Deliberation in Anarchy: Deliberative Trust-Building Outside the State

Jonathan Tomm
PhD Candidate
University of British Columbia
jonathantomm@alumni.ubc.ca

Work in progress—please consult the author before citing or distributing.
Comments welcome!

Presented at the annual meeting of the Canadian Political Science Association,
Victoria, BC, June 4–6, 2013.

I thank Mark Warren, Agustin Goenaga, Alfred Moore, Michael Morrell, Stewart Prest, and Chris Tenove for their helpful comments and insights on this paper.
1 Introduction

Political opponents who are entangled in violent and costly conflicts, and who wish to move toward negotiated settlements, face a difficult task. Where mutual coercion and strategic defection have become an established pattern, negotiating and implementing agreements requires a very significant change in the way opponents interact with each other. Successful negotiations require trust—the ability for people to count on each other’s verbal commitments, and to take those commitments to be reliable enough to act on. Without that basic trust, talks between political opponents are bound to go nowhere (Kelman 2005; Larson 1997).

Negotiations in the midst of violent conflict within or between states stand somewhat outside the usual province of deliberative democratic theory. In comparison with deliberation inside functioning democracies, negotiations between violent adversaries are more likely to be fraught with threats, bluffs, and deception—all forms of coercion that are antithetical to deliberation. Moreover, such conflicts are all-too-often marked by deep-seated animosity, contempt, and even hatred. Such attitudes, of course, are far removed from the relationships of reciprocity and mutual respect required for open and honest deliberation. Faced with such harsh realities, deliberative democratic theory might seem to be of limited relevance.

But despite this appearance, deliberative theory has much to contribute to our understanding of such negotiations. The problem faced by negotiators in difficult contexts—namely, to find ways of affirming and maintaining mutual commitments to shared goals or principles—has long been recognized as a competence of deliberation. What deliberative theorists call “collective will-formation” comprises not only shared commitments to particular goals or rationales, but also the relationships of mutual accountability that underpin the practice of giving and asking for reasons in the first place. Functioning democracies draw upon the relationship-building capacities of speech to generate the solidarity and mutual commitment needed for citizens to act together (Habermas 1996).

The trick, however, is to understand how deliberation can generate these mutual commitments outside the protection of a strong and legitimate state. In strong democracies, institutions do much of the work of guaranteeing deliberative commitments. Laws, backed up (if necessary) by force, reduce the burden on deliberation by aligning incentives to contain the risk of deception and strategic manipulation in speech. But in the relatively anarchical contexts of violent conflict, these guarantees are no longer in place.

The solution I propose is this: to look for how the function of stabilizing mutual commitments might be drawn from deliberation itself, even in the absence of protective democratic institutions. Deliberation, I argue, can generate reliable shared commitments that can temporarily stand in for the coercive institutions that normally underwrite cooperation in functioning democracies. By talking, political opponents can, in essence, pull themselves up by their own bootstraps. They can express and solidify small but

---

2 I thank Mark Warren for suggesting this way of putting it. See also Mitzen 2005, 404, 407ff.
meaningful commitments that, over time, feed into the development of institutions that later back up these nascent commitments with effective monitoring and sanctions.

The argument will unfold as follows. In part 2, I will point to the problem of trust-building in negotiations over the settlement of violent conflicts. Successful negotiators and mediators find ways to draw genuine mutual commitments out of talks even in relatively anarchical contexts, but we don’t yet have a good account of how they manage this, given the seeming ubiquity of coercion and the ever-present risk of defection.

In part 3, I will locate the problem in deliberative democratic theory. Deliberation within democratic states relies on the protection of strong institutions that can back up deliberative commitments, using force if necessary. But this reliance on institutions, appropriate in the context of a strong state, leaves deliberative theory at a loss when it comes to theorizing talk in more anarchical contexts.

In parts 4 and 5, I will argue that deliberative theory has the resources to address this problem, though they need to be clarified. Even negotiations in difficult conflicts, fraught with coercion as they are, can contain deliberative moments that have the effect of generating new, mutually acknowledged commitments. Talk itself, and not only the coercive capacities that sometimes stand behind it, can stabilize cooperation. In this way, deliberation can temporarily stand in for the kinds of institutions that political opponents must later build to further stabilize and develop productive interactions.

Finally, in part 6, I will summarize my conclusions and point toward the broader implications of this research.

2 Conflict resolution and the problem of trust

2.1 Risk and the need for trust

In drawn-out, violent (or potentially violent) conflicts, whether within or between states, a negotiated settlement is often the most attractive outcome. Violent conflicts are immensely costly, most obviously and devastatingly to the direct victims of violence, but also to the whole of societies whose resources are diverted toward supporting armies and police forces, and whose productive capacities are hampered by fundamental insecurity. Besides these costs, there are other ethical and political values that speak for the desirability of negotiations. To settle a conflict by words rather than violence is itself a value. It represents a commitment to mutual recognition, reciprocity, and human dignity—to the notion, integral to liberal societies but arguably universal, that social conflicts should be settled on a basis of mutual respect and compromise rather than blunt coercion. A preference for “the pacific settlement of disputes” is a strong and well-justified principle in international relations and diplomacy (United Nations 1945, chap. 6).

---

3 Like any generalization, this one admits of exceptions. In some cases, negotiated settlements might be, on the whole, more costly that violence, particularly if one side can win a decisive victory where a negotiated settlement would be only an unstable, and short-lived, truce (Toft 2010).

4 Here, too, there are possible exceptions to the generalization. Some conflicts may have to be fought to an uncompromising victory; the war against Nazi Germany seems like one such case.
But while negotiated settlements will often be the best we can hope for, they do not come easily. Since the Second World War, only about sixth of violent conflicts (both civil and interstate wars) ended in a negotiated peace agreement, though this number has gone up considerably since the end of the Cold War (Zartman 2008, 322–23; Babbitt and Hampson 2011, 47). More often, conflicts end with the victory of one side over the other, and some fizzle out on their own accord. Many continue to fester for decades, as the Israeli-Palestinian conflict has.

Perhaps the most common reason for the failure of negotiations is simply that one or both parties prefer to go on fighting. One of the main drivers of conflict is each party’s belief that it can get its way by further escalation: by stepping up the coercion applied to the other side. Conflicts become “ripe” for negotiation only when the parties find themselves in a “hurting stalemate,” that is, when they “feel that they can no longer expect to win the conflict through escalation (or simply holding out) at an acceptable cost” (Zartman 2008, 329). Conflicts might never become ripe, carrying on indefinitely or ending when one side finally manages to defeat the other outright.

But there is also an important set of conflicts in which ripeness occurs, but the combatants’ attempts to negotiate still founder. In these cases, both sides recognize the need to negotiate, and there are feasible settlements that both would prefer to ongoing fighting, but these settlements remain out of reach. That is, while parties get into conflicts because they are pursuing their interests, they can also remain stuck in conflicts “beyond their interests” (Zartman 2007, 15). This is, in part, what makes many protracted conflicts such tragedies.

The problem, in practical terms, is that negotiations require trust. Negotiations only work to the extent that the parties come to count on each other’s verbal commitments.\(^5\) To negotiate with an opponent invariably involves risk, and often considerable risk (Höglund and Svensson 2006). There is the risk of deception: that an opponent will make commitments only to renege later. There is the risk that negotiations are only a stalling tactic, a trick to buy time while an enemy regroups or consolidates its position. Even coming to the table involves risk: to talk the other party is to recognize it as a legitimate partner, which many adversaries are loath to do. To overcome these risks, parties have to have at least some degree of trust in each other—and trust is usually in very short supply.

Good negotiators and mediators will find ways to reduce the risks associated with negotiations so as to reduce the need for trust. They might do so by starting with small, easier issues where the risks are lower. Where possible, they will set up mechanisms of monitoring to assure each other. Best of all, perhaps, is if a strong third party can step in and place its own guarantee behind the conflicting parties’ commitments.

But these measures rarely eliminate the risk entirely. That is one of the basic elements of protracted conflicts that makes them so difficult to manage: there often is no third-party state or overarching institution that can guarantee the commitments conflicting parties make to each other. (If there were, the task of conflict resolution would

\(^5\) The rational-choice bargaining literature is a rich source of analysis on the difficulties of making credible commitments, and the related problems of private information and cheap talk (e.g. Fearon 1995; Walter 2009; Powell 2002). Broadly speaking, I interpret this literature as indicating the risks that parties have to overcome through trust-building if they are to cooperate. Bargaining failures occur when they are unable to do so.
but much easier.) And while dealing with smaller, less contentious issues is less risky, and for that reason a good place to start, at some point the more difficult and risky concessions will have to be made if the process is ever to advance beyond token talks.

2.2 Practical trust-building

We should not be surprised, then, that trust-building lies at the centre of many of the “best practices” of conflict resolution. For example, symbolic gestures of recognition and acknowledgment are often important steps in a peace process (Kelman 2005, 648; Peck 2008, 418). In many ways, recognition is the starting point for trust. Trust is appropriate among people who recognize each other as legitimate actors with rights and responsibilities. But this basic form of mutual recognition is conspicuously absent in many violent conflicts. When one party sees the other as mere terrorists, rebels, or traitors, mutual distrust is virtually assured.

A second, and related, principle of negotiation is to avoid challenging identities in ways that provoke reaction and closure. The role of identity in conflict and conflict resolution is admittedly complicated. Identities themselves can drive conflict, particularly when resisting a historical enemy becomes part of people’s membership in their own groups (Mitzen 2006). The long-term stability of a settlement likely depends in part upon oppositional aspects of identity eventually changing (O’Flynn 2006, chap. 3).

But especially in earlier stages of conflict resolution, when an initial settlement is being worked out, challenges to identity are often counterproductive. Identities are, by definition, a core aspect of personality, and thus threats to identity can provoke strong negative reactions. In order to question, examine, and possibly change values that lie close to this core, people must first feel secure. That is one reason why acts of mutual recognition are so important. It is only when people feel that their basic needs and identities are secure that they can venture the trust needed for serious negotiations (Farren and Mulvihill 2000, 122–24; Aiken 2010, 193–94).

Third, negotiators are often counselled to “focus on interests, not positions” (Fisher and Ury 1991). One reason for this principle is that exploring interests can lead to the discovery of win-win solutions. But trust-building provides another reason to focus on interests. In order to trust, conflicting parties need to know that each has a genuine interest, a stake, in peace (Kelman 2005, 646–47). If I know the settlement will be in my opponents’ interests (as well as mine, of course), that gives me reason to expect them to be serious about coming to an agreement, and maintaining cooperation. A clear-eyed understanding of interests contributes to trust.

Fourth, negotiators and mediators often use the negotiation process itself to foster trust (Curran, Sebenius, and Watkins 2004). Opponents’ confidence in a process can be a bridge by which they begin to build confidence in each other. A process might be defined by ground rules, broad goals, conditions for joining or leaving negotiations, linking certain issues together or, alternatively, keeping them apart, procedures for circulating drafts and working out troublesome language, and so on (Peck 2008, 421). Buying into, and upholding, such processes allows parties to gradually demonstrate their commitments and build the trust needed to make the necessary concessions on substance.

Fifth, getting parties to cooperate on concrete, low-risk projects can be an effective tool for building confidence and improving relations. Inter-group contact that
has a collaborative aspect—where participants are involved in accomplishing something together—appears to be more effective than talk for its own sake. As the director of an NGO working in Northern Ireland explains, “if you just get people together for the sake of having contact, I don’t think it’s terribly effective. The more you can engage people in joint projects or in some place where they’re negotiating some of the difficult issues around community safety, attacks on each other, or equality and poverty issues in the area the more you can build positive relations” (Aiken 2010, 190). The key here is that concrete, day-to-day cooperation allows conflicting parties to demonstrate their commitments, and thus to build trust.

Finally, negotiators and mediators are counselled to approach contentious issues in step with the success of trust-building (Kelman 2005, 644–45). Attempts to force discussion of an issue before there is enough trust and mutual confidence can lead to failure. Talking openly and, if necessary, making compromises on these issues presents the greatest risk to the parties involved. The greater the risk, the more trust is needed if the risk is to be overcome.

2.3 But how does trust-building work?

There is, thus, a considerable amount of practical wisdom built up around trust-building in conflict resolution. But our theories of how trust-building works lag somewhat behind practice, particularly when it comes to understanding the communicative and relational aspects of trust-building.

We know much, for example, about ways of minimizing the need for trust by arranging incentives (Babbitt and Hampson 2011, 47–50). There are rich literatures in repeated games, reputations, bargaining, and institutional design that examine the conditions under which rational, utility-maximizing actors will cooperate (e.g., Keohane 2005; Koremenos, Lipson, and Snidal 2001). Effective institutions, whether formal or informal, lower the risk of cooperation by threatening to impose sanctions on cheaters. And where cooperation is less risky, there is also less need for trust—and in low-trust environments like protracted conflicts that is a very good thing.

But minimizing the need for trust is not the same as building trust. Trust-building is something that happens in addition to the risk-management provided by sanctions and institutions. Trust comes in where institutions leave off, where despite the negotiators’ or interveners’ best efforts to incentivize cooperation, the possibility of defection remains. Trusting means stepping over this gap of uncertainty and choosing to cooperate based on a belief in the other party’s good faith (Luhmann 1979). There is, in other words, a moral element to trust—an element of mutual obligation and responsibility—that cannot be fully captured in theories of strategic interaction, power-sharing, and institutional design (Uslaner 2008; Mansbridge 1999).

Likewise, costly signalling theory captures some aspects of trust-building but misses others. Costly signals are no doubt important means of demonstrating commitments. But many of the things that are said and done in trust-building—the

---

6 On the connection between trust and risk, see Coleman 1990.
7 See Kydd 2005 for rational choice model of trust-building based on costly signals. See also Stein 1991, 441–44.
things said in off-the-record and exploratory talks, symbolic gestures of recognition, narratives that humanize the other party, the offering of new interpretations of a situation—do not always make sense as costly signals.

Socio-psychological theories of trust offer to capture emotional and identity-related aspects of trust-building that are missing from interest-based and strategic approaches. But here, too, the results are mixed. The main limitation is that the forces of habit, taken-for-granted meaning, and common identity that underpin trust in settled societies are of little help when it comes to open, violent conflict. One of the conditions of protracted conflicts is precisely that shared “lifeworld” certainties have broken up, become contested, or never existed in the first place (Warren 1996, 224; cf. Habermas 1996, 21–23). While there is little doubt that shared identities and norms help to stabilize cooperative social relations, that says little about how these identities and norms come to exist in the first place, particularly between opponents whose salient identities are largely defined in opposition to each other’s.  

3  Locating the problem in deliberative democratic theory

Deliberative democratic theory might not be an immediately obvious choice for investigating trust-building in violent conflicts. Deliberative democratic theory has been developed primarily with the democratic state in mind—a context that differs in important ways from the relatively violent and anarchical domain of most conflict resolution. Nonetheless, I will argue that this body of theory offers a rich conceptual resource for understanding conflict resolution. The first task, though, is to correctly locate the problem in deliberative democratic terms.

3.1  Collective will-formation

The promise of deliberative democratic theory for conflict resolution lies in its treatment of collective will-formation (Habermas 1996). Collective will-formation is a key competence of deliberation—really, the thing deliberation does best (Warren 2012). By “collective will-formation” I mean the process by which a number of actors—who have diverse interests, intentions, and judgments—come to endorse shared goals or principles to guide their actions.

What makes deliberation a form of collective will-formation par excellence are the inherent connections between deliberation, reasons, and will. Deliberation is best understood as the give and take of reasons. A reason is a claim about what is true, right, effective, legal, or otherwise appropriate—a statement, in other words, that claims to be valid or compelling in one way or another (Habermas 1998). Reasons are criticizable in terms of the validity they claim. The addressees of reasons can accept or reject the reasons they are offered, and their acceptances or rejections in turn stand in need of reasons. The influence that participants have on each other through deliberation comes through the force of those reasons—the way in which the people they are addressed to

---

8 Changes to identity seem more likely to come in as stabilizing factors later in a process of peace-building than to foster initial steps of cooperation. Wendt 1999, 342, 363; Kupchan 2010, 35.
find them binding or cognitively compelling. To find a practical reason compelling in this way, that is, to endorse it as right, is to will the action that the reason prescribes. Thus, by addressing reasons to each other, participants in deliberation can form a collective will: shared judgments about what ought to be done, and shared commitments to acting upon those judgments (Korsgaard 2009, 190).

3.2 Two dimensions of will-formation

It is important to notice, however, that collective will-formation has two dimensions. What I’ve just described is the first dimension, the dimension of reasons. Here, the question is about the content of public reasons, about what, in fact, is a good reason. This is what deliberation is ostensibly about. In deliberation, participants argue about the validity of particular reasons for action, demanding their justification, accepting or rejecting them.

But underlying these reasons is another dimension of collective will-formation that sometimes goes unnoticed. Here the question is not about the justifiability of this reason or that, but rather about the more fundamental question whether, and to what extent, the people deliberating are accountable to each other at all. This is the relationship dimension of collective will.\(^9\) The public reasons knocked about in deliberation are only valid if there is, in fact, a public: a group of actors who relate to each other on a basis of mutual accountability, actors who recognize and uphold obligations to each other. Reasons are creatures of relationships of mutual accountability. Those relationships form the “space of reasons” (Brandom 1994, 5) within which particular reasons have meaning. Thus, collective will-formation, and the deliberation by which it comes about, is always at the same time about both the particular reasons participants take to be good and the relationships of mutual commitment that underlie the practice of offering, accepting, and abiding by shared reasons in the first place.

Such relationships cannot be taken for granted, however, above all not in politics. Political conflicts are always overshadowed by the possibility that actors will relate to each other in purely instrumental ways. Talcott Parsons, drawing on Hobbes, puts the problem clearly: it is not only that people have conflicting goals; it is also that “the actions of men [are] potential means to each other’s ends” (Parsons 1968, 93. Cf. Hobbes 1994, part 1, chaps. 10–11, 13). In the worst conflicts, the relationship between the parties is fully instrumentalized (or nearly so). The conflicting parties do not recognize obligations to each other, only to themselves and their own constituencies.\(^10\) Collective will-formation requires the reversal of these instrumentalized relationships through the incremental recognition of mutual obligations. Much of the work of conflict resolution lies precisely here.

---

\(^9\) To use Habermas’s phrase, speech has a “double structure”: in addition to its propositional content, it also affects the intersubjective relationship between speaker and listener. Habermas 1998, 63–64. See also Mansbridge 2006, 108.

\(^10\) Fully instrumental interactions are a limit case. Even in outright war, combatants often recognize some obligations to each other. Walzer 1992.
3.3 The institutional response

Democracies (and, indeed, any effective system of government) must have ways of containing such instrumental defection. Law, backed up by coercive sanctions, is perhaps the most important means of doing so. Law, as Habermas argues, is Janus-faced (Habermas 1996, 25–39, 129–130). On the one hand, it makes a normative claim, by virtue of its legitimacy, on the citizen’s voluntary compliance. On the other hand, it also confronts its subjects as a demand backed up by coercion. The second, coercive face of law allows it to guarantee broad compliance. Even those who do not recognize the law as normatively binding will still comply in order avoid sanctions.

The key point is this: by structuring incentives to foster compliance, coercive institutions can stand in for the second, relational dimension of collective will-formation. Relationships of mutual accountability, based on a normative recognition of the other’s rights, do not have do the work of stabilizing cooperation alone (Habermas 1996, 37–38). As a result, the second dimension of collective will-formation can recede into the background. Within the context of a strong and legitimate state, individuals do not have to actively question and establish others’ commitments to shared rules. For the most part, people deliberating can expect each other to play by the rules and to abide by decisions that are made, and the threat of defection is relatively well contained.

3.4 Deliberation and anarchy

A difficulty shows up, however, when deliberative theory moves out of the democratic state into the more anarchical contexts of civil wars, weak states, and international relations. Here, the institutions that deliberative democrats rely on to back up deliberative commitments are not readily available. In the most difficult conflicts, especially where violence has taken place or looms in the offing, there often are no overarching institutions that can guarantee the commitments opponents might make to each other—or, if such institutions do exist, they are relatively weak.11 And it is not only the lack of coercive capacities. Also missing are the normative, habitual, and emotional ties that create an unspoken background of behavioural expectations—a lifeworld—in stable societies. When opponents no longer (or never did) see eye to eye on the normative expectations that regulate their common life, they confront each other as strangers—unfamiliar, and thus potentially untrustworthy, others.

As a result, the relational dimension of collective will-formation leaps back into the foreground. Political opponents must directly confront the question whether they can count on each other to observe mutual obligations or, to the contrary, they are engaged in a Hobbesian free-for-all. For people in relatively anarchical contexts, this is a live question in a way that (thankfully) it is not for people in strong, democratic states (Mitzen 2005).

But because deliberative democratic theory has been developed primarily with the assumption of a strong state in the background, it does not give us as much purchase on

---

11 While the United Nations, for example, has the authority to intervene in violent conflicts that pose a threat to international security (or, arguably, where there is an international responsibility to protect civilians), its ability to do so is depends upon powerful states, who are often reluctant to get involved.
anarchical situations as it might. Deliberative democrats tend to focus on the first dimension of collective will-formation: on the play of reasons in deliberation, and the way they come to form collective judgments. The question here is, What should we do? or, perhaps better, since deliberation is often not the best way to make decisions directly (Warren 2012, 11–13), What principles and goals should guide our decisions?

Deliberation has unique advantages as a way of answering that question. Epistemically, it allows participants to pool information, to question received wisdom, and to clarify preferences and trade-offs. Normatively, it embodies the autonomy and equality of participants, and thus grounds a presumption of legitimacy for those reasons that emerge as decisive (Habermas 1996, chap. 3). Thus, we can hope that deliberation (appropriately structured and facilitated, which is another important line of inquiry) will yield good answers: the well-considered judgments of free and equal citizens.

The problem, however, is that good answers and well-considered judgments are not the whole of collective will-formation. Agreeing about what ought to be done and actually doing it are not the same thing. The latter requires not only agreement, but the commitment to follow through on obligations to others.

This, too, is a competence of deliberation. In addition to its work in shaping reasons and opinions, deliberation also builds relationships of mutual commitment and trust. But because strong institutions reduce the need for deliberation to generate trust, deliberative democrats have, I think, paid somewhat less attention to this aspect of deliberation.

4 Deliberation as trust-building

Deliberative theory can, I think, respond to this challenge. Through deliberation, conflicting parties can make and demonstrate commitments to each other in a way that builds trust. What we need, then, is a thicker account of how deliberation does this relationship-building work. In particular, we need an account that is attuned to the challenges posed by relatively anarchical contexts.

4.1 Pushing out the edges of social order

Broadly speaking, conflicting parties who want to negotiate but face a deficit of trust need to do two things. First, they need to accurately assess each other’s commitment to upholding shared norms—in other words, each other’s trustworthiness. The notion of trustworthiness is captured in the relational dimension of collective will-formation. The question is this: To what extent does my opponent consider himself bound by obligations to me? Or, in more directly relational terms, to what extent are we relating to each other on a basis of mutual accountability? Negotiators need to answer this question well if they are to make good decisions about what risks, if any, they should take in the course of negotiations.

The second task is to actively build trust. Good negotiators are concerned not only with accurately assessing their situations but also with changing them. The goal is to shift their interactions with their opponents onto a normative footing—to generate and solidify
a relationship of mutual accountability with the other, while at the same time staying in step with their perceptions of the other’s trustworthiness (Kelman 2005).

These negotiations, in effect, take place at the edges of social order. In relational terms, this is a twilight zone, where cooperation in accordance with shared commitments can neither be assumed (as it can in more settled social and institutional contexts) nor ruled out. If negotiations are to be successful, the parties much find a way to push those edges outward, expanding the space of reasons to encompass new interactions that previously stood outside it.

4.2 The relational structure of deliberation

How can deliberation help? Answers to that question can, I think, be explored at different levels of generality. Here I want to begin at a more general level before moving, in the next section, toward a middle-range account of deliberative trust-building potentials.

At the more general level, the claim is this: deliberation builds trust because it enacts and solidifies a relationship of mutual accountability between speakers.

What gives speech its persuasive force is a presumption about how the speaker is using words (Habermas 1998; Habermas 1984, chap. 3). The listener is convinced if, and only if, she believes that the speaker is making claims properly, according to standards of valid inference. If she believes that the speaker knows what he is talking about and is speaking truthfully, she might be persuaded. But if she cannot make these assumptions—if she suspects that the speaker has been biased or sloppy in forming conclusions, is unreliable, or is engaged in outright manipulation—communication will go nowhere.

These standards of valid language use are, in effect, obligations that speakers and listeners owe to each other (Brandom 1994). Speakers and listeners are accountable to each other for the way they use words. In offering a reason or justification, the speaker undertakes responsibility for its correctness (Brandom 1994, chap. 3). In doing so, he effectively licences the listener to challenge the claim if it does not seem correct. But the listener is accountable to the speaker as well. While the speaker is responsible for making justifiable claims, the listener is responsible for accepting justifiable claims. If she does not accept the claim, she ought to be able to say why. She cannot reject it on a whim, or merely out of antipathy toward the speaker, or because it is inconvenient (Habermas 1987, 73).

The upshot is that to engage in an exchange of reasons is necessarily to place yourself in a relationship of mutual accountability with another person. Mutual accountability is a presupposition of the activity of communication, it is built into the very conception of what communication is. The activity of offering or asking for reasons only makes sense in if participants can think of each other in a certain way—namely as responsible users of language, accountable for the validity of the claims they make (Brandom 1999; Erman 2010).

There are a couple of points to bear in mind about these “idealizing presuppositions” of communication (Habermas 1996, 4).

First, these presuppositions are not factual claims. Rather, they are features of an attitude or stance that speakers must take toward each other in order to engage in the activity of deliberation (Habermas 1984, 328ff; Baynes 2007). Deliberation as an activity is premised upon a relationship of reciprocal accountability between participants. This is,
in effect, a transcendental argument. It is a claim about the conditions of possibility for communication on the basis of reasons. We could not make sense of the practice of communication or the experience of being persuaded if we did not think of ourselves and others as responsible users of language.

Concrete instances of reason-giving will, of course, conform to varying degrees to these ideals. In protracted conflicts, arguments are often recited in a ritual way, as markers of identity or allegiance rather than serious attempts to persuade. Reasons can be used to obfuscate and mislead. They can be deployed in ways that directly undermine the assumption of mutual accountability.

Thus, the relationship-building logic inherent to reason-giving should be taken as an argument about how reason-giving works when it works. These presuppositions do not describe communication per se. They describe productive communication: communication that is relatively close to the ideal of reason-giving oriented to mutual understanding and, for that reason, is capable of generating new shared meaning (Habermas 1984, 288, 331). Not all communication is productive in this way—perhaps relatively little is. But in negotiations between deeply divided antagonists, success depends upon some productive communication occurring. The point is that when it occurs, it generates and solidifies the relationships of mutual accountability that are inherent to the activity of persuasion.

Second, “reasons” should not be given a restrictive interpretation. While early formulations of deliberative theory tended to emphasize formal argumentation, many theorists now recognize that a wider range of forms of communication—including rhetoric, story-telling, expressions of emotion, and symbolic protest—can also serve as reasons (e.g. Young 2000; Dryzek 2000; Chambers 2009; Neblo 2007). A reason is simply a valid justification, an answer to a “why” question that a person endorses as, in one sense or another, right or compelling (Mansbridge 2007, 261). That is, “reason” is a formal concept; it points to the sense of conviction or endorsement that good reasons carry, and not to their particular content (Korsgaard 2008; Kant 1997, 27). Emotional expressions, stories, and other forms of less formally argumentative communication often have an implicit justificatory purpose. In this, they have the same potential as more formal argumentation to bring about rationally motivated agreement. And—more to the point here—to the extent that they are justifications directed from one person toward another, they can also do the relationship-building work of demonstrating mutual accountability.

4.3 Deliberative tasks and accomplishments

The internal connections between reason-giving and relationships of mutual accountability provide, I think, essential conceptual groundwork for understanding

---

12 The idea here is analogous to (and, indeed, a direct descendent of) Kant’s conception of moral freedom. As Kant argued, the activity of making moral judgments requires that we think of ourselves and others as exercising free choice, because only then could we be responsible for what we do. But freedom, however, is not a factual claim. Indeed, according to Kant’s mechanistic view of the physical universe, free will was not possible; human actions, like all events, are causally determined. The notion of freedom is, rather, a presupposition of the practical use of reason. Kant 1997, 26–27, 42, 79–84; Baynes 2007, 74; Korsgaard 1996, chap. 7; cf. Heath 2003.
deliberative trust-building. But this groundwork then has to be worked out into “middle level” accounts of more specific deliberative tasks and accomplishments.

What follows is a list of a few such accomplishments—ways that deliberation helps to build and solidify mutual commitments. The mechanisms I am pointing to here are not original. Many of the relationship-building capacities of deliberation have been well-theorized, and others are clearly embedded in the practical wisdom of conflict resolution and negotiation. The contribution of this paper is to articulate more clearly the practical problem to which various forms of deliberation are an answer. Negotiators in relatively anarchical contexts face the twin tasks of building trust and solidifying mutual commitments, while at the same time assessing the trustworthiness of each other. If we take the perspective of actors in this situation, certain deliberative accomplishments become salient.

Managing the effects of disagreement and repairing breakdown. One of the most important things deliberation can do in political conflicts is to mitigate the relational damage done by disappointed expectations. Political conflicts occur where standards of right and wrong are hotly contested. As a result, actions that one party takes to be justified are often viewed by the other as not only wrong, but obviously and culpably wrong—as yet more evidence of an enemy’s treachery or depravity. The crucial distinction here is between dissent, a good-faith disagreement about what, in fact, we owe each other, and deviance, the deliberate disregard for obligations to each other per se (Heath 2001, 152). In a climate of suspicion and even hatred, dissent can be mistaken for deviance, thus triggering a cycle of self-confirming distrust (Jervis 1976).

Deliberation can help to break this cycle. By offering reasons for their actions, conflicting parties present themselves as accountable to each other; they present their actions as the result of dissent, not deviance. Likewise, by demanding explanations for offending actions, they concede that a legitimate explanation is at least possible—that their opponents might not, after all, be outright scoundrels and not beyond reasoning with.

That does not mean, of course, that deliberation will result in consensus. In most cases, where disagreements and recriminations run deep, it probably will not. But deliberation does not have to achieve consensus in order to change the way political opponents interpret each other’s actions. It may be enough if they succeed in demonstrating to each other that their positions could be taken in good faith. Deliberation can work by fostering moments of perspective-taking that humanize the participants and make sense of their actions (Steiner 2012, 72–74). The act of reason-giving itself demonstrates a willingness to be accountable to others, and this can have trust-building effects even when the reasons themselves fall short of persuading all the parties involved.

Recognition through reason-giving. The give and take of reasons in deliberation builds trust through moments of recognition. Reasons are always freighted with recognitions. When one party offers its opponent a reason for the position or action it takes, it implicitly recognizes the addressee as a party with moral standing, as someone to whom it is accountable. Such recognition is not insignificant in deep-seated conflicts. Conversely, one of the surest ways to convey contempt or misrecognition is to refuse to
justify your actions or negotiating positions, or to offer a justification that is patently unacceptable to others.

In addition to these implicit recognitions, deliberation often involves explicit gestures of recognition. Rituals of politeness and respect serve this function, as can such details as the shape of a negotiating table or the order of speaking (Young 2000, 57–62; Aurisch 1989). There is a distinct deliberative rationale that underpins such gestures. Because the practice of giving and asking for reasons implies a relationship of mutual accountability and reciprocity, where that relationship is uncertain—and particularly where one or more parties feels that its status as a moral equal is shaky—it needs to be solidified. Otherwise, as Mark Warren (2006, 163) argues, “the who of the speakers undermines the what of statements, such that the speech loses its forcefulness as a means of resolving conflicts.”

Deliberation thus relies upon and enacts recognitions that reaffirm the relationship of mutual accountability.

Building a body of public reasons. One important function of deliberation is to generate a set of shared reasons or understandings that both parties can subscribe to in a principled way. John Rawls argues convincingly that the achievement of such a body of public reasons is essential to liberal democratic states (Rawls 2005). But it is all the more important in conflict where people are divided by fundamental questions of identity or sovereignty. Here, the sort of public reasons that give a moral foundation to liberal democratic states are missing, and need to be developed if the conflicting parties are to succeed in building shared institutions (O’Flynn 2006; Dryzek 2006, chap. 3).

There are a number of ways the deliberative process of reason-giving can contribute to a body of shared reasons or understandings (For a good summary, see Mansbridge 2009, 11–20). First, one party might simply be able to persuade another, or, through the give and take of argument, the parties might arrive at a win-win solution that both can subscribe to without reservation. Full-fledged consensus, if achievable, provides a stable basis for cooperation and is inherently legitimate. But in many (perhaps all) serious conflicts, such consensus will probably remain out of reach, at least on the most contentious questions.13

Second, deliberation can result in “incompletely theorized agreements” (Sunstein 1995 quoted in Mansbridge 2009, 12–15) or “overlapping consensus” (Rawls 2005). Here, conflicting parties come to agree on concrete actions or more general principles, but they do so for different reasons—or, where they do share reasons, those reasons do not go all the way down to first principles or worldviews.

Third, parties can come to agreements because they represent what Rawls calls “fair terms of social cooperation” (Rawls 2005, xlii). In this case, the parties accept the agreement as a compromise. Where consensus or incompletely theorized agreements are achieved, both parties can accept the agreement on its own substantive merits. The goals, principles, and interests it embodies are their own goals, principles, and interests. The same cannot be said of a principled compromise. Each party finds that some of the

13 Indeed, when parties have conflicting interests, we should probably view a consensus with some suspicion. Consensus can be a sign that conflicts were papered over rather than fully dealt with, or that some interests were excluded or suppressed. Mansbridge 2006; Bächtiger 2012.
principles embodied in the agreement are not its own. The parties accept the agreement, instead, because it has a “second order” validity: it enables a form of cooperation that upholds values of equality, fairness, and mutual respect (Habermas 1996, 165–67; Neblo 2007, 539; Mansbridge et al. 2009, 9–17).

Such compromises express a form of mutual recognition that makes them morally compelling. The notion of fairness or reciprocity that underpins the bargain gives it a relational significance: such compromises are, in part, what the parties owe to each other. This distinguishes them from the sort of modus vivendi that results from a purely instrumental bargaining process with no deliberative, that is, reason-giving, content (Rawls 2005, 145–47).

5 Deliberation and coercion in conflict resolution

Deliberation, then, has important relationship-building capacities. But to successfully apply deliberative theory to adversarial negotiations in conflict situations means navigating some difficult conceptual waters. On the surface of it, the enterprise might seem naive. Negotiations in violent conflicts are a rough business, suffused with coercion, hard-nosed bargaining, and strategic manoeuvres—a far cry from deliberative ideals. But deliberation does, in fact, play a significant—even necessary—role in successful conflict resolution. In order to see how, though, we have to be clear about what precisely that role is and how it relates to other, non-deliberative elements of conflict resolution.

5.1 Deliberative moments in adversarial negotiations

The task of applying deliberative theory to such negotiations is not, on the surface of it, an easy one. To propose a role for deliberation in conflict resolution immediately raises the problem of the incompatibility between deliberation and coercion. Deliberation is antithetical to coercion. Deliberative influence is exercised through the “forceless force of plausible reasons,” through claims that motivate because they are valid (right, true, or otherwise cognitively compelling) and not through deceit or the threat of force (Habermas 1996, 24; Mansbridge et al. 2009).

Negotiations in conflict situations, however, are suffused with coercion. Negotiations often involve threats, including the threat, made explicitly or implicitly by both sides, to walk away if enough of their demands are not met. Both sides seek to bring whatever power they have to bear on influencing the shape of the settlement. Likewise, deceit and misrepresentation are widespread. In contentious negotiations, both parties face the standing temptation to misrepresent their core, non-negotiable interests as being more extensive than they really are, and to overstate the attractiveness of their outside options.

Given the ubiquity of coercion in conflicts like these, deliberative democrats might be wary of applying the concept of deliberation. The risk is twofold. On the one hand, if we accept the inevitability of coercion in negotiations, we risk watering down the normative ideal and conceptual coherence of deliberation (Bächtiger et al. 2010). On the other hand, if we insist that negotiations be conducted according to a strict deliberative
Deliberation in Anarchy

Thus, it is important to be clear about how I mean to apply the notion of deliberation in conflict resolution. It is not that conflicting parties (and, for that matter, third-party mediators or interveners) will interact exclusively through deliberation. Rather, it is matter of identifying deliberative moments and deliberative effects in negotiations that will also involve coercion. We can do so without either eroding the distinctiveness of deliberation as a mode of influence, or whitewashing the real and inevitable place of coercion in conflict resolution.

Deliberative moments occur whenever conflicting parties offer reasons to each other in a way that relies upon the validity of the claim to have an effect. Such moments are not, I think, rare in negotiations. Negotiators often offer justifications of their positions, demand justifications of others, and accept or reject positions on the basis of whether they are right, fair, efficient, or otherwise appropriate (Risse 2000; Crawford 2009).

Admittedly, recognizing a deliberative moment for what it is can be tricky. The problem lies in distinguishing deliberation from coercion in the form of threats or deceit. Some “reasons” offered in negotiations are not so much reasons as threats: if they work, they work because the receiver is forced to adjust to the threatened harm and not because the receiver recognizes the claim as normatively valid. The same goes for instances of deceit or misrepresentation. Here again, it is not the validity of the claim that moves the hearer, but rather the force the speaker exerts over the hearer through deceit (Mansbridge et al. 2009, 18; Warren 2011, 12–13 citing Bok 1978).

Furthermore, it is possible for valid reasons to be offered without deliberative intent: the speaker presenting the reason believes it to be a compelling one, but he offers it only because it happens to be in his interests, and not because he is actually open to persuasion or questioning (Bächtiger 2012; Bächtiger et al. 2010). But this, too, is a form of deceit because the person offering such a reason can only hope to influence the listener if the listener believes he is serious about talking. To offer a reason is (except in contexts where everyone know it is a purely pro forma gesture) to represent yourself as accountable in terms of reasons, and in this respect, the speaker is deceiving his interlocutors.

But difficult as it may be to distinguish deliberative moments from seeming deliberation, the problem is not insurmountable. In fact, the problem is resolved, or at least managed, in practice every day. Assessing each other’s intentions and responding appropriately—which usually means responding differently—to sincerely offered reasons, disguised threats, and misrepresentations is something that skilled negotiators are good at. (Arguably, intuining each other’s intentions is something that any well-socialized human being is reasonably good at.) That is not to deny the challenges of identifying and operationalizing deliberation in empirical research (Bächtiger et al. 2010; Deitelhoff and Müller 2005). But the fact that deception and coercion can be mistaken for deliberation doesn’t mean that reliable means of distinguishing them can’t be discovered.
5.2 What deliberation does...

What do these deliberative moments do? Ultimately, the accomplishment of deliberation in conflict resolution is the same as in democracy. Deliberation is a way of forming collective will. A successful process of negotiation results in a set of bargains and understandings that the conflicting parties can commit themselves to. That aspect of commitment is what points to the deliberative moments in negotiation. For conflicting parties to commit to a settlement (which, in any case, is important if it is to be a stable one), they must see it, at some level, as a good one—as a settlement they have good reasons to uphold or, in other words, to will. To the extent that those reasons are shared they are most likely deliberative accomplishments. It is through moments of deliberation—in which reasons and justifications are offered and demanded, accepted and rejected—that conflicting parties can slowly build a body of shared commitments and common understandings.

But in this paper I have been focusing, of course, on another set of deliberative accomplishments: the relationships of mutual commitment that underpin cooperation. Much of the work of deliberation lies here—particularly in conflicts where institutional protections are weak, and thus where the parties have no choice but to rely on each other’s commitments. In these circumstances, talking is never just about the content of the words, the substance of the discussion; it’s also about the people talking, and about the relationships between those people. In giving and asking for reasons, conflicting parties are enacting and solidifying relationships of mutual accountability. They are building trust.

5.3 ...and what it doesn’t

It is also important, however, to be clear about what deliberation does not do, and thus to avoid overloading deliberation beyond its areas of strength (Warren 2012). Deliberation is not required to do the work of conflict resolution alone.

To begin with, the concept of “ripeness” applies. A conflict is ripe for more-or-less deliberative negotiations only when the parties find themselves in a “hurting stalemate,” and such stalemates are the result of a balance of coercive power. Indeed, skillful mediators may need to increase the coercive power of a weaker party, or apply coercive power themselves, in order to bring about the needed stalemate (Zartman 2007, 19). But balancing power and applying pressure are not the work of deliberation itself.

Once a conflict is ripe for negotiations, deliberation requires further help from institutions and coercive capabilities that help reduce the risks involved in negotiating and abiding by a settlement. These include the parties’ own abilities to defend themselves from, and punish, each other’s opportunistic defection. Risk reduction may also take place through third-party mechanisms of monitoring and sanctioning, such as human rights NGOs, United Nations peacekeepers, or interested states. As I have emphasized, such institutions can rarely do all the work of stabilizing cooperation—hence the need for

---

14 Paradoxically, more-or-less deliberative negotiations can be thwarted by a lack of coercion. This occurs when efforts to reduce violence also make the conflict less costly to the parties involve, and thus lower their incentives to negotiate a settlement. This is one of the dilemmas of conflict management. Zartman 2007, 18; Babbitt and Hampson 2011, 52–53.
trust, and deliberation that builds trust. But that does not mean that institutions and incentives do nothing. On the contrary, the more they contain the risk of defection and deceit, the lighter the trust-building burden deliberation carries, and the more easily conflicting parties can find their way through to productive agreements.

Finally, I must emphasize the symbiotic relationship between deliberative trust-building and institutions. As deliberation generates trust, conflicting parties become increasingly capable of working together to build and maintain institutions. Trust is often needed to overcome the risk of negotiating and implementing new institutions, but once institutions are up and running, they back up those initial commitments with effective monitoring and sanctions.

Deliberative trust-building, in other words, is not meant to be a permanent replacement for coercive institutions. What we should look for, instead, is for talking to provide a source of mutual commitment that can give conflicting parties the initial trust they need to get the cycle of cooperation and institution-building going.

6 Conclusion: Deliberation in the worst circumstances

Negotiations in violent conflicts are undoubtedly a tough case for deliberation, so much so that to even talk of deliberation in such circumstances might seem utopian. But it is not. While deliberative theory has been developed primarily within the setting of the democratic state, it also has the resources to uncover how collective will can be generated outside the protections of strong and legitimate institutions. Indeed, it is through deliberation—through the give and take of reasons—that conflicting parties develop the mutual commitments needed to begin the work of building and maintaining shared institutions.

References


