# Transforming Power Relationships: Leadership, Risk, and Hope

James H. Read
Ian Shapiro

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

# Transforming Power Relationships: Leadership, Risk, and Hope

**James H. Read & Ian Shapiro**

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

# Transforming Power Relationships: Leadership, Risk, and Hope

James H. Read & Ian Shapiro

May 2013

**Contact:**

James H. Read
Professor of Political Science
College of St. Benedict/ St. John's University
jread@csbsju.edu

Ian Shapiro,
Sterling Professor of Political Science
Yale University
ian.shapiro@yale.edu

---

# Abstract

Chronic communal conflicts resemble the prisoner's dilemma. Both communities prefer peace to war. But neither trusts the other, viewing the other's gain as its own loss, so potentially shared interests often go unrealized.

Achieving positive-sum outcomes from apparently zero-sum struggles requires a kind of risk-embracing leadership. To succeed leaders must: a) see power relations as potentially positive-sum; b) strengthen negotiating adversaries instead of weakening them; and c) demonstrate hope for a positive future and take great personal risks to achieve it.

Such leadership is exemplified by Nelson Mandela and F.W. de Klerk in the South African democratic transition. To illuminate the strategic dilemmas Mandela and de Klerk faced, we examine the work of Robert Axelrod, Thomas Schelling, and Josep Colomer, who highlight important dimensions of the problem but underplay the role of risk-embracing leadership. Finally we discuss leadership successes and failures in the Northern Ireland settlement and the Israeli-Palestinian conflict.

## Keywords

**General note on content**

The opinions expressed in this paper are those of the author and not necessarily those of the IHS.

# Contents

# I.    Introduction

Political communities marred by longstanding, bitter conflict find it extremely difficult to end that conflict, even if most members of the community suffer its effects and wish it could be resolved. The conflicts are often chronically self-reinforcing: violent eruptions reproduce incentives for parties, leaders, and ordinary individuals to act in ways that perpetuate the dynamic. Such conflicts were modelled in Thomas Hobbes's account of the state of nature, and formalized by game theorists as the prisoner's dilemma. The "Troubles" in Northern Ireland, the racial conflicts in South Africa, the Israeli-Palestinian conflicts today – others could be added – display the incentives and behavior patterns that keep such cycles going. It is difficult to discover a way out.

Hobbes insisted that only an all-powerful sovereign could break the cycle and bring civil peace (Hobbes [1651] 2010). But communities characterized by chronic conflict either lack an effective sovereign, or perhaps worse, possess a state asserting sovereign authority which itself propels the conflict because it is perceived as illegitimate by a substantial proportion of the population. The South African apartheid state, for example, possessed a powerful military, an effective legal and bureaucratic apparatus, and strong political support among white South Africans; but it was powerless to impose a solution to a racial conflict driven in large part by the very character of that state.

Nor is partition an assured solution. Territorially separate states can sometimes minimize conflict by disengaging. That is more difficult where communities in conflict share the same territory, as in South Africa, Northern Ireland, and the former Yugoslavia, where the worst instances of ethnic cleansing and genocide occurred after partition, not before. In South Africa overturning the "Homelands" policy, a forced racial partition on extremely unequal terms, was one of the chief aims of the anti-apartheid movement.  Even to agree on fair terms of partition, and ultimately separate states (an outcome generally favored in the Israeli-Palestinian case), requires a degree of inter-communal cooperation that is difficult to achieve given the conflict's bitter history.

Democratic franchise is not by itself a solution, without democratic institutions regarded on all sides as fair and legitimate. What characterizes contemporary communities marked by chronic conflict is not that they lack democratic practices altogether, but that they are flawed democracies: polities that depend on appeals to democratic legitimacy which the system itself conspicuously fails to embody (Jung et. al. 2011, 83-86). The leaders of battling groups are not monarchs able to arrange a pact and then impose it on their respective subjects. In flawed democracies political leaders -- even as they engage in difficult negotiations with leaders of rival groups -- are accountable to separate constituencies, which retain the power to replace their leaders, and to block any agreement the constituency has not been persuaded to support. In commencing a risky reform process in 1990, for example, South

African President F.W. de Klerk looked to a future in which South Africans of all races would enjoy political rights in some form. But his reforms would have been stillborn if in the early 1990s he had lost majority support among white South Africans. Democratic institutions at their best hold leaders accountable to the community as a whole, potentially rewarding leaders who successfully bridge longstanding divides. But flawed democratic institutions often reinforce the very conflicts that prevent agreement on better democratic institutions.

It is tempting to categorize some conflicts as simply beyond resolution: to view the interests at stake as inherently zero-sum, one side's gain being the other's loss, and to dismiss as illusory any lasting, mutually-acceptable settlement. Today it is not uncommon to hear the Israeli-Palestinian conflict described this way: both sides make exclusive claims to the same territory, invoke irreconcilable religious identities to support those claims, and resort to violence that the other side can neither forgive nor forget. Perpetual war thus seems inevitable. But similar claims were made about South Africa in the 1980s. Indeed at the time, a peaceful resolution in South Africa looked less likely than in the Middle East. Few observers then believed that white South Africans, whose military power remained unmatched on the African continent, would accept a multiracial state ruled by a black majority. They did not understand that to characterize a conflict as inherently zero-sum because of rival parties' irreconcilable interests and values is to miss important ways in which events themselves, and the words and actions of leaders with a hand in those events, can reshape participants' political preferences and their perceptions about the nature of the conflict.

In this essay we argue that risk-taking acts of political leadership, marked by what we call strategically hopeful action, are needed to extract positive-sum outcomes from apparently zero-sum conflicts. These acts clear the way for agreement on new political institutions that are cured of the flaws that undermined the old order's legitimacy. Our theme is not leadership in general, but leadership of a specific kind exercised under especially difficult circumstances. Our principal case study is the leadership of Nelson Mandela and F.W. de Klerk that made possible (though by no means guaranteed) the South African democratic transition. Later we discuss the contrasting leadership records of the Northern Ireland settlement and the Israeli-Palestinian conflict. We illuminate strategically hopeful leadership by examining game theoretic models of conflict and cooperation in the work of Robert Axelrod, Thomas Schelling, and Josep Colomer, all of whom highlight important dimensions of the problem but underplay the specific risks and contributions of leadership. Finally we show how strategically hopeful leadership embodies a variable-sum understanding of power.

## II.    Strategically Hopeful Action

By strategically hopeful action we mean a certain kind of calculated risk-taking in the face of imponderably complex circumstances, the aim of which is to improve some aspect of the political and social world (at least in the actor's judgment). The paradoxical overtones of "strategically hopeful" are intentional: the phrase blends aspects of political action typically kept separate. The action we have in mind is strategic, in that attempts to resolve chronic communal conflicts without hardheaded calculation would be merely wishful thinking, and in some circumstances dangerously reckless. The action in question is hopeful in that it communicates a willingness to strengthen historical adversaries and take personal risks for a better future when key determinants of that future are, at best, imponderable. Displaying hope adds a performative ingredient at a time when it is vitally necessary. In periods of rapid political and institutional flux (like South Africa in the early 1990s) it is impossible for any leader to foresee, much less control, the wider flow of events. But a skilled leader can estimate to a degree the likely responses of an adversary to an unexpected opening. Among other skills this requires empathy, the capacity to imagine oneself in another's place and comprehend how the conflict looks from their perspective.  This was a skill both Mandela and de Klerk possessed to an impressive degree.

We illuminate strategically hopeful action by drawing from game theoretic models, and in particular the iterated prisoner's dilemma featured in the work of Robert Axelrod (Axelrod 1984; 1997). But in contrast to Axelrod's unitary rational actors, we highlight the strategic dilemmas that confront individual leaders as they try to resolve chronic conflict while retaining the support of refractory constituencies.  Bitter communal conflicts are usefully viewed as prisoners' dilemmas because most people on both sides would gain, compared to the miserable status quo, were they able to cooperate. But in the absence of either mutual trust, or a third party able to enforce an agreement, no party has a good reason to cooperate with the result that potential mutual gains are not realized. This will hold both for a single-shot prisoner's dilemma and for a repeated interaction where the stakes are high and the number of rounds is known beforehand.

Actual communal conflicts are often of uncertain duration, which opens up more interesting possibilities explored by Robert Axelrod. If A knows that she will face B again in the future (in an iterated interaction of indefinite length), and that B can retaliate in round (n+1) for what A does in round (n), then it might be in A's interest to cooperate in hopes that B will reciprocate; and in B's interest to reciprocate because if he defects, A will retaliate in the following round. This is the essence of "TIT FOR TAT": cooperate on the first round and then do whatever the other player does (Axelrod 1984, 13). But A takes a risk by cooperating first. If instead of reciprocating, B exploits A's cooperative move by defecting, A then defects on the next round, B is likely to defect again in turn, and the result is a self-reinforcing and mutually-costly cycle of defection. In contrast to a one-shot prisoner's dilemma, where the preferred

strategy is always the same (defect whatever the other does), the indefinitely iterated prisoner's dilemma is more open-ended and selecting a strategy is a genuine choice.

Axelrod found that the "nice" strategy of TIT-FOT-TAT won in computer simulations of iterated dilemmas of indefinite length, beating out a wide range of alternative strategies. The irony is that although TIT-FOR-TAT offers the highest average payoff, in any given round it never scores better than the other player, and always scores lower than an opponent who responds to cooperation by defecting. But players who pursue "nasty" strategies score badly against other "nasty" players: each drags the other down. TIT FOR TAT won "not by beating the other player, but by eliciting behavior from the other player which allowed both to do well" (Axelrod 1984, 112).

The communal conflicts that concern us here rarely permit a fresh beginning and exhibit instead a long, oft-repeated history of mutual violence and distrust. Axelrod recognized that with simple TIT-FOR-TAT, "once a feud gets started, it can continue indefinitely" (Axelrod 1984, 138). In later work Axelrod investigated tentative strategies for restoring cooperation when it has broken down. These include "generous" TIT FOR TAT (occasionally cooperating in the face of an opponent's defection) and "contrite" TIT FOR TAT (cooperating in response to defection that was in response to your own previous defection).[1] Such restorative strategies can work if they are employed about ten percent of the time, not more, otherwise they will be exploited (Axelrod 1997, 33-39). They are useful here because they allow us both to pinpoint the potential contributions of leaders to cooperative outcomes, and to highlight the risks they must be willing to take. Because generous TIT FOR TAT will be exploited if used too frequently, any leader who employs it assumes a significant risk -- for she or he must commit one hundred percent to the move once it is chosen.

Even more daunting is that in longstanding conflicts like those of South Africa, Northern Ireland, and Israel-Palestine, it is misleading to speak of "restoring" cooperation (as in Axelrod's model) because no golden age of cooperation ever existed. Moreover, even if most people on both sides prefer peace to war, there typically remains a determined minority on each side that prefers war to any negotiated settlement (as repeatedly demonstrated in Northern Ireland and the Israeli-Palestinian conflict.) Strategically hopeful leaders face the task of constructing inter-communal cooperation for the first time and creating an enduring constituency for it.

Axelrod did not address the role of leaders; the players in his models could just as well be whole communities, or even distinct species "cooperating" over the long span of evolutionary time (Axelrod 1984, 88-105). His narrative of the tacit "Live and Let Live" ethic (which emerged across trench lines mutually to limit casualties in World War I) shows that

---

[1] Axelrod explored the contrite and generous versions of TIT FOR TAT on the theory that we all know that people make mistakes from time, and would not want to be victims forever of an inadvertent error.

leaderless cooperation can sometimes reduce the death toll (73-87). But it cannot end the war. One can imagine a comparable ethic occasionally emerging in neighborhoods of Belfast or East Jerusalem, without resolving the larger conflict.

An individual leader can attempt a new, tentatively cooperative approach more readily than can whole organizations or whole communities, but in the process accepts a different degree of risk than does the community as a whole. Besides political failure and repudiation, which is penalty enough, such leaders make themselves targets for assassination by extremists on both sides. Others in a community can take a longer view: if one leader fails, they can elevate another; if a cooperative initiative fails, the community can pivot back to a warlike approach. Communities in long-term conflict thus mirror Axelrod's indefinitely iterated prisoner's dilemma. Individual leaders, on the other hand, may personally have only one major shot at cooperation across the divide, and if it fails, they could be finished. Their predicaments more closely resemble one-shot prisoner's dilemmas: the safe money bets the other way. To press forward with a cooperative opening, despite the daunting odds, is a mark of strategically hopeful leadership. In effect, such leaders internalize the costs of solving collective action problems that will otherwise continue plaguing their constituencies.

Game theorists typically assume that preferences remain fixed throughout the game and that players seek to maximize their payoffs while remaining indifferent to one another's utilities (Morrow 1994, 19-20, 34). Axelrod makes different and more illuminating psychological assumptions. He argues that in real-world approximations of the iterated prisoner's dilemma, "the very experience of sustained mutual cooperation altered the payoffs of the players, making mutual cooperation even more valued than it was before"; participants on the two sides began "to care about each other's welfare." The reverse also occurred: where cooperation failed, "a powerful ethic of revenge was evoked." Thus "not only did preferences affect outcomes, but behavior and outcomes also affected preferences" (Axelrod 1984, 85). Failed attempts at cooperation can reinforce tendencies to view conflicts as zero-sum, even when successful cooperation would have resulted in mutual gains. Axelrod observes that human beings need to learn that "most of life is not zero-sum" and that mutual cooperation is possible even if not always achieved. The successful experience of mutual cooperation can itself predispose human beings to view interactions in variable-sum rather than zero-sum terms (Axelrod 1984, 110-112).

To have any chance of bridging longstanding divisions, strategically hopeful leaders must first judge for themselves that the longstanding conflict in which they are enmeshed is potentially positive-sum, and therefore capable of resolution. They have to believe, and persuade others to believe, that all parties lose so long as the conflict continues unabated (even if they lose unequally), and that all sides stand to gain if the conflict can be resolved. Yet they recognize that accumulated bitterness and past failures to resolve the conflict have led many to perceive it as zero-sum: as a contest to be resolved only by victory for one side and defeat for the other. Game theoretic models (like the prisoner's dilemma) mask this

judgment problem when utilities over outcomes are assigned to actors. The payoff schedule of a prisoner's dilemma indicates that both actors are better off if they cooperate; the problem is that neither can trust the other. Leaders attempting to resolve historic conflicts understand the trust problem, but they also face another, equally difficult challenge: convincing skeptics on both sides that there is any potential outcome that could leave both parties better off. The mutual cooperation outcome in the prisoner's dilemma matrix, if never realized, may disappear altogether from participants' consciousness (Axelrod 1984, 110-112), so that the interaction appears as a game of pure conflict. Thus strategically hopeful leaders must not only play the game strategically; they must also persuade others about which strategic game is being played.

Strategically hopeful leaders must also judge when and how to set aside preconditions. Parties locked in chronic conflict typically insist on mutually-contradictory preconditions for negotiations. Preconditions highlight and often magnify the prisoner's dilemma character of the conflict because each side insists on securing before negotiations concessions that could only be obtained with great difficulty through negotiations; and in the unlikely event that its preconditions were met, it would no longer have any incentive to negotiate. Waiving one's own preconditions looks like waving the white flag. Strategically hopeful leaders must find ways to sidestep their own constituency's preconditions without acceding to their adversary's, and their risk will fail if "the enemy" refuses to meet them halfway. Strict preconditions -- for example a complete cessation of violence by the other side -- also empower the hardliners in each camp by enabling them to kill negotiations by deliberately violating those preconditions.

Strategically hopeful leaders recognize that their bridge-building attempts will fail without some modification of political preferences on both sides. If leaders instead merely seek to maximize the satisfaction of prevailing preferences – which in practice means the preferences of the particular group to whom they are immediately accountable – then the peace effort is doomed from the start. In times of regime transition, when stakes are high, institutions in rapid flux, and the outcome uncertain, political preferences become interdependent to an important degree: the aims of each side are conditioned in important ways upon the aims signaled by the other side. It is here that strategically hopeful acts of leadership can have their greatest impact. A cooperative opening that, against the odds, achieves early if limited success can change each side's perception of the other side's intentions. Changes in what is perceived to be possible alter in turn what is possible, to a degree that might have been unforeseeable to many before the process began. The assumption before 1994 that white South Africans would never relinquish their monopoly of political power peacefully – as though this were their definitive political goal as a people – turned out under the right set of circumstances not to be true.

The dynamic of interdependent preferences cuts both ways. On one hand it encourages the hope for a new opening. However risky and difficult the first step might be, if it appears

tentatively successful, its perceived success can alter the preferences and perceptions of other actors, generate a surplus of good will, and make successive moves possible that previously were not. But the same effect works in reverse: any serious misstep risks modifying preferences in negative directions, making future cooperation harder than before. Strategically hopeful leaders recognize that, if they fail, their failure constricts the future options of their successors.

Strategically hopeful leaders implicitly comprehend power, at least potentially, in variable-sum terms. Political power is frequently, perhaps even normally, gained at another's expense: one candidate wins an election, another loses. More fatefully, in the regime transitions that concern us here, the success of leaders committed to reform depends upon facing down their own radical flanks, whose members will see reformists' gain of power as their own loss and vice-versa. To perceive power as variable-sum does not mean that literally everyone gains, or that gains are equally shared, only that gains can exceed losses or losses exceed gains. Even during the horrific meltdown of the former Yugoslavia in the 1990s (where, it is often remarked, "Everyone lost,"[2] and whose path South Africa might have followed), every atrocity had its own relative winners and losers. A variable-sum understanding of power requires looking past the more obvious political gains and losses and taking into consideration the wider community's collective power over its own future, which can be enhanced or diminished by leaders' success or failure in risk-taking efforts at cooperation.

Strategically hopeful leaders need not make theoretical pronouncements about the nature of power. But they must recognize that their own power to secure a resolution acceptable to their own constituency depends upon preserving, and where necessary reinforcing, the power of the "enemy" leader across the table. The strategic temptation to weaken an antagonistic rival and divide the opposition must here be resisted; otherwise "enemy" leaders will be unable to persuade their constituencies to support an agreement. Leaders committed to resolving chronic conflict must recognize that in important respects their own power stands or falls with the power of their counterpart on the other side. This is easy to admit in principle but difficult to put into practice.

The narrative that follows examines two decisions, one taken by Nelson Mandela in 1985 and the other by F.W. de Klerk in 1992, that exemplify strategically hopeful leadership. Our narrative draws significantly, though not exclusively, from Mandela's and de Klerk's own accounts as expressed in public statements, autobiographies, and interviews. We do not regard either leader as possessing any comprehensive or impartial understanding of the events in which they participated. Both were limited in what they could see and even more limited in what they could control. But in order to understand their actions, we need to

---

[2] One of this essay's co-authors (Read) heard this characterization of the 1990s -- "everyone lost" -- repeated by many contacts during study trips in 2003 and 2006 to the former Yugoslavia.

reconstruct how they perceived the South African racial conflict, their own strategic options in attempts to resolve it, and the political and personal risks they faced in venturing a cooperative move under unpromising circumstances. Different leaders in Mandela's or de Klerk's place might well have chosen differently, with very different consequences for the country.

Nor do we treat either Mandela or de Klerk as saints, untainted by ordinary political ambition. Yet their political ambitions led them, at decisive moments, to risk career and perhaps life on making cooperative moves across the divide when the usual political incentives, and the advice of trusted colleagues, prescribed the opposite. Some rational choice analyses of leadership (see Frohlich et. al. 1971) treat leaders as political entrepreneurs who solve collective action problems in exchange for personal "profit" (e.g. appropriating revenues, distributing patronage, continuing in office). This describes many political leaders much of the time, but fails adequately to capture strategically hopeful efforts to resolve longstanding, bitter conflicts when all previous attempts have failed. Even success here may ultimately spell the end of a political career, as it did for F.W. de Klerk and David Trimble; and failure will end a career even sooner. Strategically hopeful leaders act as they do because they recognize the human cost of continued conflict. Their political ambition is to be remembered as one of those who resolved it.

## III.    Two leaders, two gambles

In 1985 Nelson Mandela (in Pollsmoor Prison at the time) decided to initiate secret talks with representatives of the National Party government, without the knowledge or approval of the executive committee of the African National Congress (ANC) - neither its leadership in exile, nor those ANC leaders also imprisoned at Pollsmoor with whom he was in regular communication. Mandela did not propose formal negotiations with the government; he only invited "talks about talks." Yet even this step was exceedingly risky, both for the ANC as an organization and for Mandela personally. The ANC's policy was that negotiations could begin only after the government had satisfied a number of preconditions, none of which had been met: revoking the legal ban on the ANC and other anti-apartheid groups, releasing all political prisoners, and allowing free and open political opposition. On the contrary, the government's repression of anti-apartheid activity was on the increase (Prime Minister P.W. Botha would declare a new State of Emergency in 1986). Meanwhile the government's own preconditions ruled out negotiations unless and until the ANC permanently renounced violent struggle and dismantled its military wing, Umkhonto we Sizwe.

Mandela recognized that overthrow of apartheid through outright military victory was "a distant if not impossible dream" (Mandela 1994, 457). But the *idea* of armed struggle, even on a limited scale, was enormously important to the ANC rank and file and could not be easily relinquished. By 1985 the Soviet Union had cut off the financial and military support it earlier provided to the ANC and the South African Communist Party (Jung et. al. 2011, 94). This would have increased the ANC's motivation to negotiate seriously with the government, but it also heightened the risk in doing so, because eagerness to talk might telegraph the ANC's diminished military capability and worsen its bargaining position. Under the circumstances, for Mandela to invite talks might betray ANC weakness in an increasingly high-stakes struggle. The government might have accepted Mandela's invitation, but with the ulterior aim of trapping him, discrediting him among his people and decapitating the ANC leadership at a crucial moment.[3]

Nevertheless Mandela decided to go forward. In his autobiography *Long Walk to Freedom* he describes why he did so, and also why he kept it secret from his ANC colleagues until he was committed:

> I had concluded that ….if we did not start a dialogue soon, both sides would be plunged
> into a dark night of oppression, violence, and war…It simply did not make sense for both

---

[3] During the 1980s competing forces within the South African government pushed policy toward the ANC in opposite directions: some sought to "decapitate" the ANC by dividing its leadership, while others sought to negotiate with the ANC in hopes of attaining a political settlement. The latter strategy presupposed an ANC leadership strong enough to negotiate. The immediate catalyst of the talks was a 1985 letter from Nelson Mandela to South African Justice Minister Kobie Coetsee. The "talks about talks" between Mandela and Coetsee began in early 1986 (Lodge 2006, 148-160).

sides to lose thousands if not millions of lives in a conflict that was unnecessary. They [the government] must have known this as well. It was time to talk.

This would be extremely sensitive. Both sides regarded discussions as a sign of weakness and betrayal. Neither would come to the table unless the other made significant concessions…Someone from our side needed to take the first step, and my new isolation [in a different floor of the prison from his colleagues upstairs] gave me both the freedom to do so and the assurance, at least for a while, of the confidentiality of my efforts.

I chose to tell no one of what I was about to do. Not my colleagues upstairs or those in Lusaka [the ANC leadership in exile]…I knew that my colleagues upstairs would condemn my proposal, and that would kill my initiative even before it was born. There are times when a leader must move out ahead of his flock, go off in a new direction, confident that he is leading his people the right way. Finally, my isolation furnished my organization with an excuse in case matters went awry: the old man was alone and completely cut off, and his actions were taken by him as an individual, not a representative of the ANC.

Once his initiative was underway, Mandela informed his ANC colleagues. Their responses were sharply divided, but most of them were, after the fact, guardedly willing to permit Mandela to continue an action they would not have approved in advance (Mandela 1994, 457-459, 466-467; see also Lodge 2006, 158-160).

Mandela's observations could be summarized in strategic terms as follows:

1) The conflict between ANC and government, and by extension between black and white in South Africa, was potentially variable-sum. Even though most people on both sides perceived it as zero-sum, in fact both sides stood to gain from a political settlement and both sides would lose terribly in the event of an escalating racial civil war.

2) Nevertheless, if present trends continued, the lose-lose scenario ("dark night of oppression, violence, and war") was the more likely, because "both sides regarded discussions as a sign of weakness and betrayal." Without using the term, Mandela was describing a classic prisoners' dilemma.

3) A leader's duty is to resolve this impasse, to "move out ahead of the flock," hoping this act will get people to perceive, and act on, the positive-sum possibilities. But this is risky because the leader cannot control the process he/she sets in motion; the other side may indeed suspect weakness and escalate its demands.

4) In taking this step on his own authority, Mandela internalized much of the risk, thereby diminishing it for other ANC leaders. He realized that his initiative might backfire, and if it did

his position as a top ANC leader would have been finished. Indeed he recognized bargaining leverage with the government in the fact that, if "matters went awry," his ANC colleagues could limit the damage by renouncing the initiative of an irrelevant old man. Mandela might "move out ahead of his flock," but he knew that the "flock" ultimately retained the freedom to denounce him and refuse to follow.

Mandela's 1985 decision in prison to "talk to the enemy" was a vital link in the chain of events by which apartheid was replaced by racially inclusive democracy (Lodge 2006, 165-166). But an equally crucial and risky decision was taken in 1992 by President F.W. de Klerk, the man who had released Mandela from prison two years earlier. Unlike Mandela's 1985 decision to invite "talks about talks," which occurred far from public view, De Klerk's 1992 gamble took place in the full glare of publicity.

De Klerk was a member of P.W. Botha's cabinet during the 1980s, but had not participated in, and initially was not informed of, the government's secret talks with the still-imprisoned Mandela. But he would have known about those talks when he first met with Mandela in November 1989, a few months after he took office as South African president (de Klerk 1998, 109, 156-158). Before his presidency de Klerk had been a typical conservative Afrikaner politician who displayed no reformist behavior.[4] But he had become convinced that apartheid had failed and that only fundamental reform "could pull South Africa back from the edge of the chasm on which we were teetering" (89). De Klerk's predecessor P.W. Botha had publicly admitted as early as 1979 that South Africans must "adapt or die," but Botha was unwilling or unable to follow through on this imperative or (in de Klerk's view) clearly to communicate the intention to do so (103-105). And even if Botha had been fully committed to reform, there was at the time no one on the ANC side with whom he could have negotiated. In contrast, de Klerk knew from the outset of his presidency that Mandela had shown himself to be someone with whom "it would be possible for us to do business" (158) - thus providing de Klerk with options none of his predecessors had enjoyed. In February 1990 de Klerk surprised South Africans and the world by lifting the ban on the ANC, the South African Communist Party, and the Pan Africanist Congress (PAC) unconditionally. He announced plans to release all political prisoners including Mandela, and to begin negotiations toward democracy (Jung and Shapiro 1996, 194-195; de Klerk 1998, 229-232).

But by 1992 the success of de Klerk's reform effort was greatly endangered. Political violence had escalated, especially in Natal province between ANC supporters and followers of Mangosuthu Buthelezi's ethnically Zulu Inkatha Freedom Party (IFP) – a conflict that brought South Africa as close as it ever came to civil war in the 1990s, and which de Klerk reports "almost destroyed our efforts to achieve a negotiated constitutional settlement" (de Klerk 1998, 193). Negotiations on a new constitutional settlement (Convention on a Democratic South Africa, or CODESA) had collapsed. There was an apparently unbridgeable

---

[4] De Klerk insists that his public image as a conservative in the 1980s was mistaken. De Klerk 1998, 101.

gulf between the ANC's demand for full majority rule and the government's insistence on permanent constitutionally-guaranteed veto rights for all minorities (the white minority of course included). Instead of trusting the government to lead the reform process, as de Klerk had hoped, the ANC had stepped up its campaign of mass action, and though it had announced a "suspension" of the armed struggle in 1990, it had not dismantled its military wing. The economy was spiraling downward. Personal relations between de Klerk and Mandela had soured. Mandela accused de Klerk of complicity in government-instigated violence, a charge the latter vehemently denied; de Klerk in turn suspected Mandela of deliberately concocting such charges as a tactic to maximize his bargaining leverage.[5] The country appeared headed for the abyss.

By this point de Klerk's standing among the white electorate, to whom he owed his presidency and thus any opportunity to engage in reform, was eroding dangerously. In late 1991 and early 1992 the Conservative Party, which was intensely opposed to de Klerk's reform projects, began winning by-elections in former National Party strongholds (Jung and Shapiro 1996, 198-199). The election results were widely interpreted as a rejection of the National Party's reform proposals on the part of white voters, who were at that point still the only voters who mattered.[6] "The mandate that I had received in 1989 from the white electorate was visibly slipping away from me and the National Party" (de Klerk 1998, 229-232). According to conventional wisdom, de Klerk at this point should have moved to the right, shoring up his political base and salvaging his party's governing status - even at the cost of scaling back or postponing his plans to dismantle apartheid. Many of de Klerk's supporters, advisors, and cabinet members urged this course.

Instead de Klerk made a high-stakes move in the opposite direction. On February 20, 1992, he surprised both the public and many in the National Party leadership by calling a snap referendum among white voters on the question: "Do you support the continuation of the reform process that the state president started on 2 February 1990 and which is aimed at a new constitution through negotiations?" De Klerk typically practiced consensus-seeking leadership in his party and cabinet, but in this case he consulted no one: "If I had put my decision to the vote the majority of the caucus would have opposed what they then regarded as an over-hasty and risky decision" (de Klerk 1998, 232). His act here parallels Mandela's unilateral decision in 1985 to initiate talks with the government. De Klerk made clear that if the referendum failed he would resign. He campaigned vigorously for the March referendum, winning an impressive 68.7 percent Yes vote.

---

[5] For de Klerk's version of this dispute, see de Klerk 1998: 199-204, 258-267, 384-385; for Mandela's, see Mandela 1994, 509-515. Subsequent investigations established that some members of the South African security forces were involved in illegal and violent covert activities. There is no evidence that De Klerk authorized these actions, and he attributed them to reactionary forces aiming to undermine his reforms.

[6] South Africa's 1983 Constitution had created a so-called "tricameral parliament" in which Coloured (mixed race) and Indian voters were permitted to vote for members of two separate and largely powerless chambers. Black South Africans, who made up the vast majority of the population, were still wholly disenfranchised under that constitution.

In a subsequent interview de Klerk explained his reasoning in calling the referendum: that in the negative by-election results people were expressing their fears and dissatisfactions, but that when confronted squarely with the issue of the country's future they would rise to the occasion and embrace the need for change. There was no polling data on the subject. He was relying on his intuitive sense of what moved his compatriots and why. Calling the referendum was the only unilateral decision of his presidency; his cabinet and the National Party leadership would have been unwilling to risk that his judgment about the white electorate was accurate (de Klerk, interview with Ian Shapiro, 2003).

But de Klerk survived one risk only to face another. The successful referendum had not ratified any particular post-apartheid constitutional settlement, but merely given de Klerk clearance to continue a process whose ultimate outcome he could not fully control. De Klerk would likely have been finished politically if, after winning the referendum, he had failed to secure an agreement with the ANC that was also acceptable to a critical mass of white South Africans. And in 1992 the two sides' constitutional demands were still fundamentally opposed. De Klerk and the National Party had promised during the referendum campaign that in any constitutional settlement they would never give up permanent, constitutionally-guaranteed power sharing for the white minority, while the ANC rejected any such provision as apartheid in another form. Only in late 1993 did de Klerk relinquish the permanent power-sharing demand, settling instead for constitutional guarantees on property rights, civil freedoms, and a two-thirds majority to alter the Constitution. The white minority's willingness ultimately to accept in 1994 what they still appeared implacably to reject in 1992 – black majority rule over which whites possessed no guaranteed veto rights – illustrates the fluid character of political preferences as events unfolded. (Whether de Klerk himself knew during the referendum campaign that he was promising more than he would be able to deliver later is an open question.)

De Klerk recognized that he could not impose a constitutional settlement unilaterally: he needed the cooperation of Mandela and his supporters in the ANC. De Klerk needed not only to face down the Afrikaner hard right; he also had to convince skeptics in the ANC that he could be trusted – this after the National Party government and the ANC had been through decades of bitter conflict in which countless numbers of people had been uprooted, imprisoned, tortured, and killed.

In characterizing his unilateral decision to initiate "talks about talks" Mandela pictured a shepherd moving out ahead of a flock. De Klerk, in characterizing his own challenge, favored river metaphors - in particular crossing the Rubicon. He observes that P. W. Botha realized the Rubicon would someday have to be crossed, but was unwilling to "take the unavoidable plunge" and accept that "partners in negotiation would have to include the ANC and its radical allies" (de Klerk 1998, 103-106). Unlike his predecessor de Klerk was willing to plunge into the Rubicon and start swimming. He had no guarantee, however, that anyone on

the other bank – the ANC and its radical allies – would do anything to help him to the other shore if he began to drown.

Both Mandela and de Klerk recognized the interdependent character of the negotiated transition: neither could succeed without the cooperation of the other. Both viewed the situation as variable-sum (they and their respective constituencies stood to gain or lose together, but each also hoped to gain more or lose less than the other). Both recognized that the other also viewed matters this way. Both knew there were hard-liners on both sides who saw the conflict as irreducibly zero-sum and anticipated a violent showdown. De Klerk observes of Mandela and himself: "We realized that we both bore the ultimate responsibility for ensuring that there would be a negotiated settlement and we were both committed to carrying out this responsibility" (de Klerk 1998, 169). Mandela voiced the same idea in recognizing de Klerk's "genuine and indispensable contribution to the peace process" and observed that, "To make peace with an enemy one must work with that enemy, and that enemy becomes one's partner" (Mandela 1994, 533). Both leaders believed that failure to cooperate would produce terrible consequences and horrifying losses for all sides – "a dark night of oppression, violence, and war" (Mandela 1994, 457), "a prolonged struggle so bitter and destructive that there would be little left for anyone to inherit" (de Klerk 1998, 121).

Yet both also recognized conflict over the terms on which these common interests would be met. De Klerk and Mandela acted "in the full knowledge that we were opponents with divergent goals" (de Klerk 1998, 169). Mandela claimed that de Klerk "did not make any of his reforms with the intention of putting himself out of power" but "for precisely the opposite reason: to ensure power for the Afrikaner in a new dispensation" (Mandela 1994, 503). In addition to differences of interest among the constituencies they represented and deep divergences between their constitutional proposals, there was little personal trust between Mandela and de Klerk during the negotiations process. Yet both were willing to take significant personal risks to bridge the racial divide, hoping that the other would reciprocate rather than exploit the move.

# IV. Leadership, Strategic Interaction, and the Prisoner's Dilemma

Mandela and de Klerk realized that resolution of the conflict depended on an interdependent decision: neither party could impose a unilateral solution, but instead had to offer something ultimately acceptable both to the adversary and to their own constituencies. The tragedy of interdependent decision is that outcomes worse for all sides may result because leaders fail to agree on the terms of cooperation.

There is an enormous literature on leadership, scholarly and popular. But little of it addresses the problem that concerns us here: how leaders democratically accountable to only one constituency of a longstanding communal conflict can initiate cooperation across the divide and then persuade their own constituency to cross that bridge. Analyses of democratic leadership that presuppose effectively functioning institutions widely perceived as legitimate (like the large literature on U.S. presidential leadership) do not reach the case at hand.

Rational choice theories of leadership often feature what William Riker calls "heresthetics" or "the art of political manipulation" – leaders strategically deploying words and actions to structure decision situations to their advantage, so that they get their way without having to persuade anyone to modify their preferences (Riker 1983; 1986; see also Dewan and Myatt 2012, 432). For example, by manipulating the order in which a set of alternatives is voted on - without changing anyone's preferences - one can often win a decision that would have been lost had the voting order been different. Both Mandela and de Klerk were, among other things, skilled "political manipulators" in Riker's sense. Recall that at key moments both took risks that their colleagues would have opposed had they been consulted beforehand but were guardedly willing to support later. But for both Mandela and de Klerk short-term "agenda manipulation" was directed to longer term ends that could only be realized if other key elites and substantial numbers of their constituents were persuaded, through both word and act, to modify their political preferences. Riker's heresthetic cannot account for the choices made by Mandela and de Klerk.

The notion that leaders merely manipulate fixed preferences provoked an initially healthy, but now excessive and faddish, counter-reaction in the enormous literature on "transforming" leadership pioneered by James MacGregor Burns (Burns 1978; 2003; Conger 2004). In contrast to merely "transactional" leaders, "transformative" leaders create new possibilities by modifying the character and preferences of followers. Transformative leadership theories identify an important element missing from rational choice accounts, but too often treat transformation as a quality inhering in a charismatic leader, or emerging from that leader's interactions with committed followers. At least equally important for understanding acts of communal bridge-building are leaders' strategic interactions with rival leaders, and with their constituencies – for whom an adversary's charisma might actually be a threatening liability.

In the South African democratic transition, Mandela's transformative capacities depended on de Klerk's responses and vice-versa, and on their securing the support of their skeptical constituencies – as both leaders fully understood. For these reasons we find theories of strategic interaction, even when they exclude or underplay leadership, more useful to the problem at hand than much of the leadership literature. But at the same time we encourage greater attention to the way leadership can shape the outcome of strategic interactions. Robert Axelrod's iterated prisoner's dilemma has already been discussed; here we consider strategic interaction theorists Thomas Schelling and Josep Colomer.

Schelling's *The Strategy of Conflict*, despite its vintage, remains among the most fruitful examinations of the type of decision problems considered here. Schelling makes a fundamental distinction between zero-sum (or "constant sum") conflicts, where "more for one participant inexorably means less for another"; and variable-sum, or "mixed motive" conflicts, in which "there are common as well as conflicting interests among the participants" and "mutual dependence as well as opposition." Strategy in a variable-sum game must take into account not only "the division of gains and losses between two claimants" but also "the possibility that particular outcomes are worse (better) for both claimants than certain other outcomes…There is a common interest in reaching outcomes that are mutually advantageous." Neither participant can fully control the outcome; instead, "the ability of one participant to gain his ends is dependent to an important degree on the choices or decisions that the other participant will make" (Schelling 1980, 5). Cooperation might fail because whoever first indicates willingness to cooperate signals potential weakness and thus invites exploitation by the other side; or because either or both sides are unable credibly to commit themselves to positions from which they will not later be tempted or pressured to move. For negotiations to succeed, the expectations of both sides "must somehow converge on a single point at which each expects the other not to expect to be expected to retreat" (35, 70).

This means that if either party weakens the other too much, he will jeopardize the possibility of an agreement. Mandela illustrates Schelling's point when he writes that despite the frictions "I never sought to undermine Mr. de Klerk" because "the weaker he was, the weaker the negotiations process" (Mandela 1994, 533). Mandela did on occasion push ANC demands close to the limits of what de Klerk could accept, but never exceeded those limits; he never forgot that de Klerk had his own constituencies to satisfy.

Schelling's variable-sum model aptly describes the South African racial power contest as perceived by both Mandela and de Klerk. Both leaders agreed that through cooperation each side stood to gain – at least compared to the likely outcome without cooperation. Both recognized the interdependent character of the decision: if the leadership on either side refuses to sign, or agrees to a pact they cannot persuade their constituency to support, then both sides lose, as do the leaders themselves, politically and personally. Each recognized that the other also recognized these things.

But Mandela and de Klerk also understood something not provided for by Schelling: that whether a conflict is perceived as zero-sum or variable-sum may itself be at stake in what leaders do and say. Schelling takes as given that a conflict is either zero-sum or variable-sum, which in turn is a function of the respective (and independently derived) preferences on each side. This initial classification is decisive because "the intellectual processes of choosing a strategy in pure conflict and choosing a strategy of coordination are of wholly different sorts" (Schelling 1980, 4, 96). But Schelling does not explain how participants decide whether they believe the conflict to be zero-sum or variable-sum. Mandela's inviting "talks about talks" with the government in 1985 was risky because, though he himself saw the conflict as potentially variable-sum, he did not know whether his colleagues or the National Party government would agree. For Mandela to frame the conflict as involving potential shared gains was a hopeful act.

Even if leaders on both sides view a conflict as potentially variable-sum, cooperation might fail because their respective constituencies – who retain the power to block agreements and replace their leaders – might instead see the conflict as a fight to the finish. Mandela noted that many anti-apartheid activists in the early 1990s demanded "a victory on the battlefield, not the negotiating table" and he noticed signs at rallies that read, "MANDELA, GIVE US GUNS" and "VICTORY THROUGH BATTLE NOT TALK" (Mandela 1994, 506, 526). De Klerk was no less aware of the white right who anticipated a violent showdown (1998, 316-319). Both leaders thus faced the twin challenge: first skillfully to manage a difficult bargaining process with adversaries, and then to persuade a critical mass of the rank and file on both sides to agree with them.

Unlike some more formalized versions of rational choice theory, Schelling does not insist dogmatically on invariant preference functions. But his bargainers maximize their share of a fixed supply of gains from cooperation. He ignores the ways in which players with interdependent preferences can expand (or shrink) the available surplus depending on their behavior in successive rounds of bargaining. This is the performative dimension of cooperation: when successful, it creates a new reality (Austin 1975).

Schelling also downplays the moral character of leaders. For him, negotiations between warring underworld gangs, or between kidnappers and victims, do not differ in their strategic logic from any other bargaining situation in which conflict and cooperation are mixed (Schelling 1980, 12, 43-44). Schelling's undemanding leadership assumptions reach their limits in explaining anything as difficult as the South African settlement. What Mandela was asking white South Africans to give up – their monopoly on political power, in exchange for an uncertain future under black majority rule – went well beyond the kind of concessions featured in *The Strategy of Conflict*. It was unsurprising that many people on all sides saw the conflict as zero-sum, and in particular that white South Africans anticipated harsh treatment from an empowered black majority they had oppressed for so long. White South Africans still retained decisive military and economic superiority, and some on the right would

have gladly pushed de Klerk aside to spearhead a violent contest if the white minority had refused to be ruled by a black majority. Both de Klerk and Mandela kept a close eye on the South African military throughout a democratic transition process that the military could have halted had it chosen to do so.[7]

Yet in the end white South Africans peacefully turned over power to a black-majority government headed by Nelson Mandela, who promised them no more – but also no less – than to live as equal citizens in a multiracial democracy. That "South Africa belongs to all who live in it, black or white" was affirmed in the ANC's Freedom Charter of 1955. But few South Africans of any race had personally experienced a political community of this kind. One of Mandela's accomplishments as a leader was to make this vision believable, both to its advocates after nearly a century of frustration, and to white South Africans who – no matter what the Freedom Charter said – feared that black rule meant oppressor and oppressed exchanging positions.

Mandela understood this fear and addressed it. "I have fought against white domination, and I have fought against black domination. I have cherished the ideal of a democratic and free society in which all persons live together in harmony and with equal opportunities" (Mandela 1994, 322; 2004, 59-62). But it was his actions that made the statement believable, just as de Klerk's turning on the white right and plunging into the Rubicon established his credibility. Taking these risks enabled them to persuade their supporters to rethink the conflict in variable-sum terms. Once they did, they could focus on their enduring common interests beneath the more obvious conflicts.

Schelling does not specifically discuss democratic transitions. Josep M. Colomer's *Strategic Transitions: Game Theory and Democratization* (2000) is more specifically directed to the problem at hand. Colomer illuminates the dynamics of democratic transitions by examining possible coalitions among six types of strategic actors: radical democratic opposition; moderate democratic opposition; radical soft-liners (reformists in the regime); moderate soft-liners (less committed reformists); moderate hard-liners; and radical hard-liners (Colomer 2000, 49). His model is based on the actors, strategies, and outcomes of the democratic transitions in the former Soviet bloc, but with some modification it can also illuminate South Africa, Northern Ireland, and other cases.

Colomer seeks to explain why confrontation between regime and democratic opposition sometimes produces stable compromises in favor of democratic or semi-democratic institutions; and other times triggers "frontal conflict" whereby both sides "fight to eliminate each other," and consequently both risk "becoming an absolute loser" (Colomer 2000, 1-3).

---

[7] The military did not signal decisive support for the transition until March 1994, when they accepted an order from the Transitional Executive Council (by then the caretaker government) to put down a white separatist group in that was supporting a local black leader in the Tswana homeland of Bophuthatswana who opposed the coming elections (Jung and Shapiro 1996, 201-4).

For example, strategic bargaining between moderate democratic opposition and moderate hard-liners in the regime resembles the prisoner's dilemma: there is potential for democratic compromise if each actor settles for its second choice, but each will resort to civil war rather than allow the other its first choice (57). Like Schelling, Colomer presupposes variable-sum interdependent decision: neither regime nor democratic opposition is powerful enough unilaterally to determine outcomes, and failure to cooperate leads to consequences worse for both. Like Axelrod, Colomer takes the prisoner's dilemma as a starting point (57-61) and seeks to explain how "cooperation can develop even among antagonists" (Colomer 2000, 51, quoting Axelrod 1984).[8] But Colomer argues that critical transitions more closely resemble single-shot prisoner's dilemmas than Axelrod's iterated prisoner's dilemma, because "…actors' opportunities to repeat interactions in a process of regime change may be remote" (Colomer 2000, 51).

To explain how cooperation can emerge from a single-shot prisoner's dilemma, Colomer endows actors with the capacity for foresight. Mutually destructive outcomes result from "short-term-looking decisions…made in ignorance of other actors' choices." In contrast, cooperation might result even in a single-shot prisoner's dilemma because actors are "able to foresee not only the immediate consequences of their choices but also the consequences of the other actors' reactions, their further counter-reactions, and so on" (Colomer 2000, 2). Colomer's prescient actors "do not make simultaneous or blind choices of strategies," as in the classic prisoner's dilemma. Instead they "enter into open, dynamic interactions in which they make choices precisely in the expectation of other actors' reaction, as is typically the case in political exchanges" (52). In contrast to the common game theoretic assumption that a move once taken cannot be reversed, Colomer stipulates that strategic moves can be retracted without paying a high price; for example, "if the hard-liners reject an intermediate formula proposed by the opposition, the latter can resume mobilizations and protests" (57-58). By introducing foresight, political exchanges, and retractable moves, Colomer builds flexibility into the prisoner's dilemma and prevents it from converging inevitably on mutual defection.

Colomer's introduction of foresight into the prisoners' dilemma interaction – the principal innovation of his model – simultaneously reveals a limitation: the model includes no specific role for leadership. If foresight was abundantly and equally distributed among human beings and operative in all strategic interactions, prisoner's dilemmas would always be resolved and civil wars would seldom occur. Most people most of the time operate with limited and inflexible cognitive capacities (Kahneman 2011, 19-108; 259-376). Colomer employs generic actors representing strategic positions (like "moderate democratic opposition"), not individual leaders, even when he illustrates those positions with proper names like Gorbachev, Jaruzelski, and Havel (Colomer 2000, 41). If the classic prisoner's dilemma underestimates

---

[8] Colomer also examines what he labels "mugging games," which resemble prisoner's dilemmas in some respects but, unlike the PD, give one of the actors a strategic advantage over the other. Colomer 2000, 53-57. For simplicity our commentary will focus on Colomer's treatment of the PD.

the cooperative possibilities, Colomer's revision, while insightful, downplays the difficulty of exercising foresight in periods of rapid and extraordinarily complex change. Nor does the model provide for the different degree of risk run by individual leaders as compared to collectivities during such periods.

Colomer assumes, for example, that actors' strategies will be weighted by a preference for peace over war: "A transition by agreement is less risky for the actors involved than a civil war" (Colomer 2000, 2). This is true as a general proposition (otherwise negotiated solutions would be impossible). But for individual leaders it is often far more risky, politically and personally, to initiate a cooperative opening where there has been a long history of conflict and distrust than to do nothing. What would be ultimately risk-minimizing for the group, if it can be achieved, may be risk-augmenting for the leader who makes the first move. Colomer assumes that parties can retract a failed cooperative opening without significant cost. But the leader responsible for a failed opening might pay a higher price: disgrace, the end of a career, perhaps assassination. A strategically hopeful leader must take into account not only the risks and opportunities faced by both communities in any attempted reform, but also the individual risks he or she runs as leader, and the corresponding risks faced by leaders on the other side of the divide. The foresight required under such circumstances is more complex than the general recognition (which Colomer assumes) that both groups lose if they fail to cooperate.

Colomer also understates the difficulties of negotiated transitions by assuming "the absence or weakness of maximalist actors…the radical opposition and the radical hard-liners must be weak enough to be considered irrelevant in the main interactions" (Colomer 2000, 61). Whether this assumption fits the post-Soviet transitions may be debated; it is clearly unrealistic in many other circumstances. Even if large majorities on both sides favor peace, the process can be held captive by a committed minority that prefers war to any negotiated solution – as demonstrated with depressing regularity in Northern Ireland and the Middle East. Hard-liners cannot be assumed to be marginal; they must be actively marginalized if they cannot be convinced or coopted. This is among the most difficult and essential tasks of a strategically hopeful leader. If one cannot be confident that maximalist actors (in the adversary's camp as well as one's own) can no longer be effective spoilers, then one cannot know what strategic game is being played. For among maximalist actors it is a purely zero-sum conflict, not a prisoner's dilemma.

## V. Zero-sum and Variable-sum Power

The prisoner's dilemma is a parable of power and powerlessness. The interrogators are powerful because they make and enforce the rules of the game. The prisoners lack power to realize their common interests because they can neither alter the rules, nor trust one another to cooperate under those rules. With respect to utility, the prisoner's interaction is potentially variable-sum: mutual cooperation yields the highest aggregate utility and mutual defection the lowest. But with respect to power, their interaction appears zero-sum. For the only aspect of the interaction lying within each prisoner's control is deciding whether to cooperate or defect. Cooperation leaves one powerless and vulnerable. Defection promises each prisoner a small degree of power, though of course at the other's expense.

An iterated prisoner's dilemma creates a wider range of strategic options: if each actor must take into account future retaliation or cooperation by the other, defection is no longer obviously a dominant strategy. An invitation to mutual cooperation is sometimes successful, though it offers no guarantee. This expanded range of strategies, which Axelrod and others have examined, suggests an alteration of the power relations among the actors. This latter possibility has not received much discussion, perhaps because the concept of power is not explicitly employed in formal game theory, though assumptions about power are implicit in the structure of the game.[9] If under the classic one-shot prisoner's dilemma, each actor's power comes at the expense of the other, and aggregate power is fixed at a low level; then it should follow that, under an iterated prisoner's dilemma where reciprocity and trust can potentially emerge, the power of each actor does not necessarily come at the other's expense, and the power available to both actors is potentially variable, not fixed. This hypothesis requires that we examine the concept of power itself.

The meaning of power has long been contested (Lukes 2005, 14-38, 60-107). Hobbes defined it as one's "present means to obtain some future apparent good" (Hobbes 2010, 93). This definition of power is not inherently zero-sum; in principle we might secure our own good by cooperating with others in realizing theirs. But where interests conflict, or where shared interests go unrealized because trust is absent (as in the state of nature), power for Hobbes becomes zero-sum in practice. Thus elsewhere Hobbes provides the classic zero-sum description of power: "Because the power of one man resisteth and hindereth the effects of the power of another: power simply is no more, but the excess of the power of one above that of another. For equal powers opposed, destroy one another, and such their opposition is called contention" (Hobbes 1928, 26).

For Hobbes, one's power is demonstrated through victory in a head-to-head contest. This premise is shared by the realist school of international relations theory, who perceive an

---

[9] A typical instance is Morrow (1994) who nowhere mentions power. For the implicit presence of power in game theory see Dowding (2011, xxiv).

anarchic realm in which war-winning capacity counts above all (Gilpin 1981, 94; Mearsheimer 2001, 2). But Hobbes's idea is also echoed in the way many political scientists have defined power even under the operation of peaceful democratic rules. According to Nelson Polsby - speaking here as a "pluralist" in the so-called three faces of power debate that began in the 1960s - in studying power within a political community we should focus above all on who wins and who loses in a "direct conflict between actors" because this is the best measure of their respective "capacities to affect outcomes" (quoted in Lukes 2005, 18). The "radicals" in this faces of power debate pointed out that conflict and domination were not always readily visible. But they did not call into question the assumption that power always comes at another's expense; on the contrary they drove that premise further. (See Lukes 2005 for that debate, including Lukes's recent revision of his original view on the matter [63-65]).

Conflict and inequality are always relevant to describing political power. The more problematic assumption is that A's capacity to "affect outcomes" is a direct function of B's lack of capacity to affect those same outcomes. This doubtless describes many power relations, but not the negotiations between Mandela and de Klerk. If those interactions were conflictual and "subtractive" in some respects, they were cooperative and "additive" in others. Conflict and cooperation fused in the same complex relation; both equally shaped outcomes. To comprehend as power only the conflict and ignore the cooperation – as though cooperation were passive, something other than power – is to misunderstand the power relationship itself.

The assumption that power is inherently zero-sum has not gone unchallenged. Power has alternatively been defined as the cooperative capacity to achieve collectively shared aims (Talcott Parsons) or the general capacity of human beings to "act together" (Hannah Arendt). But Parsons's and Arendt's arguments, though insightful, are flawed by their treating power as almost wholly consensual and cooperative, thus neglecting the element of conflict that preoccupies other theorists of power (for discussion and critique see Lukes 2005, 30-35; Read 2012). Some other challenges to the zero-sum view do incorporate both conflict and cooperation, but do not develop an alternative variable-sum description in much detail (for example Giddens 1984; Baldwin 2002; Wrong 1995; for a more fully-developed variable-sum argument incorporating both cooperation and conflict see Haugaard 2012). The zero-sum view of power, despite its one-sidedness, has produced rich descriptions of political life (for example Gaventa 1980). The variable-sum alternative view, to be fully persuasive, must pass this same test.

The political interactions between Mandela and de Klerk, and by extension between the constituencies they led, suggest that a variable-sum theory of power can persuasively describe significant dimensions of political life - without neglecting inequality, domination, the potential for violence, and the persistence of conflict. Both Mandela and de Klerk implicitly perceived the racial struggle in variable-sum, not zero-sum terms. When Mandela reminded

South Africans that "I have fought against white domination, and I have fought against black domination," he was challenging a pernicious zero-sum assumption about power: that liberation for either community entailed domination for the other. Mandela's strategically hopeful view – one shared with de Klerk, despite their divergent constitutional visions – was that black South Africans could be liberated without dominating whites (or any other group of this multiracial society).[10]

But Mandela and de Klerk also realized that many South Africans on all sides saw one group's gain as the other's loss, and feared that this zero-sum view would produce a negative-sum outcome unless the leaders actively cooperated to prevent it. Both were fully aware of the conflicting aims of their respective constituencies and of their personal conflicts as rival leaders. Yet both also recognized that each had to keep the other strong enough to be able to secure the support of their constituencies for any negotiated settlement, and for this reason could not push an advantage too far. They can be imagined as actors in a prisoner's dilemma who, by taking risks to generate trust, have found a way out of the prison. In the classic prisoner's dilemma the powerlessness of each ensures the defection of both. Mandela and de Klerk overcame this dilemma because their interactions generated new power on both sides.

The variable-sum view of power sketched out above does not "refute" the more common assumption that gains of power for some entail losses of power for others. The "your gain is my loss" suspicion remains relevant for at least two reasons. First, even where power is cooperatively generated, there remain important respects in which power gains for some involve power losses for others. In a stable democracy rival parties compete for a fixed number of powerful offices even as they cooperate to maintain the power of the democratic system itself (whose long-term survival can never be taken for granted).  In the South African case, the reversal-of-domination scenario so feared by white South Africans did not come to pass, in part because de Klerk and Mandela cooperated to prevent it. In this sense the variable-sum model best captures events. Yet it is undeniable that white South Africans had to relinquish significant political power in order for black South Africans to gain new political power in the transition to non-racial democracy; in this sense the zero-sum view retains its attraction.

Another reason the zero-sum hypothesis cannot be dismissed is that it retains the power of a self-fulfilling prophesy. If one views power as radically zero-sum, one can indeed make it zero-sum.  Had either Mandela or de Klerk been replaced by leaders who perceived the racial power contest in zero-sum terms and acted accordingly, political meltdown and an escalation of racial war would have been the likely outcome. A zero-sum power strategy, once decided upon, will "prove itself correct" because the adversary responds in kind. A

---

[10] For present purposes we ignore the interactions between both sides and the then two million or so "Cape Coloureds" as well as the ethnic Zulu Inkatha Freedom Party. For discussion see Jung and Shapiro 1996, 185-99, 201-13; and Jung, Lust-Okar and Shapiro 2011, 105, 122, 146-53.

strategically hopeful opening, in contrast, might fail. Either strategy is rationally defensible. Either understanding of power can fit the facts of the conflict. There may be no "best strategy" but instead a morally-charged choice among strategies.

# VI.    Unionists and Nationalists, Israelis and Palestinians

The South African constitutional resolution was not replicated in Northern Ireland or Israel. Those conflicts differ in their histories, the ideologies invoked, the role of outside actors, and the design of the settlements realized or proposed. It is illuminating, nonetheless, to take account of some shared similarities from the perspective developed here. First, they are (or were, in South Africa's case) high-intensity, self-reinforcing, protracted conflicts with no visible end in sight. If anything, the Israeli-Palestinian conflict appeared to many commentators during the late 1970s the most likely to be resolved, and South Africa the least. Particular circumstances have always shaped the direction and intensity of these conflicts, but none of them seemed likely to dissipate merely by a favorable turn of events.

Second, none of the conflicts could be resolved by a decisive victory followed by an imposed solution.  They are what Samuel Huntington described as "transplacements" and Adam Przeworski as "extrications"— situations in which any resolution must be negotiated because, although both parties lack the power to impose unilateral solutions, the potential exists for reformers in the government and moderates in the opposition to negotiate a settlement that would command enough support to survive (Huntington 1991, 113-114, 152; Przeworksi 1991, 67-69).

Finally, all three of the conflicts could be described either as zero-sum or variable-sum, depending on how one frames them and whether one regards the preferences of the principal actors as fixed or dynamic. There was nothing about the South African conflict that made it inherently more variable-sum than the other two. Apartheid was obviously incompatible with non-racial democracy. Viewing Northern Ireland as an integral part of the Republic of Ireland contradicts the ideology that it is a permanent part of the United Kingdom. Israelis and Palestinians advance irreconcilable historical and religious claims to exclusive possession of the same land. If human beings were no more than replicas of the ideologies they espouse, then all of these conflicts would be inescapably zero-sum. In fact, people who live through chronic conflicts often suffer greatly, and most appear to place more value on physical and economic security and personal liberty than on pursuing ideological stances to their ultimate conclusion.[11] The desire for these goods can generate a politics of battle, a politics of reconciliation, or some mix – depending the strategies chosen and the responses they call forth.

---

[11] In the 2010 Northern Ireland Life and Times survey, for example, only 13% of respondents said they would find it "almost impossible to accept" if Northern Ireland were ultimately to be joined to the Republic of Ireland; 85% of respondents would either "happily accept" or "could live with" this result. Alternatively, if Northern Ireland were never to join the Republic of Ireland, only 2% found this "almost impossible to accept" while 93% would either "happily accept" or "could live with" this result. "Improving cross-community relations" and reducing unemployment were ranked as higher priorities than resolution of the national affiliation question. http://www.ark.ac.uk/nilt/2010/Political_Attitudes/index.html [09-12-2012]. Recent surveys of Palestinians indicate that practical concerns like employment and physical security have highest priority. Palestinian Center for Public Opinion, www.pcpo.org/polls.htm [09-12-2012].

Protracted, self-reinforcing conflicts will continue unless a leader of at least one of the parties stakes their career on altering the dynamic. Even so, the odds of failure are high; there are many more ways for negotiations to fail than for them to succeed. And leaders who successfully bridge the abyss against the odds may receive few political rewards for doing so, as the Northern Ireland case demonstrates.

The 1998 Good Friday (or Belfast) Agreement should be considered a qualified success because it has ended most of the political violence, and all significant players have remained committed to employing only peaceful means in pursuit of their political goals. (Peace efforts stalled whenever parties insisted on hard preconditions and advanced only when preconditions were waived; see Mitchell 1999, 22-38). The fundamental question - whether Northern Ireland shall remain part of the United Kingdom or merge with the Republic of Ireland - remains unresolved, and the Protestant and Catholic communities remain highly politically and culturally segregated (McGlynn et. al. 2012; McGarry and O'Leary 2009, 65-69.) But all major parties, including those most committed to a united Ireland, seem to have accepted what is called the principle of consent: that "Northern Ireland should remain in the UK as long as a majority of Northern Ireland's citizens support this status" and that any unification with Ireland requires majority support of Northern Ireland's people in referendum (McGarry and O'Leary 2009, 56).  The Agreement was facilitated by external actors – especially Britain, the Republic of Ireland, and the United States – but its consummation would have been impossible without risk-embracing political leaders from both sides of Northern Ireland's communal divide. Yet those leaders occasionally lost their courage or dragged their feet in ways that weakened the example they set.

In Northern Ireland the political incentives for communal reconciliation have always been slim. Ian Paisley, Unionist firebrand and scourge of Catholicism, expressed the once-prevailing view in both camps: "A traitor and a bridge are very much alike, for they both go over to the other side" (quoted in Powell 2008, 54-55). In the four decades preceding the Good Friday Agreement, Unionist leaders like Terence O'Neill and Brian Faulkner who reached across communal lines found their political careers abruptly cut short. For Nationalists and Republicans[12] (who seek to unite Northern Ireland to the Republic of Ireland) the political rewards for compromise have been equally meager. Among Nationalists the most significant bridge-builder has long been John Hume of the Social Democratic and Labour Party (SDLP), who since the 1960s has advocated what ultimately became central principles of the 1998 Agreement: commitment to exclusively peaceful measures, the legitimacy of both Northern Ireland political traditions, and the principle of consent (McLoughlin 2010). But the violent repression of Bloody Sunday in 1972 discredited Hume's nonviolent approach and energized the Irish Republican Army. Any Republican leader who

---

[12] The "Nationalist" label refers broadly to all who seek to unite the North to the rest of Ireland, but is also employed to distinguish parties like John Hume's SDLP that work within the system and endorse only peaceful methods from "Republicans" like the Irish Republican Army and its political affiliate Sinn Féin, who consider all existing Irish governments illegitimate and have in the past condoned the use of violence.

agreed, even provisionally, to a partitioned Ireland invited the fate of Michael Collins in 1922. Gerry Adams, the Sinn Féin leader who eventually turned the republican movement toward politics and (gradually) away from violence, never forgot that in doing so he risked assassination (Powell 2008:100, 147-148).

Yet by 2007 Paisley and his Democratic Unionist Party (DUP) were willing to cross a bridge (the 1998 Agreement) whose builders they had denounced as traitors, and accept power-sharing with Sinn Féin and Gerry Adams, who in entering government tacitly accepted a divided Ireland, at least for the present. The 2007 electoral success of the DUP and Sinn Féin came at the expense of the parties and leaders who did the most to make the 1998 Agreement possible. In the 1980s John Hume, recognizing that no lasting peace was possible without the participation of Sinn Féin, risked his reputation by entering into initially secret talks with Gerry Adams, hoping to persuade Republicans to declare a ceasefire (McLoughlin 2010, 153-167). The IRA ceasefire was so long delayed that Hume's eventual success appeared failure at the time. Hume insisted that Sinn Féin be treated as a legitimate party to the Good Friday settlement -- perhaps recognizing that a successful peace agreement would boost Sinn Féin's political fortunes at the expense of Hume's own party, as indeed happened (McGlynn et. al. 2012, 10-14).

Because other unionist parties were opposed to the Good Friday Agreement, its success critically depended on the Ulster Unionist Party (UUP) and its leader David Trimble. Trimble's party and political career afterward suffered eclipse for his efforts, in part because the IRA's long delay in disarming left him hanging (Powell 2008, 203-205). But Trimble's own limitations as a leader contributed. In 1997 he made the risky decision to enter negotiations that could have gutted his support base, and he kept the UUP at the table despite significant opposition within his own party and right-wing charges that he was betraying his own people (Mitchell 1999, 108-117). But when it came to the May 1998 referendum on the Agreement, Trimble's political courage wavered. In contrast to F.W. de Klerk, who called the 1992 referendum and personally led the successful campaign for a Yes vote, Trimble – though he continued to voice support for the Agreement – stayed largely on the sidelines for most of the campaign, apparently in response to the heat he had taken earlier. The task of persuading Protestants to support the Agreement fell instead to a coalition of unaffiliated citizen groups. Leaders of this Yes Campaign repeatedly advised Trimble "to become more forthright and to engage in more active campaigning," and in the final weeks before the vote Trimble did become more active (Hancock 2011, 103, 111). The Yes vote ultimately succeeded among Protestants, but just barely, and many of the Yes votes were shaky (Hayes and McAllister 2001).

Trimble's political future was uncertain in any case, but his episodes of hesitation did nothing to restore his political fortunes, and clouded the legacy of his more courageous moments. The same is true Adams's unwillingness or inability to persuade the IRA to disarm in the years immediately following the 1998 Agreement. The IRA's very-delayed decommissioning

in 2005 was grudging, not an act of communal reconciliation. Trimble and Adams deserve credit for the risks they took. But in the end John Hume, who first attempted a bridge to Unionists, then risked his reputation and his party's electoral future to persuade Sinn Féin to cross that bridge, set the strongest example for Northern Ireland's future leaders.

Today the Israeli-Palestinian conflict appears impossibly difficult to resolve, despite its continuing urgency. Both sides view even the hint of compromise as signaling weakness, and insist on preconditions each knows in advance the other will reject. Repeated failures to secure peace have reinforced a penchant to view as some essential, primordial antagonism what has in fact resulted from contingent choices and repeated failures of leadership. Here we focus on one particular missed opportunity.

In 1993 Israeli Prime Minister Yitzhak Rabin decided for the first time to talk to the Palestinian Liberation Organization (PLO) and its chairman Yasser Arafat, who was then the most powerful Palestinian leader and essential to any deal. Arafat was no Nelson Mandela. He headed a corrupt and ineffective organization, and his leadership of the Palestinian cause was often shaky. He also lacked Mandela's strategic judgment, as evidenced by his reckless decision to side with Iraq in the first Gulf War. But Arafat had taken a historic and politically-risky step when he publicly acknowledged the State of Israel in 1988. In the mid-1990s Arafat, whatever his faults, appeared willing and able to secure Palestinian support for an agreement that showed strong signs of being potentially viable (Jung et.al. 2011, 107-114).

By 1995 both Rabin and Arafat were favorably positioned to manage the hard-liners on their respective flanks and consummate an agreement, the main elements of which had been hammered out in secret negotiations in Oslo and then announced in the fall of 1993. Rank-and-file support was strong in both communities for the two-state solution envisaged at Oslo. Rabin was a war hero whose dedication to Israel's security was not in doubt. Arafat committed himself to policing the West Bank to secure Israel from Palestinian attacks. Both were personally invested in the process and both recognized a common interest in preventing outbreaks of terrorism. Had the process continued, they might well have reached a provisional agreement delivering benefits that would have replenished their political capital for further negotiations.

Rabin's assassination in November 1995 by an Israeli right-winger opposed to the peace process was a stunning blow. But even then tragedy might have been turned to opportunity had Rabin's successor, Shimon Peres, been willing to take greater political risks. Peres could have called a snap election in the wake of the assassination and won an endorsement from the Israeli public for continuing the peace process, analogous to de Klerk's March 1992 referendum. At the time, public opinion on both sides strongly favored a two-state solution and outrage at Rabin's assassination had all but the most fanatical Israeli right on the

defensive.[13] Major issues remained unresolved, but this was also true in South Africa in 1992. One round of successful negotiations shifts perceptions of what might be possible in the next round, which in turn changes what is possible.

But Peres missed the opportunity, tacking instead to the right. He permitted the assassination of Hamas militant Yahva Ayyash in January 1996, reinforcing the cycle of violence and closures on the West Bank and Gaza Strip. Peres responded to attacks from Southern Lebanon by bombing Lebanese refugee camps. A wave of suicide bombings in spring of 1996 hardened the Israeli stance in negotiations. Palestinian radicals thus helped secure the victory of the Israeli right. Peres alienated himself from Israeli supporters of the negotiations, and in the May 1996 elections he lost to Benjamin Netanyahu who had made no secret of his hostility to the Oslo accords. Meanwhile Arafat's political support among Palestinians had been decisively weakened by his failure to secure an agreement. When President Clinton summoned Arafat and Israeli Prime Minister Ehud Barak to Camp David in 2000, Arafat was offered a deal similar to what he had been willing to accept five years earlier, but it was too late; Palestinian support for the deal had long since evaporated – as had support among Israelis.[14] Arafat was no longer in a position to secure Palestinian consent to anything Barak could have offered. Mandela and de Klerk both realized, despite their conflicts, that each needed to keep the other strong enough to be able to close the deal. Recognition of this strategic reality has been in chronically short supply in the Israeli-Palestinian conflict.

Rabin's fate demonstrates that taking political risks can literally be fatal. Most other leaders involved in the Israeli-Palestinian conflict have shown little appetite for the kind of strategically hopeful action that made the South African transition possible. Leaders on both sides (which, since their victory in the 2006 elections in Gaza and the West Bank have included Hamas) have not placed a high enough priority on resolving the conflict to build support for a new dispensation and take the risks that would be needed to advance its prospects for success. Indeed, both have countenanced policies that seem likely to worsen it. But further entrenching the status quo is not risk-free either. Leaders who refuse to take risks to recast festering conflicts in positive sum terms thereby increase the costs and dangers faced by someone else, somewhere else, sometime in the future.

---

[13] According to polls conducted by the Tami Steinmetz Center for Peace Research (TSC), the Oslo Peace Index of Israeli public opinion rose from 46.9 in October 1995 to 57.9 on November 8, immediately following Rabin's assassination. TSC, "Peace Index, 1995," http://spirit.tau.ac.il/socant/peace/  In October 1995, 72.5 percent of Palestinians polled supported the peace process. "JMCC Public Opinion Poll #10," http://www.jmcc.org/publicpoll/results/1995/no10htm

[14] Barak went out on a limb under strong pressure from President Clinton, offering new concessions on Jerusalem that infuriated many in the Knesset and subsequently cost Barak his premiership.

# VII. Conclusion

We have argued here that strategically hopeful action by leaders who are willing to take calculated risks is necessary for transforming political orders characterized by histories of divisive, violent, communal conflict. One vital characteristic of such leaders is their capacity, despite the bitter history, to view a conflict that is widely perceived to be zero-sum as potentially positive sum.

We do not claim that strategically hopeful leadership is sufficient; any number of factors outside a leader's control may cause even a carefully-calculated, risk-taking effort to fail. And sometimes, as we noted with respect to the Israel-Palestine conflict since the mid-1990s, the chances might be vanishingly slim. But we do argue that such action is necessary: without it the chronic conflict will remain beyond the possibility of resolution, even if the stars align in every other respect. Political leadership of a more routine character, whereby leaders respond "rationally" to the most visible and concrete forms of political reward and punishment, cannot bridge the abyss. On the contrary, all of the ordinary political incentives favor action reinforcing the status quo.

Grassroots efforts by civic or religious associations to foster inter-communal trust may be advantageous in preparing the ground for political efforts at the top, and in consolidating an agreement once its institutional arrangements have been delineated. And initiatives of this kind keep alive some glimmer of hope during long stretches of political darkness. But bottom-up action cannot by itself overcome the enormous collective action problems that stand in the way of negotiating, ratifying, and consolidating political support for new institutions acceptable to all sides. Only highly-placed political leaders, accountable to the groups engaged in conflict, can accomplish this, and only if they are willing to take significant risks of the kind we have described.

If our analysis here challenges leaders to take risks for peace, it also challenges social scientists and political theorists to look differently at leadership, interdependent decision, conflict, and power. Leadership theories that feature profit-maximizing elites who manipulate fixed preferences, or alternatively attribute extraordinary transformative powers to heroic individual leaders, miss vital dimensions of Mandela's and de Klerk's joint accomplishment. Theories of mixed-motive interdependent decision, exemplified here by Axelrod, Schelling, and Colomer, provide better starting-points for the problem at hand. But their theories also face limitations from failing to explore the dynamic possibilities opened by leaders who are willing to take unusual risks, in the hope that others will reciprocate rather than exploit them. Our understanding of political power would be deepened by attention to the peculiar blend of conflict and cooperation in power interactions of the sort that facilitated South Africa's relatively peaceful – and surprising – democratic transition.

When Nelson Mandela was released from prison in 1990, the world seemed to be blissfully proceeding toward resolution of many long-festering conflicts, not only in South Africa but also in Eastern Europe with the fall of communism, in Latin America, and elsewhere. At the time it may have seemed that Mandela, de Klerk, and South Africa achieved what they did by riding some "wave of history."  But the optimistic global mood of the early 1990s soon gave way to some ghastly nightmares, as in the former Yugoslavia at the same moment that Mandela and de Klerk were heading off disaster in South Africa. A cascading mood of optimism will not substitute for leaders who are willing and able to take risks at the critical junctures when these are required. In retrospect it is clear that Mandela and de Klerk did not merely ride a wave. They bet on one another when they had good reasons not to, displaying hope for the future that transformed the present and its possibilities. This enabled them to dismantle apartheid without destroying their country in the process, an achievement that commends itself to our attention today no less than it did in 1994.

## VIII.  References

Austin, John L. 1975. *How to do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton: Princeton University Press.

Baldwin, David A. 2002. "Power in International Relations" in Walter Carlsnaes, Thomas Risse, and Beth Simmons, eds., *Handbook of International Relations.* Thousand Oaks: Sage, 177-191.

Burns, James MacGregor. 1978. *Leadership*. Harper & Row.

Burns, James MacGregor. 2003. *Transforming Leadership: A New Pursuit of Happiness*. New York: Atlantic Monthly Press.

Colomer, Josep M. 2000. *Strategic Transitions: Game Theory and Democratization*. Baltimore: The Johns Hopkins Press.

Conger, Jay. 2004. "Transforming and Visionary Leadership." In George R. Goethals et. al. eds., *Encyclopedia of Leadership*, Vol. 4. London: Sage.

de Klerk, F.W. 1998. *The Last Trek – A New Beginning: The Autobiography*. London: Macmillan.

de Klerk, F.W. December 9, 2003. Interview with author, Cape Town, South Africa.

Dewan, Torun and David P. Myatt. 2012. *Journal of Theoretical Politics* 24: 431-460.

Dowding, Keith. 2011. "Introduction" to Dowding, ed., *Encyclopedia of Power*. London: Sage.

Elliott, Sydney and W.D Flackes. 1999. *Conflict in Northern Ireland: An Encyclopedia.* Belfast: Blackstaff Press.

Frohlich, Norman, Joe A. Oppenheimer and Oran R. Young. 1971. *Political Leadership and Collective Goods*. Princeton: Princeton University Press.

Gaventa, John. 1980. *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley.* Champaign: University of Illinois Press.

Giddens, Anthony. 1984. *The Constitution of Society.* Berkeley: University of California Press.

Gilpin, Robert. 1981. *War and Change in World Politics.* Cambridge: Cambridge University Press.

Hancock, Landon E. 2011. "There is No Alternative: Prospect Theory, the Yes Campaign and Selling the Good Friday Agreement." *Irish Political Studies* 26: 95-116.

Haugaard, Mark. 2012. "Rethinking the Four Dimensions of Power: Domination and Empowerment." *Journal of Political Power* 5: 33-54.

Hayes, Bernadette C. and Ian McAllister. 2001. "Who Voted for Peace? Public Support for the 1998 Northern Ireland Agreement." *Irish Political Studies* 16: 73-94.

Hobbes, Thomas. 1928. *Elements of Law*. Cambridge, UK: Cambridge University Press.

Hobbes, Thomas. 2010. *Leviathan*. Ed. Ian Shapiro. New Haven: Yale University Press.

Huntington, Samuel. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman, OK: University of Oklahoma Press.

Jung, Courtney and Ian Shapiro. 1996. "South Africa's Negotiated Transition." In Ian Shapiro, *Democracy's Place*. Ithaca: Cornell University Press.

Jung, Courtney, Ellen Lust-Okar, and Ian Shapiro. 2011. "Problems and Prospects for Democratic Settlements: South Africa as a Model for the Middle East and Northern Ireland?" In Ian Shapiro, *The Real World of Democratic Theory*. Princeton: Princeton University Press.

Kahneman, Daniel. 2011. *Thinking Fast, and Slow*. New York: Farrar, Straus & Giroux.

Lodge, Tom. 2006. *Mandela: A Critical Life*. Oxford: Oxford University Press.

Lukes, Steven. 2005. *Power: A Radical View*, 2nd ed. New York: Palgrave Macmillan.

Mandela, Nelson. 1994. *Long Walk to Freedom: the Autobiography of Nelson Mandela*. Boston: Little Brown.

Mandela, Nelson. 2004. *Nelson Mandela: In his own words*. Boston: Little Brown.

McGarry, John and Brendan O'Leary. 2009. "Power Shared after the Deaths of Thousands." In Rupert Taylor, ed., *Consociational Theory: McGarry and O'Leary and the Northern Ireland Conflict*, 15-84. London: Routledge.

McGlynn, Catherine, Jonathan Tonge and Jim McAuley. 2012. "The Party Politics of Post-devolution Identity in Northern Ireland." *The British Journal of Politics and International Relations*. Online Early View, September 2012.

McLaughlin, P.J. 2010. *John Hume and the Revision of Irish Nationalism*. Manchester, UK: Manchester University Press.

Mearsheimer, John. 2001. *The Tragedy of Great Power Politics.* New York: Norton.

Mitchell, George. 1999. *Making Peace*. New York: Knopf.

Morrow, James D. 1994. *Game Theory for Political Scientists*. Princeton: Princeton University Press.

Powell, Jonathan. 2008. *Great Hatred, Little Room: Making Peace in Northern Ireland.* London: Bodley Head.

Przeworski, Adam. 1991. *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America*. Cambridge, UK: Cambridge University Press.

Read, James H. 2012. "Is Power Zero-sum or Variable-sum? Old Arguments and New Beginnings." *Journal of Political Power* 5: 5-31.

Riker, William H. 1983. "Political Theory and the Art of Heresthetics." In Ada W. Finifter, ed., *Political Science: The State of the Discipline*. Washington, D.C.: American Political Science Association.

Riker, William H. 1986. *The Art of Political Manipulation*. New Haven: Yale University Press.

Schelling, Thomas C. 1980. *The Strategy of Conflict*. 2$^{nd}$ ed. Cambridge, MA: Harvard University Press.

Wrong, Dennis H. 1995. *Power: Its Forms, Bases, and Uses*. New Brunswick: Transaction Publishers.