she would receive at (x_{21}, y_{21}) if she switched from s_1 to s_2. Similarly, F prefers (x_{11}, y_{11}) to (x_{12}, y_{12})—and so would not switch from t_1 to t_2—because of the dominance of t_1, assumed in inequalities (3).

3. *Nondominance for A.* Given (4), to prevent s_1 from being dominant requires that

$$x_{22} > x_{12}. \tag{5}$$

4. *Lack of control.* F's two worst outcomes (1 and 2) are associated with A's nondominant strategy of s_1. From the first of inequalities (3), it follows that

$$y_{11} = 2; \quad y_{12} = 1.$$

To ensure the dominance of s_1, the second of inequalities (3) requires that

$$y_{21} = 4; \quad y_{22} = 3,$$

which gives a complete ordering of payoffs for SF of

$$y_{21} > y_{22} > y_{11} > y_{12}.$$

On the other hand, the fact that (x_{11}, y_{11}) is a Nash equilibrium but A does not have a dominant strategy associated with it—as given by inequalities (4) and (5)—defines a partial ordering of payoffs for A that admits six different complete orderings:

$$x_{11} > x_{21} > x_{22} > x_{12}; \qquad x_{22} > x_{11} > x_{12} > x_{21};$$
$$x_{11} > x_{22} > x_{12} > x_{22}; \qquad x_{22} > x_{11} > x_{21} > x_{12};$$
$$x_{11} > x_{22} > x_{21} > x_{12}; \qquad x_{22} > x_{12} > x_{11} > x_{21}.$$

The one ordering of SF, and the six orderings of A, define a total of six different strict ordinal games subsumed by the generic Self-Frustration Game, which are divided into two classes in Figure 3.[13]

I summarize properties of these games with two propositions, with the second proposition distinguishing between the two classes:

Proposition 4. *In the six specific games subsumed by the generic Self-Frustration Game, A's ranking of his payoff at the Nash equilibrium is better than SF's ranking in five games and the same as SF's in one game (game 28).*

Of the five games in which A's ranking is better, three are games in which A obtains his best payoff (4), and two are games in which he obtains his next-best payoff (3). In game 28, which is also a Frustration Game (when the players are interchanged), both players obtain their next-worst payoffs (2).

Like F in the Frustration Game, SF in the Self-Frustration Game always obtains her next-worst payoff (2) at the unique Nash equilibrium in the six games, whereas A does better than SF—at least in terms of comparative rankings—in 83% of the games. But unlike F, SF can escape her frustration in all six specific Self-Frustration Games, provided she has threat power (which, it will be recalled, offered an escape in only four of the 12 specific Frustration Games):

Proposition 5. *The exercise of either compellent or deterrent threat power by SF always results in a better outcome [i.e., either (3,3) or (3,4)] for her than that which she obtains at the Nash equilibrium. In the class I*

---

[13]Ignore for now the bracketed outcomes, shown below the outcomes in parentheses, in the so-called anticipation game.

games, this outcome is, in fact, better for both players; it can also be induced by A's exercise of deterrent threat power.

But what if the players do not have threat power? And what if the exercise of threat power leads to different outcomes—as it does in the three class II games—depending on which player (if either) possesses it?

Before discussing the class II games in which threat power is effective, I offer a dynamic perspective on the possible play of *all* 2 x 2 games, based on the "theory of moves" (TOM). As I will show, this theory offers SF, in particular, the opportunity to break out of the Nash equilibrium in the class I and II Self-Frustration games—without relying on threat power—given that the players are nonmyopic in a sense to be described.

5. The Theory of Moves (TOM)

The starting point of TOM is a payoff matrix, or *configuration*, in which the order of play is not specified. In fact, players are assumed not even to choose strategies but instead to move and countermove from outcomes, or states, by looking ahead and using "backward induction" to determine the rationality of both their moves and those of an opponent.

Because game theory assumes that players choose strategies simultaneously in games in normal or strategic form,[14] it does not raise questions about the rationality of moving or departing from outcomes—at

---

[14]Strategies may allow for sequential choices, but game theory models, in general, do not make endogenous who moves first, as TOM does, but instead specify a fixed order of play (i.e., players make either simultaneous or sequential choices). There are some recent exceptions, however, including Hamilton and Slutsky (1990, 1993), Rosenthal (1991), van Damme and Hurkens (1993), and Amir (1995). Typically, these models allow a player in the preplay phase of a game to choose when he or she will move in the play of the game. Yet the choice of when to move applies only to a player's initial strategy choice, whereas the nonmyopic calculations to be developed here assume that players, starting at states, make moves and countermoves that depend on thinking several steps ahead.

least beyond an immediate departure, à la Nash. In fact, however, most real-life games do not start with simultaneous strategy choices but commence at outcomes. The question then becomes whether a player, by departing from an outcome, can do better not just in an immediate or myopic sense but, rather, in an extended or nonmyopic sense.

In the case of 2 x 2 games, TOM postulates four *rules of play*, which describe the possible choices of the players at different stages:

1. Play starts at an outcome, called the *initial state,* which is at the intersection of the row and column of a 2 x 2 payoff matrix.
2. Either player can unilaterally switch his or her strategy, and thereby change the initial state into a new state, in the same row or column as the initial state.[15] The player who switches, who may be either row (R) or column (C), is called player 1 (P1).
3. Player 2 (P2) can respond by unilaterally switching his or her strategy, thereby moving the game to a new state.
4. The alternating responses continue until the player (P1 or P2) whose turn it is to move next chooses not to switch his or her strategy. When this happens, the game terminates in a *final state,* which is the *outcome* of the game.

[15] I do not use "strategy" in the usual sense to mean a complete plan of responses by the players to all possible contingencies allowed by rules 2-4, because this would make the normal form unduly complicated to analyze. Rather, *strategies* refer to the choices made by players that define a state, and *moves and countermoves* to their subsequent strategy switches from an initial state to a final state in an extensive-form game, as allowed by rules 2-4. For another approach to combining the normal and extensive forms, see Mailath, Samuelson, and Swinkels (1993).

Note that the sequence of moves and countermoves is strictly alternating: first, say, R moves, then C moves, and so on, until one player stops, at which point the state reached is final and, therefore, the outcome of the game.[16]

The use of the word "state" is meant to convey the temporary nature of an outcome, before players decide to stop switching strategies. I assume that no payoffs accrue to players from being in a state unless it is the final state and, therefore, becomes the outcome (which could be the initial state if the players choose not to move from it). But here I draw attention to the emotions that are evoked before an outcome is reached.

Rule 1 differs radically from the corresponding rule of play in classical game theory, in which players simultaneously choose strategies in a matrix game that determines an outcome. Instead of starting with strategy choices, I assume that players are already in some state at the start of play and receive payoffs from this state if they stay. Based on these payoffs, they decide, individually, whether or not to change this state in order to try to do better.[17]

To be sure, some decisions are made collectively by players, in which case it would be reasonable to say that they choose strategies from scratch, either simultaneously or by coordinating their choices. But if, say, two countries are coordinating their choices, as when they agree to sign a treaty, the most important strategic question is what individualistic calculations led them to this point. The formality of jointly signing the treaty is the

---

[16]An emendation in the rules of TOM that allows for backtracking would be appropriate in games of incomplete information, wherein players make mistakes that they wish to rectify. For more on possible rule changes under TOM, see Brams (1994b).

[17]Alternatively, players may be thought of as choosing strategies initially, after which they perform a thought experiment of where moves will carry them once a state is selected. The concept of an "anticipation game," developed later, advances this idea, which might be considered dynamic thinking about the static play of a matrix game. Generally, however, I assume that "moves" describe actions, not just thoughts, though I readily admit the possibility of the thought interpretation.

culmination of their negotiations, which does not reveal the move-countermove process that preceded it. This is precisely what TOM is designed to uncover.

In summary, play of a game starts in a state, at which players accrue payoffs only if they remain in that state so that it becomes the outcome of the game. If they do not remain, they still know what payoffs they would have accrued had they stayed; hence, they can make a rational calculation of the advantages of staying versus moving. They move precisely because they calculate that they can do better by switching states, anticipating a better outcome if and when the move-countermove process finally comes to rest.

Rules 1–4 say nothing about what causes a game to end, but only when: termination occurs when a "player whose turn it is to move next chooses not to switch its strategy" (rule 4). But when is it rational not to continue moving, or not to move in the first place from the initial state?

To answer this question, I posit a rule of *rational termination* (first proposed in Brams, 1983, pp. 106-107), which has been called "inertia" by Kilgour and Zagare (1987, p. 94). It prohibits a player from moving from an initial state unless doing so leads to a better (not just the same) final state:

5. A player will not move from an initial state if this move
   (i) leads to a less preferred final state (i.e., outcome); or
   (ii) returns play to the initial state (i.e., makes the initial state the outcome).

I will discuss in section 6 how rational players, starting from some initial state, determine by backward induction what the outcome will be.

Condition (i) of rule 5, which precludes moves that result in an inferior state, needs no defense. But condition (ii), which precludes moves to the same state because of cycling back to the initial state, is worth some elaboration. It says that if it is rational for play of the game to cycle back to the initial state after Pl moves, Pl will not move in the first place. After all, what is the point of initiating the move-countermove process if play simply returns to "square one," given that the players receive no payoffs along the way (i.e., before an outcome is reached)?

Not only is there no gain from cycling, but in fact there may be a loss because of so-called transaction costs—including the emotional energy spent—that players suffer by virtue of making moves that, ultimately, do not change the situation.[18] Therefore, it seems sensible to assume that P1 will not trigger a move-countermove process if it only returns the players to the initial state, making it the outcome.

I call rule 5 a *rationality rule,* because it provides the basis for players to determine whether they can do better by moving from a state or remaining in it. Still another rationality rule is needed to ensure that both players take into account each other's calculations before deciding to move from the initial state. I call this rule the *two-sidedness rule*:

6. Given that players have complete information about each other's preferences and act according to the rules of TOM, each takes into account the consequences of the other player's rational choices, as as well as his or her own, in deciding whether to move from the initial state or subsequently, based on backward induction. If it is rational for one player to move and the other player not to move from the initial state, then the player who moves takes *precedence*:

---

[18]However, other rules of play of TOM allow for cycling; see Brams (1994b, ch. 4; 1995a). In future work, I plan to explore why players might invest emotional energy to prolong conflict through cycling.

his or her move overrides the player who stays, so the outcome is that induced by the player who moves.

Because players have complete information, they can look ahead and anticipate the consequences of their moves. I next show how, using backward induction, they do this in the context of an example illustrating the emotions evoked when players make moves, first to inferior states, in order to get to still better states. I will also mention some theoretical tie-ins of the Self-Frustration Game to other generic games.

6. From Self-Frustration to Murder in *Macbeth*[19]

A central feature of Shakespeare's great tragedy, *Macbeth*, written in 1606, is the conflict between Lady Macbeth and her husband, Macbeth, over murdering King Duncan. Duncan's demise would facilitate their ascent to the throne as king and queen of Scotland, but there are risks.

Lady Macbeth's ambition is fed by a letter she receives from Macbeth prophesying his greatness, based on his meeting with three "Weird Sisters" (who are considered to be witches). After mentioning in his letter that the Sisters saluted him with "Hail King that shalt be," he writes his wife—"my dearest partner of greatness"—of "what greatness is promised thee" (I, v, 10-12).

But Lady Macbeth worries—with good reason, it turns out—that her husband's "nature . . . is too full o'th' milk of human kindness/To catch the nearest way" (I, iv, 15-17). Consequently, she wishes

That I may pour my spirits in thine ear,

---

[19]This section is adapted, with significant changes, from a student paper of a former NYU undergraduate, Nancy Wang.

And chastise with the valour of my tongue
All that impedes thee from the golden round (I, iv, 25-27).

The advice she gives to Macbeth that will advance him to the "golden round" (i.e., crown) is preceded by her own musings.

Upon receiving news that King Duncan will be visiting Macbeth at his castle that very evening, Lady Macbeth thinks about his "fatal entrance . . . under my battlements" (I, v, 38-39). She immediately steels herself psychologically for his murder:

Come, you spirits
That tend on mortal thoughts, unsex me here,
And fill me from the crown to the toe, top-full
Of direst cruelty. Make think my blood,
Stop up the'access and passage to remorse,
That no compunctious visitings of nature
Shake my fell purpose (I, v, 40-45).

As if intending to murder Duncan herself, she says

Come, thick night,
And pall thee in the dunnest smoke of Hell,
That my keen knife see not the wound it makes (I, v, 49-51).

But, in fact, it is not Lady Macbeth—shed of the womanliness in herself that she despises—who wants to carry out the dastardly deed. Instead, she earnestly hopes that her husband, though he looks "like th'innocent flower," is "the serpent under't" (I, v, 63-64). Macbeth, however, has grave doubts, especially that the murder will "return/To plague th'inventor" (I, vii, 9-10), and that King Duncan's "virtues/Will plead like angels" (I, vii, 18-19) if he is assassinated.

While Macbeth confesses to "vaulting ambition, which o'erleaps itself/And falls on the' other" (I, vii, 27-28), assassinating King Duncan to satisfy this ambition is another matter. Indeed, Lady Macbeth is furious when he tells her, while feasting with Duncan at their castle, that "we will proceed no further in this business/He hath honoured me of late" (I, vii, 31-32). She immediately accuses her husband of cowardice, even unmanliness, saying that *she* would, as a mother, have "dashed the brains out [of her baby], had I so sworn/As you have done to this [sworn to murder King Duncan] (I, vii, 57-58).

Still wavering, Macbeth asks, what "if we should fail?" (I, vii, 59). Lady Macbeth brushes him off:

> We fail?
> But screw your courage to the sticking place,
> And we'll not fail (I, vii, 60-62).

Lady Macbeth then outlines how, when Duncan sleeps, the assassination will be carried out. Macbeth finally agrees, but he is haunted by fear—an emotion that even Lady Macbeth's ferocity cannot dispel—of this "terrible feat" (I, vii, 80).

The game played between Lady Macbeth and Macbeth involves her choosing to incite him to murder (I) or not inciting him ($\bar{I}$), and Macbeth's killing Duncan (K) or not killing him ($\bar{K}$). Starting in the upper left cell of the matrix in Figure 2b, and moving clockwise, I rank the states from best (4) to worst (1) for (Lady Macbeth, Macbeth) as follows:

IK—Murder motivated: (3,3). Lady Macbeth is pleased that Macbeth carries out the murder—even if she must strenuously prod him to do so (3)—and Macbeth is pleased to satisfy her desire and demonstrate his manliness (3).

$I\overline{K}$—Extreme frustration: (1,1). To the great chagrin of both Lady Macbeth (1) and Macbeth (1), her pleas are ignored and his courage is thrown into question.

$\overline{I}K$—Murder unmotivated: (4,2). Lady Macbeth is most pleased if Macbeth murders King Duncan without her having to prod him (4), but Macbeth is plagued by guilt and self-doubt when not incited by Lady Macbeth (2).

$\overline{I}\overline{K}$—Status quo: (2,4). Lady Macbeth is angry at Macbeth and disgusted with herself when neither she nor Macbeth acts to change the status quo (2), whereas Macbeth is happy to be honored by King Duncan and not have to act treacherously against him, especially as his host (4).

Structurally, this is game 56 in Figure 3, except that the row and column players, and their strategies, are interchanged in its Figure 2b representation.

Play commences at the status quo of (2,4), which is upsetting to Lady Macbeth once she has read the letter from her husband: she realizes that the throne is within their grasp, but her husband may fail her in the clutch. Because her dominant strategy of $\overline{I}$ is associated with this state, however, standard game theory predicts that she will not move from it; neither will Macbeth, because it is a Nash equilibrium. But TOM makes a different prediction.

7. Nonmyopic Equilibria (NMEs) in *Macbeth*

From the perspective of TOM, Lady Macbeth would calculate the rational consequences of moving from (2,4)—what countermove, on the part of Macbeth, her move from this state would trigger, what counter-countermove she would make, and so on.

Given that players have complete information about each other's preferences, I assume they base their calculations on *backward induction*, which I will next illustrate for game 56 in Figure 2b. Starting from (2,4) and cycling back to this state, I will show where R (Lady Macbeth) and C (Macbeth) will terminate play:[20]

If R moves first, the counterclockwise progression from (2,4) back to (2,4)—with the player (R or C) who makes the next move shown below each state in the alternating sequence—is as follows (see Figure 2b):

	State 1		State 2		State 3		State 4		State 1
	R		C		R		C		
R starts:	(2,4)	→	(1,1)	→	<u>(3,3)</u>	→|	(4,2)	→	(2,4)
Survivor:	(3,3)		(3,3)		(3,3)		(2,4)		

The *survivor* is determined by working backward, after a putative cycle has been completed. Assume that the players' alternating moves have taken

[20]Where, of course, depends on the endstate, or *anchor,* from which the backward induction proceeds, which I assume here—for reasons given in section 6—is after one complete cycle. This assumption defines a finite extensive-form game, but it is dropped in other parts of TOM, where alternative rationality rules are applied to "cyclic games," which may cycle until the player without "moving power" quits (Brams, 1994b, ch. 4). In only two of the six cyclic games (28 and 35) subsumed by the Frustration Game does moving power provide relief from the Nash equilibrium for F, and in these games a deterrent threat works as well in game 28 and better in game 35. In the six specific games subsumed by the Self-Frustration Game, the NME (to be defined in the text) from the Nash equilibrium provides SF with at least as much relief as does SF's moving power, except in game 48, in which SF can induce (2,4), rather than the NME of (4,3), with moving power. For more details on the effects of moving power in these six games, see Brams (1994b, 1995a).

them counterclockwise from (2,4) to (1,1) to (3,3) to (4,2), at which point C must decide whether to stop at (4,2) or complete the cycle and return to (2,4). Clearly, C prefers (2,4) to (4,2), so (2,4) is listed as the survivor below (4,2): because C *would* move the process back to (2,4) should it reach (4,2), the players know that if the move-countermove process reaches this state, the outcome will be (2,4).

Knowing this, would R at the prior state, (3,3), move to (4,2)? Because R prefers (3,3) to the survivor at (4,2)—namely, (2,4)—the answer is no. Hence, (3,3) becomes the survivor when R must choose between stopping at (3,3) and moving to (4,2)—which, as I have just shown, would become (2,4) once (4,2) is reached.

At the prior state, (1,1), C would prefer moving to (3,3) than stopping at (1,1), so (3,3) again is the survivor if the process reaches (1,1). Similarly, at the initial state, (2,4), because R prefers the previous survivor, (3,3), to (2,4), (3,3) is the survivor at this state as well.

The fact that (3,3) is the survivor at initial state (2,4) means that it is rational for R initially to move to (1,1), and C subsequently to move to (3,3), where the process will stop, making (3,3) the rational choice if R has the opportunity to move first from initial state (2,4). That is, after working *backward* from C's choice of completing the cycle or not at (4,2), the players can reverse the process and, looking *forward,* determine that it is rational for R to move from (2,4) to (1,1), and C to move from (1,1) to (3,3), at which point R will stop the move-countermove process at (3,3).

Notice that R does better at (3,3) than at (2,4), where it could have terminated play at the outset, and C does better at (3,3) than at (1,1), where it could have terminated play, given that R is the first to move. I indicate that (3,3) is the consequence of backward induction by underscoring this state in the progression; it is the state at which *stoppage* of the process occurs. In addition, I indicate that it is not rational for R to move on from (3,3) by the vertical line blocking the arrow emanating from (3,3), which I refer to as *blockage:* a player will always stop at a blocked state, wherever it is in the progression. Stoppage occurs when blockage occurs for the *first* time from some initial state, as I illustrate next.

If C can move first from (2,4), backward induction shows that (2,4) is the last survivor, so (2,4) is underscored when C starts. Consequently, C would *not* move from the initial state, where there is blockage (and stoppage), which is hardly surprising since C receives its best payoff in this state:[21]

<table>
<tr><td></td><td>State 1</td><td></td><td>State 2</td><td></td><td>State 3</td><td></td><td>State 4</td><td></td><td>State 1</td></tr>
<tr><td></td><td>C</td><td></td><td>R</td><td></td><td>C</td><td></td><td>R</td><td></td><td></td></tr>
<tr><td>C starts:</td><td>(2,4)</td><td>→|</td><td>(4,2)</td><td>→|</td><td>(3,3)</td><td>→</td><td>(1,1)</td><td>→</td><td>(2,4)</td></tr>
<tr><td>Survivor:</td><td>(2,4)</td><td></td><td>(4,2)</td><td></td><td>(2,4)</td><td></td><td>(2,4)</td><td></td><td></td></tr>
</table>

As when R has the first move, (2,4) is the first survivor, working backward from the end of the progression, and is also preferred by C at (3,3). But then, because R at (4,2) prefers this state to (2,4), (2,4) is temporarily displaced as the survivor. It returns as the last survivor, however, because C at (2,4) prefers it to (4,2).

Thus, the first blockage and, therefore, stoppage occurs at (2,4), but blockage occurs subsequently at (4,2) if, for any reason, stoppage does not terminate moves at the start. In other words, if C moved initially, R would then be blocked. Hence, blockage occurs at two states when C starts the

---

[21]But it is rational in several 2 x 2 games for a player to depart from his or her best state (4), because in these games he or she would do worse if the other player departed first (Brams, 1994b, ch. 3).

move-countermove process, whereas it occurs only once when R has the first move.

The fact that the rational choice depends on which player has the first move—(3,3) is rational if R starts, (2,4) if C starts—leads to a conflict over what outcome will be selected when the process starts at (2,4). However, because it is not rational for C to move from the initial state, R's move takes precedence, according to rule 6, and overrides C's decision to stay. Consequently, when the initial state is (2,4), the result will be (3,3), which is indicated for game 56 in Figure 3, and in Figure 2b, by placing [3,3]—in brackets—below (2,4).

The outcome into which a state goes is called the *nonmyopic equilibrium* (NME) from that state. NMEs may be viewed as the consequence of both players' looking ahead and making rational calculations of where the move-countermove process will transport them, based on the rules of TOM, from each of the four possible initial states.

To take another example, backward-induction analysis from each state in game 27 shows that each state will go into (4,3). Thus, wherever play starts, the players can anticipate that they will end up at (4,3), making it the unique NME in game 27. This is also true of (4,3) in game 28, but not in game 48, the third class I game. Starting at (2,4) in game 48, the players will not depart from this state, making (2,4) as well as (4,3) an NME in this game.

Like game 48, all games in class II contain at least two NMEs. But some of these NMEs are *indeterminate*, because there is a conflict over who will move first. In game 50, for example, if (1,1) is the initial state, [3,4]/[4,2] indicates that when R moves first from (1,1), (3,4) will be outcome, whereas when C moves first, (4,2) will be the outcome.[22]

Because R prefers (4,2) whereas C prefers (3,4), each player will try to hold out longer in order to induce the other player to move first. Who wins in this struggle will depend on which player has "order power"—that is, who can determine the order of moves, starting at (1,1) (Brams, 1994, ch. 5).

Every 2 x 2 game contains at least one NME, because from each initial state there is an outcome (perhaps indeterminate) of the move-countermove process. If this outcome is both determinate and the same from every initial state, then it is the only NME; otherwise, there is more than one NME.

In game 56 in Figures 2a and 3, there are three different NMEs, which is the maximum number that can occur in a 2 x 2 strict ordinal game; the minimum, as already noted, is one. All except two of the 78 2 x 2 games (game 56 and Chicken, which is not shown) have either one or two NMEs.

The four bracketed states of each game in Figure 3 define what I call the *anticipation matrix*, with each state in this matrix an NME. Insofar as players choose strategies as if they were playing a game based on this matrix, one can determine which NMEs are Nash equilibria in the *anticipation game* and therefore likely to be chosen.[23]

---

[22]Actually, the result of backward induction by R from (1,1) in game 50 is (2,3) rather than (3,4). But, as I argue in Brams (1994b, p. 114, ftn. 20), the players would have a common interest in implementing the Pareto-superior (3,4) to (2,3) when there is clockwise movement from (1,1). However, the implementation of (3,4) would require a binding commitment on the part of R not to move on from (3,4) to (4,2), which is not assumed possible in noncooperative game theory. I conclude (with an interchange of players and their payoffs): "I do not see an airtight case being made for either (2,3) or (3,4) as *the* NME from (1,1) when R moves first, which nicely illustrates the nuances that TOM surfaces that the rules of standard game theory keep well submerged." Incidentally, game 50 is the only game of the 78 2 x 2 strict ordinal games in which this kind of ambiguity about NMEs arises.

[23]Because the NMEs in games 27 and 28 are all the same, the strategies of the players in their anticipation games are indistinguishable, making all four states Nash equilibria. In

To summarize, where players start in the original games in Figure 3—including the unique dominant-strategy Nash equilibria—may not be where they end up, according to TOM. Thus, an original game may mask a good deal of instability when the players can move and countermove from states.[24]

This instability is evident in *Macbeth*. Lady Macbeth relentlessly hounds Macbeth to murder King Duncan, which brings the players temporarily to (1,1) when Macbeth temporizes. Frustrated, Lady Macbeth explodes with anger, which impels Macbeth finally to act, bringing the players to (3,3) after the murder is committed.[25]

The couple is able, for a while, to escape suspicion, placing the blame on the servants, who are also murdered, and King Duncan's sons, who flee. When Banquo, who had with Macbeth heard the prophecy of the Weird Sisters, seems likely to uncover the Macbeths' plot, Macbeth has him killed. In the end, however, the previously indomitable Lady Macbeth comes apart emotionally and commits suicide. Without her formidable presence, Macbeth loses faith and is slain by Macduff, whose entire family he had had murdered.

---

the four other games, t_1 is C's (weakly) dominant strategy; R's best response leads to a dominant-strategy Nash equilibrium in each anticipation game, some of which contain other Nash equilibria. Only in game 49 is the Nash equilibrium in the anticipation game, [4,2], also the Nash equilibrium, (4,2), in the original game.

[24]Even the choice of a state associated with a dominant-strategy Nash equilibrium in the anticipation games offers no assurance that players will stay at this state. Indeed, except for one of the two states associated with the [3,4] Nash equilibria in the anticipation game of game 50, the players will move from *every* such state to some different NME in the six anticipation games in Figure 3.

[25]By drugging the servants who were supposed to guard Duncan and making sure the door to his chamber was unlocked, Lady Macbeth was certainly an accessory to murder. But this legalistic interpretation glosses over her preeminent role, which she maintained by an implacable display of cold fury and hot anger. In fact, the NME of (3,3) is reinforced by Lady Macbeth's compellent threat of choosing strategy I and sticking with it, but I prefer the NME explanation for (3,3), because it relies only on the farsighted thinking, not any special powers, of the players.

Remorseless as Lady Macbeth is until the end, she is a character, in my opinion, brimming with emotion. She uses her frustration to great advantage to push her reluctant and vacillating husband over the brink, so to speak, which Delilah did to Samson, on a different matter, in the famous Bible story (also modeled as game 56 in Brams, 1980, 1994b).

Characters in real-life self-frustration games have displayed similar emotions, which often caught their opponents off guard (e.g., the United States by the 1941 Japanese attack on Pearl Harbor; Israel by Egyptian President Anwar Sadat's 1977 peace initiative and visit to Jerusalem). Their moves engendered great shock and surprise (Brams, 1995b), which is an emotional reaction in its own right triggered by the surpriser's (Japan's and Israel's) unanticipated actions.

TOM, by explicating the rationale of such moves, makes them less surprising. It turns out that the six specific games subsumed by the Self-Frustration Game in Figure 3 are precisely the games subsumed by a so-called generic Surprise Game (Brams, 1994b) and a generic Freedom Game (Brams, 1995a), in which the player with the dominated strategy (SF in Figure 3) finds it rational, according to TOM, to switch to his dominated strategy, in turn inducing A to switch her strategy, too.

This move and countermove from the unique Nash equilibrium (upper-left states in games in Figure 3) brings both players to the lower-right states. From the latter states, the players would return to the Nash equilibrium (if given the chance). Unfortunately for Macbeth and Lady Macbeth, they met their end before this could happen.

8. Conclusions and Future Research

Frustration is expressed, I have argued, when players who are dissatisfied and lack control are forced to make "desperate" choices—choices that hurt them, at least temporarily, because they involve moves from the Nash equilibrium. But they are not unmotivated choices. By erupting with an emotion like anger or rage when they move to their disadvantage, at least initially, these players signal that they want this loss of control to end—in particular, for the other player to respond by helping them.

I have suggested that the theory of moves (TOM), which offers a rationale for players' making dynamic choices in games, can explain these choices. In the case of the Frustration and Self-Frustration Games, players express frustration when they threaten the other player, or try to induce a nonmyopic equilibrium (NME), that hurts the frustrated or self-frustrated player initially.

I showed that in the Frustration Game, Frustrated (F) can break out of the unique Nash equilibrium, at which she suffers her next-worst outcome, with a deterrent threat in four of the 12 specific games. But this threat is effective only if F can endure the breakdown outcome better than Advantaged (A), who has a dominant strategy. The women (F) in *Lysistrata* had such threat power in the form of withholding sex from the men (A), which induced the men to stop fighting, making the women ecstatic and the men at least reasonably happy.

The Self-Frustration Game leaves the player with the dominant strategy, Self-Frustrated (SF), dissatisfied at the unique Nash equilibrium. In this game, I assumed that each player considers the consequences of moving from that state, the other player's countermoving, and so on, which eventually brings both players to what I call a nonmyopic equilibrium (NME). This is the outcome that rational players would be expected to migrate to—and from which, I assume, they derive all their payoffs—if it is rational for them to move at all (based on backward induction).

It turns out that in the six specific games subsumed by the Self-Frustration Game, SF can, starting at the Nash equilibrium, induce an NME better for herself, and sometimes for A, than the Nash equilibrium. Threat power of either the compellent or deterrent kind is also effective in these games for one or both players. In *Macbeth*, however, it seemed less that Lady Macbeth could threaten Macbeth with anything—except, perhaps, withdrawal of her love—than switch to her dominated strategy to induce Macbeth in turn to switch his strategy.

This move and countermove led the couple to plot the murder of King Duncan and successfully execute it. The fact that their crime later unraveled, however, does not undermine the rationality of their calculations in the beginning, undergirded, as it was, by Lady Macbeth's logic and the Weird Sisters' prophecy of greatness.

In the past, I have used TOM to interpret the actions, but not the emotional responses, of the players along the path to NMEs. It is *during* the moves along this path, I suggest, when human emotions come into play, even if (long-term) payoffs to the players accrue only when an NME is reached.

I believe other emotions besides frustration can be fruitfully studied using TOM. Thus, for example, an Envy Game might be one in which the good fortune of one player is tied to the bad fortune of the other, suggestive of pure-conflict games or something close to them.

For the game theorist, unlike the psychologist, it will be the structure of situations that trigger emotions, rather than the emotions themselves, that will be of primary interest. No generic game that mirrors such structures, however, will capture all the subtleties of emotional interactions that people have in literary, as well as real-life, games. But one psychologist has argued that "a necessary first step is to develop a theoretical system that can account successfully for a reasonably large number of 'emotional' phenomena" (Mandler, 1994, p. 243).[26]

A dynamic game-theoretic model offers, in my opinion, a promising start on understanding when frustration and other emotions are likely to arise, and how best to cope with them. Nonetheless, the classification of the different possible paths, the assessment of their payoff consequences, and a psychological interpretation of the emotions these different paths are likely to evoke will require considerable investigation and analysis, even for 2 x 2 games.

To further this endeavor, I propose as a working hypothesis that the more players suffer in non-preferred states along the path to a nonmyopic equilibrium, the more intense will be their emotions. Insofar as the nonmyopic equilibrium offers "relief" to the players—in the sense of providing them with a preferred state—this emotion will, in the end, be positive.

But if one player remains dissatisfied at this equilibrium state, he or she may want to depart, rendering this equilibrium unstable. Thereby one might analyze the *stability* of nonmyopic equilibria—a major refinement of

---

[26]Mor (1993) indicates other linkages of the psychology literature to the game-theoretic modeling literature, including that based on TOM, especially in international relations.

the current concept—showing how it depends on the nature of the path taken to reach it and the emotions evoked along this path.

A subsidiary hypothesis is the following: The more difficult the path to a nonmyopic equilibrium, the more likely this equilibrium will remain stable. "Difficulty" might be measured by the number of Pareto-inferior states that a 2 x 2 game contains, especially if it is a "cyclic" game (36 of the 78 2 x 2 games) in which both players have a continuing incentive to move.

According to this hypothesis, a resolution of a crisis in cyclic games that contain two Pareto-inferior states (22 games) is more likely to hold than in cyclic games that contain only one such state (12 games) or no such state (two games, both of which are games of pure conflict). The underlying reason for this greater stability is that the players in games with more Pareto-inferior states suffer together when an equilibrium is upset.

As a case in point, the relative stability of the resolution achieved in the 1962 Cuban missile crisis may stem not only from the fact that a mutually disastrous nuclear war was averted but also—from an emotions viewpoint—that future superpower leaders had no desire to repeat such a traumatic experience.[27] In this sense, the high drama of agreements forged in crises may be an antidote to their later breakdown, compared with agreements reached in noncrisis situations in which there is no possibility of mutual disaster.

But, of course, this is an untested hypothesis. It and related hypotheses can, I believe, successfully be investigated within the dynamic framework developed here.

---

[27]Nikita Khrushchev's sense of loss of control in this crisis is reflected in the following statement he made in a letter written to John Kennedy: "If people do not show wisdom, then in the final analysis they will come to clash, like blind moles" (Divine, 1971, p. 47).

References

American Heritage Dictionary, New College Division (1980). Boston: Houghton Mifflin.

Amir, Rabah (1995). "Endogenous Timing in Two-Player Games: A Counterexample. *Games and Economic Behavior* 9, no. 2 (May): 234-237.

Aristophanes (411 B.C.E.). *Lysistrata*, tr. by Allan H. Sommerstein. London: Penguin, 1973.

Aumann, Robert J., and Mordecai Kurz (1977). "Power and Taxes." *Econometrica* 45, no. 5 (July): 522-539.

Berne, Eric (1964). *Games People Play*. New York: Ballantine.

Brams, Steven J. (1980). *Biblical Games: A Strategic Analysis of Stories in the Old Testament*. Cambridge, MA: MIT Press.

Brams, Steven J. (1983). *Superior Beings: If They Exist, How Would We Know? Game-Theoretic Implications of Omniscience, Omnipotence, Immortality, and Incomprehensibility*. New York: Springer-Verlag.

Brams, Steven J. (1993). "Theory of Moves." *American Scientist* 81, no. 6 (December): 562-570.

Brams, Steven J. (1994a). "Game Theory and Literature." *Games and Economic Behavior* 6, no. 1 (January): 32-54.

Brams, Steven J. (1994b). *Theory of Moves*. Cambridge, UK: Cambridge University Press.

Brams, Steven J. (1995a). "Modeling Free Choice in Games." Preprint, Department of Politics, New York University.

Brams, Steven J. (1995b). "The Rationality of Surprise: Unstable Nash Equilibria and the Theory of Moves." Preprint, Department of Politics, New York University.

Brams, Steven J., and Marek P. Hessel (1984). "Threat Power in Sequential Games." *International Studies Quarterly* 28, no. 1 (March): 23-44.

Brams, Steven J., and Walter Mattli (1993). "Theory of Moves: Overview and Examples." *Conflict Management and Peace Science* 12, no. 2 (Spring): 1-39.

Damasio, Antonio R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.

de Sousa, Ronald (1987). *The Rationality of Emotion*. Cambridge, MA: MIT Press.

Divine, Robert A. (ed.) (1971). *The Cuban Missile Crisis*. Chicago: Quadrangle.

Elster, Jon (1994). "Rationality, Emotions, and Social Norms." *Synthese* 98, no. 1 (January): 21-49.

"Emotions and Rational Choice" (1993). *Rationality and Society* (special issue) 5, no. 2 (April).

Frank, Robert H. (1988). *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton.

Frijda, Nico H. (1986). *The Emotions*. Cambridge, UK: Cambridge University Press.

Geanakopolos, John, David Pearce, and Ennio Stacchetti (1989). "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1, no. 1 (March): 60-79.

Gerlernter, David (1994). *The Muse and the Machine: Computerizing the Poetry of Human Thought*. New York: Free Press.

Gilboa, Itzhak, and David Schmeidler (1988). "Information Dependent Games: Can Common Sense be Common Knowledge?" *Economics Letters* 27: 215-221.

Hamilton, Jonathan H., and Steven M. Slutsky (1990). "Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria." *Games and Economic Behavior* 2, no. 1 (March): 29-46.

Hamilton, Jonathan H., and Steven M. Slutsky (1993). "Endogenizing the Order of Moves in Matrix Games." *Theory and Decision* 34, no. 1 (January): 47-62.

Hirshleifer, Jack (1987). "On the Emotions as Guarantors of Threats and Promises." In John Dupré (ed.), *The Latest on the Best: Essays on Evolution and Optimality*. Cambridge, MA: MIT Press, pp. 305-326.

Hirshleifer, Jack (1994). "The Dark Side of the Force." *Economic Inquiry* 32 (January): 1-10.

Howard, Nigel (1994). "Drama Theory and Its Relation to Game Theory. Part 1: Dramatic Resolution vs. Rational Solution; Part 2: Formal Model of the Resolution Process." *Group Decision and Negotiation* 3, no. 2 (June): 187-206; 207-235.

Howard, Nigel *et al.* (1993). "Manifesto for a Theory of Drama and Irrational Choice." *Systems Practice* 6, no. 4: 429-434.

Kilgour, D. Marc, and Frank C. Zagare (1987). "Holding Power in Sequential Games." *International Interactions* 13, no. 2: 91-114.

Kroll, John A. (1993). "The Complexity of Interdependence." *International Studies Quarterly* 37, no. 3 (September): 321-347.

Lazarus, Richard S. and Bernice N. (1994). *Passion and Reason: Making Sense of Our Emotions.* Oxford University Press.

Lewis, Michael (1995). "Self-Conscious Emotions." *American Scientist* 83, no. 1 (January-February): 68-78.

Mailath, George J., Larry Samuelson, and Jeroen Swinkels (1993). "Extensive Form Reasoning in Normal Form Games." *Econometrica* 61, no. 2 (March): 273-302.

Mandler, George (1994). "Emotions and the Psychology of Freedom." In Stephanie H. M. van Goozen, Nanne E. Van de Poll, and Joseph A. Sergeant (eds.), *Emotions: Essays on Emotion Theory.* Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 241-262.

Mor, Ben D. (1993). *Decision and Interaction in Crisis: A Model of International Crisis Behavior.* Westport, CT: Praeger.

Rabin, Matthew (1993). "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83, no. 5 (December): 1281-1302.

Rapoport, Anatol, and Melvin J. Guyer (1966). "A Taxonomy of 2 x 2 Games." *General Systems: Yearbook of the Society for General Systems Research* 11: 203-214.

Rosenthal, Robert W. (1991). "A Note on Robustness of Equilibria with Respect to Commitment Opportunities." *Games and Economic Behavior* 3, no. 2 (May): 237-243.

Schelling, Thomas C. (1966). *Arms and Influence*. New Haven, CT: Yale University Press.

Shakespeare, William (1606). *Macbeth*, ed. by Nicholas Brooke. Oxford, UK: Oxford University Press, 1994.

Solomon, Robert C. (1993). *The Passions: Emotions and the Meaning of Life*. Indianapolis, IN: Hackett.

Sommerstein, Alan M. (1973). "Introduction." Aristophanes, *Lysistrata*. London: Penguin, pp. 9-38.

Thibaut, John W., and John W. Kelley (1959). *The Social Psychology of Groups*. New York: John Wiley.

van Damme, E., and S. Hurkens (1993). "Commitment Robust Equilibria and Endogenous Timing." Discussion Paper No. 9356, Center for Economic Research, Tilburg University, The Netherlands (July).

FIGURE 1. FRUSTRATION GAME

Generic Game

Frustrated (F)

Advantaged (A)		t_1	t_2	
	s_1	$(x_{11}, y_{11}) = 2$	$(x_{12}, y_{12}) = 1$	← Dominant strategy for A ($x_{11} > x_{21}$ and $x_{12} > x_{22}$), containing the two worst outcomes for F.
	s_2	(x_{21}, y_{21})	(x_{22}, y_{22})	

12 Specific Games Subsumed by Generic Game

Class I (8 games): No relief for F from Nash equilibrium of (x_{11}, y_{11})

5 (17)

(4,2)	(3,1)
(2,4) ↑	(1,3)

6 (18)

(4,2)	(3,1)
(1,4) ↑	(2,3)

10 (10)

(3,2)	(4,1)
(2,4) ↑	(1,3)

11 (11)

(3,2)	(4,1)
(1,4) ↑	(2,3)

18 (35)

(4,2)	(3,1)
(2,3)	(1,4)

19 (36)

(4,2)	(3,1)
(1,3)	(2,4)

25 (45)

(3,2)	(4,1)
(2,3)	(1,4)

26 (46)

(3,2)	(4,1)
(1,3)	(2,4)

Class II (4 games): Relief through the exercise of deterrent threat power

22 (39)

(4,2)	(2,1)
(3,3)^t	(1,4)

28 (48)

(2,2)	(4,1)
(1,3)	(3,4)^{t*}

32 (12)

(2,2)	(4,1)
(1,4) ↑	(3,3)^{t*}

35 (21)

(4,2)	(2,1)
(3,4)^t ↑	(1,3)

Key: (x, y) = (payoff to A, payoff to F)
4 = best; 3 = next best; 2 = next worst; 1 = worst
Nash equilibrium underscored
t/*t** = deterrent threat outcome for F/both players
↑ = dominant strategy for F

FIGURE 2. A FRUSTRATION AND A SELF-FRUSTRATION GAME

2a. Frustration Game 35 (of *Lysistrata*)

Men

Women	Fight (F)	Don't fight ($\bar{F}$)	
Refrain (R)	Frustration (1,2)	Partial success for women (3,1)	← Deterrent threat
Don't refrain ($\bar{R}$)	Success for men (2,4)	Resolution $(4,3)^{td}$	← Dominant strategy
	↑ Dominant strategy		

2b. Self-Frustration Game 56 (of *Macbeth*)

Macbeth

Lady Macbeth	Kill (K)	Don't kill ($\bar{K}$)	
Incite (I)	Murder motivated $((3,3)^{tc})$ [2,4]	Extreme frustration (1,1) [3,3]/[2,4]	← Compellent threat
Don't incite ($\bar{I}$)	Murder unmotivated ((4,2)) [4,2]	Status quo ((<u>2,4</u>)^{tc}) [3,3]	← Dominant strategy
		↑ Compellent threat	

Key: (x, y) = (payoff to R, payoff to C)
[x, y] = [payoff to R, payoff to C] in anticipation game
4 = best; 3 = next best; 2 = next worst; 1 = worst
t = state induced by compellent (*c*) or deterrent (*d*) threat power
Nash equilibria underscored
Nonmyopic equilibria (NMEs) circled in original game 56

FIGURE 3. SELF-FRUSTRATION GAME

Generic Game

Self-Frustrated (SF)

		t_1	t_2	
Advantaged (A)	s_1	(x_{11}, y_{11}) = 2	(x_{12}, y_{12}) = 1	← Nondominant strategy for A ($x_{11} > x_{21}$ and $x_{22} > x_{21}$), containing the two worst outcomes for SF.
	s_2	(x_{21}, y_{21}) = 4	(x_{22}, y_{22}) = 3	
		↑ Dominant strategy for SF ($y_{11} > y_{12}$ and $y_{21} > y_{22}$)		

Six Specific Games Subsumed by Generic Game

Class I (3 games): (x_{22}, y_{22}) Pareto-superior to Nash equilibrium of (x_{11}, y_{11})

27 (47)

(3,2) [4,3]	(2,1) [4,3]
(1,4) [4,3]	((4,3)tc) [4,3]

28 (48)

(2,2) [4,3]	(3,1) [4,3]
(1,4) [4,3]	((4,3)tc) [4,3]

48 (57)

(3,2) [4,3]	(1,1) [4,3]
((2,4)) [2,4]	((4,3)tc) [4,3]

Class II (3 games): (x_{22}, y_{22}) not Pareto-superior to Nash equilibrium of (x_{11}, y_{11})

49 (44)

((4,2)tc) [3,3]	(2,1) [4,2]/(3,3)
(1,4) [4,2]	((3,3)tc) [4,2]

50 (45)

((4,2)tc) [3,4]	(1,1) [3,4]/[4,2]
((3,4)td) [3,4]	(2,3) [4,2]/[3,4]

56 (56)

((4,2)tc) [3,3]	(1,1) [3,3]/[4,2]
((2,4)) [2,4]	((3,3)tc) [4,2]

Key: (x, y) = (payoff to A, payoff to SF)
[x, y] = [payoff to A, payoff to SF] in anticipation game
4 = best; 3 = next best; 2 = next worst; 1 = worst
t = state induced by compellent (*c*) or deterrent (*d*) threat power of SF
Nash equilibria underscored (except in anticipation games 27 and 28)
Nonmyopic equilibria (NMEs) circled in original game