

Investigating diplomatic transformations

NICHOLAS J. WHEELER*

How does interstate conflict become transformed into cooperation? Many scholars regard this as a question not even worth exploring, on the basis that the dark shadow cast by international anarchy and the security dilemma dooms experiments in sustained security cooperation to failure.¹ This article rejects such a fatalist prognosis, but, in doing so, it acknowledges the importance of the security dilemma as an obstacle to cooperation. The key phrase in my article—*diplomatic transformations*—is designed to open up the question as to whether the practice of diplomacy, which Martin Wight classically defined as the ‘art of communication’,² holds an important key to mitigating, and indeed escaping, the security competition which can be generated by the security dilemma.

Wight believed that diplomacy could mitigate a conflict between two adversaries, but not transform it. The starting point for this article is the proposition that Wight had too timid a conception of the possibilities of diplomacy. As Wight put it in the chapter on war in the 1978 edition of *Power politics*, diplomacy could ‘circumvent the occasions of war, and ... extend the series of circumvented occasions ... but the causes of war ... will remain so long as a multiplicity of governments are not reduced to one government and international politics transformed into domestic politics’.³ Diplomacy for Wight was a practice undertaken by diplomats, and he recognized the contribution which human agency and the qualities of empathy, mutuality and reciprocity made to sustaining the society of states. But what remained at best submerged in his account was an appreciation that these qualities were also crucially important at the level of top policy-makers. This level is crucial because, given the nature of domestic political hierarchies, it

* This is a revised version of the 38th Martin Wight Memorial Lecture, delivered at Chatham House on 8 November 2012. I am grateful to all those who shared their time so generously and for the insights they gave me into Martin Wight’s intellectual universe. I especially want to thank Robert Ayson, Ken Booth, Ian Clark, Tim Dunne, Ian Hall, Jan Ruzicka and Silviya Lechner for their contributions. I benefited greatly from the comments provided by Will Bain, Andrew Linklater and Christian Reus-Smit in developing the article version; I also want to thank Ken Booth, Laura Considine, Tim Dunne, Anne Harris, Dani Nedal, Jan Ruzicka, Mark Webber and especially Justin Morris for their insightful comments on earlier versions. I would also like to thank Josh Baker for his research assistance and comments on the article.

¹ Kenneth N. Waltz, *Theory of international politics* (New York: McGraw Hill, 1979); John J. Mearsheimer, *The tragedy of Great Power politics* (New York: Norton, 2001).

² Martin Wight, *Power politics* (Leicester: Leicester University Press, 1978), p. 137.

³ Wight, *Power politics*, p. 138.

is only here that the decisions can be taken which have the potential to transform conflicts into something more positive.

Wight did explore, in one of his little-studied international theory lectures,⁴ the possibilities for two adversaries ending a conflict, but he struggled with the question of how the critical moves that might de-escalate a conflict can come about. He ended up suggesting that one side needed to take the steps to 'create confidence on the other side'.⁵ But this begged the crucial question as to whether one side would move in this way in the absence of trust. It is my contention that Wight never squared this circle, and in particular that he did not explicitly discuss the concept of trust and the role it could play in transforming adversarial relationships. This article contends that a focus on trust—and, crucially, how it can be developed between leaders—might help both scholars and practitioners to better understand how conflicts become transformed in adversarial contexts.

Trust between policy-makers, and their counterparts in an adversary state, is critical to the kind of moves that can transform a conflict. But it is important to recognize that leaders might communicate their willingness to make a cooperative move—or, as happens more often, respond to a conciliatory move by an opponent—through meetings with emissaries from the enemy state. At the same time, assessments at the top level of the potential trustworthiness of a current adversary might be importantly shaped by the interactions between diplomats on both sides at the lower levels of the bureaucracy. This is the key link between Wight's focus on diplomacy as the practice of diplomats and what Winston Churchill, referring to the 1953 Geneva summit (a term coined by Churchill) meeting between the leaders of the United Kingdom, the United States, the Soviet Union and France, called 'diplomacy at the highest level'.⁶

I argue below that a key precondition for the growth of trust at the 'highest level' of diplomacy is a capacity to empathize with the other side's security fears and motivations—what Ken Booth and I have called 'security dilemma sensibility' (a concept to be explored more fully later).⁷ Decision-makers cannot be sure about the military plans of an adversary—or potential adversary—because of the security dilemma: the existential condition of uncertainty that confronts actors as to the motives and intentions of others with the capabilities to do them harm.⁸ The challenge facing actors in an adversarial relationship is to decide whether their conflict has been generated by what Robert Jervis called the 'spiral model' of international conflict, in which actors mistakenly impute malign motives to actions that were taken for defensive purposes, thereby creating a conflict or

⁴ Martin Wight, *International theory: the three traditions*, edited by Gabriele Wight and Brian Porter (Leicester: Leicester University Press, 1991), pp. 180–206.

⁵ Wight, *International theory*, p. 185.

⁶ David Reynolds, *Summits: six meetings that shaped the twentieth century* (London: Penguin, 2007), ch. 1; David H. Dunn, ed., *Diplomacy at the highest level: the evolution of international summitry* (Basingstoke: Palgrave Macmillan, 1996).

⁷ Ken Booth and Nicholas J. Wheeler, *The security dilemma: fear, cooperation and trust in world politics* (Basingstoke: Palgrave Macmillan, 2008), p. 7.

⁸ This definition is developed in Booth and Wheeler, *The security dilemma*, pp. 4–5.

escalating it to higher levels than might have been the case had actors appreciated that they were trapped in such a spiral.⁹

The crux of my argument (and the wider project from which the article draws)¹⁰ is that communication between top-level policy-makers—especially face-to-face meetings—has the potential to transform conflicts between adversaries because it makes possible the growth of trust between these key actors. The mechanisms by which diplomatic encounters of this kind might generate trust have not been systematically explored in the literature on interstate security cooperation,¹¹ and this article seeks to contribute to rectifying that omission by showing the importance of this communicative dimension in understanding the transformation of the US–Soviet conflict in the second half of the 1980s.

Trust, by which I mean *the expectation of no harm in contexts where betrayal is always a possibility* (about which I will say more later), is vitally important to processes of diplomatic transformation, because without it actors cannot be expected to make the kind of ‘frame-breaking conciliatory’¹² moves that might de-escalate a conflict.¹³

⁹ Robert Jervis, *Perception and misperception in international politics* (Princeton, NJ: Princeton University Press, 1976), pp. 58–113. Jervis contrasted the spiral model with what he called the ‘deterrence model’. In the latter, decision-makers assume that they are dealing with an aggressor state, and the policy prescription is to maximize their deterrent capabilities and avoid showing any sign of weakness or lack of resolve (Jervis, *Perception and misperception*, pp. 58–62).

¹⁰ Nicholas J. Wheeler, ‘The challenges to trust building in nuclear worlds’, ESRC/AHRC-funded project under RCUK’s Global Uncertainties Programme: Security for All in a Changing World, <http://www.birmingham.ac.uk/research/activity/conflict-cooperation-security/trust-building-conflict-transformation/index.aspx> (accessed 20 Feb. 2013). Although this article focuses on the role that Reagan and Gorbachev’s summitry played in the de-escalation of the Cold War, the wider project considers two other key case-studies where the role of trust in promoting cooperation is explored. These are the India–Pakistan nuclear conflict, with a particular focus on the origins, dynamics and demise of the 1999 Lahore peace process; and the conflict between the United States and Iran over its nuclear programme in the 2000s. For treatments of aspects of these cases from the perspective of trust, see Nicholas J. Wheeler, ‘“I had gone to Lahore with a message of goodwill but in return we got Kargil”: the promise and perils of “leaps of trust” in India–Pakistan relations’, in Bhumitra Chakma, ed., *The politics of nuclear weapons in South Asia* (Farnham: Ashgate, 2011), pp. 155–76; Nicholas Wheeler and Dani Nedal, ‘Iranian nuclear negotiations: a long way from trust’, <http://www.birmingham.ac.uk/research/activity/conflict-cooperation-security/news/2012/07/Iranian-Nuclear-Negotiations-A-Long-Way-from-Trust.aspx>, accessed 25 Feb. 2013; Josh Baker, Scott Lucas and Nicholas Wheeler, ‘Breaking the deadlock in the Iranian nuclear negotiations’, <http://www.birmingham.ac.uk/research/activity/conflict-cooperation-security/news/2012/12/Breaking-the-Deadlock-in-the-Iranian-Nuclear-Negotiations---Josh-Baker-Scott-Lucas-and-Nicholas-J-Wheeler-.aspx>, accessed 25 Feb. 2013.

¹¹ Some scholars have recognized the importance of the communicative dimension in understanding how interstate conflicts become transformed. See K. M. Fierke, *Changing games, changing strategies: critical investigations in security* (Manchester: Manchester University Press, 1998), and *Diplomatic interventions: conflict and change in a globalizing world* (Basingstoke: Palgrave Macmillan, 2005); Jennifer Mitzen, ‘Reading Habermas in anarchy: multilateral diplomacy and global public spheres’, *American Political Science Review* 99: 3, 2005, pp. 401–17; Neta C. Crawford, *Argument and change in world politics: ethics, decolonization, and humanitarian intervention* (Cambridge: Cambridge University Press, 2002). The communicative dimension and its impact on the possibilities for trust and mistrust is becoming a central theme for some trust researchers in International Relations. See Booth and Wheeler, *The security dilemma*, pp. 155–8, 166, 237–8, 247–8; Harmonie Toros, *Terrorism, talking and transformation: a critical approach* (London: Routledge, 2012), esp. pp. 42–65; Naomi Head, *Justifying violence: communicative ethics and the use of force in Kosovo* (Manchester: Manchester University Press, 2012), and ‘Transforming conflict: trust, empathy, and dialogue’, *International Journal of Peace Studies* 17: 1, 2013, pp. 33–55 (forthcoming).

¹² I am grateful to Roderick Kramer for suggesting this formulation.

¹³ Brian Rathbun also explains international cooperation, and in his case-studies the growth of international organizations, in terms of trust as the antecedent of cooperation. However, he explains the growth of trust in terms of leaders operating with a ‘generalized model’ of trust as against my focus on the specific contextual possibilities of summitry and other communicative encounters. See Brian Rathbun, *Trust in international cooperation: the creation of international security institutions and the domestic politics of American multilateralism* (Cambridge: Cambridge University Press, 2012).

For the purposes of this article, I am defining a *diplomatic transformation* as a process in which two adversaries go through a series of steps of de-escalation which progressively reduce the role that the threat or the use of force plays in their relationship. Such a process, if sustained, might eventually lead to the growth of what Karl Deutsch and his co-researchers called a 'security community', where the threat or use of force has become de-legitimized as an instrument of state policy among members.¹⁴ A process of diplomatic transformation is characterized by the following indicators: key decision-makers no longer impute malevolent motives and intentions to an adversary (now in the process of becoming a former adversary); decision-makers on both sides recognize the role that their past actions have played in making the other side fearful and insecure; and there is reciprocation (positive or negative)¹⁵ of any cooperative moves that promote mutual reassurance.

This discussion of trust in achieving diplomatic transformations raises a perennially thorny issue in trust research at the international level, namely the question of agency. A key question, though one rarely explicitly analysed, is whether collectivities can trust each other, or whether trust can only exist at the interpersonal level. Here I argue for the latter position, namely, that the concepts of trust and trustworthiness belong to the individual and interpersonal level. But—and this is crucial to my argument—those human agents who are central to whether diplomacy succeeds or fails at all levels speak and act in the name of collectivities called states. In so doing, they are enabled and constrained by a set of intervening structural factors related to the roles and responsibilities of the offices they hold; the domestic political systems in which they operate; and wider societal narratives. Many trust researchers in International Relations obscure these important analytical boundaries by changing the referent for trust in their writings from individual leaders to collectivities without explaining what is at stake here.¹⁶

By framing the question of the referent in the way I propose above, it is possible to make a clear distinction between notions of *trust* and of *cooperation*. In my argument, it is appropriate to talk about collectivities cooperating, as it is to talk about former adversary-states gaining increased confidence in the peaceful motives and intentions of one another. Using this level-of-analysis approach is also the best way of conceiving of the difference between trust and confidence at the international level.¹⁷ Thus, it is meaningful to say the United States and Soviet

¹⁴ Karl Wolfgang Deutsch, Sidney A. Burrell, Robert A. Kann, Maurice Lee Jr, Martin Lichterman, Raymond E. Lindgren, Francis L. Loewenheim and Richard W. Van Wagenen, *Political community and the north Atlantic area: international organization in the light of historical experience* (Princeton, NJ: Princeton University Press, 1957). For later treatments, see Emanuel Adler and Michael Barnett, 'Security communities in theoretical perspective', in Emanuel Adler and Michael Barnett, eds, *Security communities* (Cambridge: Cambridge University Press, 1998); Booth and Wheeler, *The security dilemma*, pp. 182–200.

¹⁵ These terms are explained below (see n. 47).

¹⁶ I am grateful to Naomi Head and Laura Considine for highlighting this issue of changing referents. For examples of this analytical slippage between levels, see Andrew Kydd, *Trust and mistrust in international relations* (Princeton, NJ: Princeton University Press, 2005); Deborah Welch Larson, *Anatomy of mistrust: US–Soviet relations during the Cold War* (Ithaca, NY: Cornell University Press, 1997). For the view that trust can be applied to collectivities, see Vincent Keating and Jan Ruzicka, 'No need to hedge: trusting relationships in international politics' (copy on file with the author).

¹⁷ For a different way of conceiving of the relationship between trust and confidence, see Keating and Ruzicka, 'No need to hedge'.

Union cooperated when they signed a treaty, or that one or other of these states made a unilateral move which started a process of cooperation. But my principal proposition is that a key *enabling* condition (necessary if not sufficient) for setting in motion and sustaining a process of diplomatic transformation (as against limited and temporary episodes of cooperation) is trust at the interpersonal level between policy-makers on the two opposing sides.

Despite the centrality of trust to diplomatic transformations, there is a surprising lack of literature on how it can be built up in conflict situations, especially at the interstate level.¹⁸ Wight recognized that unilateral conciliatory moves (he did not enter into the question of where agency should be located for such moves, illustrating the point above concerning the analytical fuzziness of these boundaries) might inspire confidence in the peaceful motives of that side in the perceptions of an adversary, but he did not consider (nor could he, given his inattention to the agency question) how such moves might depend upon a prior foundation of trust between leaders at the highest levels of government.

The article uses a case-study of the summitry between President Ronald Reagan and his Soviet counterpart, Mikhail Gorbachev, and their key advisers in the mid- to late 1980s to show how these diplomatic encounters built trust. Moreover, I argue that, as a consequence of this growth of trust at the level of policy-makers, Gorbachev felt able to make a series of unilateral moves that significantly de-escalated the conflict between the United States and the Soviet Union. My justification for focusing on the Reagan–Gorbachev case is twofold. First, the Cold War was the backdrop against which Wight was writing; and second (and more importantly), while this case has been extensively researched, there are few conceptual works that study it through the prism of trust.¹⁹ Moreover, those IR theorists who have applied this lens have failed to recognize the centrality of Gorbachev's and Reagan's face-to-face communicative encounters to the building of trust and the role that this played in making possible the Soviet Union's conciliatory frame-breaking moves.²⁰ Conversely, some historians have recognized the importance of the interpersonal dynamics between Reagan and Gorbachev in promoting trust and cooperation,²¹ but they have not considered what lessons emerge from this case for the kind of communicative encounters that might promote trust in other cases.

¹⁸ Key exceptions are Adler and Barnett, 'Security communities in theoretical perspective'; Larson, *Anatomy of mistrust*; Kydd, *Trust and mistrust in international relations*; Aaron M. Hoffman, *Building trust: overcoming suspicion in international conflict* (Albany, NY: SUNY Press, 2006); Booth and Wheeler, *The security dilemma*; Wheeler, 'I had gone to Lahore with a message of goodwill', pp. 155–76. For an important argument that emphasizes the promise of trust in conflict situations, but does so by arguing for the mutual interrelationship between notions of trust, empathy and dialogue, see Head, 'Transforming conflict'.

¹⁹ The key exceptions are Kydd, *Trust and mistrust in international relations*; Larson, *Anatomy of mistrust*; Booth and Wheeler, *The security dilemma*.

²⁰ In making this argument, I am drawing on ideas initially explored in Booth and Wheeler, *The security dilemma*, pp. 155–6, 166–7; Nicholas J. Wheeler, 'To put oneself in the other fellow's place: John Herz, the security dilemma and the nuclear age', *International Relations* 22: 4, 2008, pp. 493–509 at p. 495; Nicholas J. Wheeler and Laura Considine, 'Reagan may have had trust in Gorbachev, but the United States demanded verification of the Soviet Union', in Martin Klimke, Reinhild Kreis and Christian Ostermann, eds, 'Trust, but verify': the politics of uncertainty and the transformation of the Cold War order, 1969–1991 (forthcoming).

²¹ Notably Reynolds, *Summits*, pp. 317–70.

It is necessary at this point to acknowledge the methodological challenges that confront researchers seeking to identify how particular communicative encounters built trust between policy-makers in adversarial relationships. This article draws heavily on accounts in memoirs by those who were present at these meetings, as well as declassified transcripts of the summit discussions that are available, and which have been mined by scholars working on this period. The usual methodological health warnings about materials of this kind apply; it is necessary to be on guard against deliberately self-serving accounts or, more problematically, actors who interpret their role in past events through the prism of later thinking and evidence. Nevertheless, an important defence can be made of the methodological approach adopted here. Crucially, an explanation is needed of how the Soviet Union—in the absence of any US positive reciprocation—made significantly bolder moves in the period 1987–88. As I argue below, existing explanations for these frame-breaking conciliatory moves are deficient, and until compelling evidence is produced that disqualifies the personal accounts of the key actors upon which the trust argument advanced here is based, the latter stands as an important explanation (though not a sufficient one) of the diplomatic transformation that took place between the United States and the Soviet Union. Put differently, given our present state of knowledge, neither position on the veracity of the accounts is falsifiable.

The second half of the article is taken up with the case-study. The first part identifies two important (though not exhaustive) pathways—*reciprocity* and *unilateralism*—by which a conflict might be de-escalated, and considers the role that trust might play in opening up these possibilities. Here, the discussion develops the central importance of the security dilemma, and the challenge of overcoming—through the exercise of empathy—the conflictual dynamics to which it gives rise. Crucially, the article shows how two key mindsets—*peaceful/defensive self-images*²² and *ideological fundamentalism*²³—block actors from exercising security dilemma sensibility, a particular form of empathy, in conflict situations.

Pathways to diplomatic transformation

The importance of showing empathy for an adversary's security concerns and interests is an idea rooted in the classic works of Herbert Butterfield, John Herz and Robert Jervis, and further enriched by more recent contributions to the debate.²⁴ Drawing on this literature, Ken Booth and I developed the concept

²² The term was coined by the author in 'To put oneself in the other fellow's place'.

²³ This term is discussed in Booth and Wheeler, *The security dilemma*, p. 65. It was first developed by Booth in 'New challenges and old mindsets: ten rules for empirical realists', in Carl G. Jacobsen, ed., *The uncertain course: new weapons, strategies and mindsets* (Oxford: Oxford University Press for Stockholm International Peace Research Institute, 1987), pp. 39–66.

²⁴ Herbert Butterfield, *History and human relations* (London: Collins, 1951), p. 21; John H. Herz, *International politics in the atomic age* (New York: Columbia University Press, 1959), p. 249; Jervis, *Perception and misperception in international politics*, pp. 65–76; Robert Jervis, 'Cooperation under the security dilemma', *World Politics* 30: 2, 1978, p. 181. The post-Cold War period has seen a growing interest in International Relations in the concept of empathy and its applicability to conflict issues. One highly innovative approach here has been the work of James Blight and Janet Lang, who have pioneered the use of what they called 'critical oral history'. The

of security dilemma sensibility, which we defined as ‘an actor’s intention and capacity to perceive the motives behind, and to show responsiveness towards, the potential complexity of the military intentions of others. In particular, it refers to the ability to understand the role that fear might play in their attitudes and behaviour, including, crucially, the role that one’s own actions may play in provoking that fear.’²⁵

The first operational implication of the exercise of security dilemma sensibility would be an increased awareness that an adversary might be operating with a peaceful/defensive self-image. It was Butterfield who drew attention to the dynamics of mutual fear that can arise when governments with defensive intent (through their failure to see themselves as others see them) provoke potential adversaries to behave in ways that increase the level of mutual insecurity. In a much-quoted passage, he wrote:

It is the peculiar characteristic of the situation I am describing—the situation of what I should call Hobbesian fear—that you yourself may vividly feel the terrible fear that you have of the other party, but you cannot enter into the other man’s counter-fear, or even understand why he should be particularly nervous. For you know that you yourself mean him no harm, and that you want nothing from him save guarantees for your own safety; and it is never possible for you to realize or remember properly that since he cannot see the inside of your mind, he can never have the same assurance of your intentions that you have. As this operates on both sides the Chinese puzzle is complete in all its interlockings—and neither party sees the nature of the predicament he is in, for each only imagines that the other party is being hostile and unreasonable. It is even possible for each to feel that the other is wilfully withholding the guarantees that would have enabled him to have a sense of security.²⁶

Actors exercise security dilemma sensibility when they show an appreciation of how an interstate conflict might be driven by dynamics of ‘Hobbesian fear’ and how, as a result, it is no good expecting an adversary to make the decisive first move in ending a conflict.

Unfortunately, the exercise of security dilemma sensibility is a rarity among decision-makers, who as a result tend to expect their adversary to make the critical

key purpose of this method is to bring former officials and decision-makers, who had experienced conflict situations, back together to reflect on their previous encounters with a view to learning lessons for ending conflicts. Empathy is at the heart of this process, and Blight and Lang have generated a treasure trove of data on the cases of Vietnam, the Cuban Missile Crisis and the end of the Cold War. For an excellent overview of the method and its application to the failure of detente, see James M. Blight and Janet M. Lang, ‘When empathy failed: using critical oral history to reassess the collapse of US–Soviet détente in the Carter–Brezhnev years’, *Journal of Cold War Studies* 12: 2, 2010, pp. 29–74. See also Ralph K. White, ‘Empathizing with Saddam Hussein’, *Political Psychology* 12: 2, 1991, pp. 291–308; Neta C. Crawford, ‘The passion of world politics: propositions on emotion and emotional relationships’, *International Security* 24: 4, 2000, pp. 116–56; Richard Ned Lebow, ‘Reason, emotion and cooperation’, *International Politics* 42: 3, 2005, pp. 283–313; Roland Bleiker and Emma Hutchison, ‘Fear no more: emotions and world politics’, *Review of International Studies* 34 (supplement S1), 2008, pp. 115–35; Head, ‘Transforming conflict’.

²⁵ Booth and Wheeler, *The security dilemma*, p. 7.

²⁶ Butterfield, *History and human relations*, p. 21. Robert Jervis has been the most prominent thinker since Butterfield to develop and elaborate the psychological dynamics that underpin the security dilemma. Jervis argued that the key driver behind the ‘spiral model’ of international conflict is policy-makers’ failure to ‘recognize that one’s own actions could be seen as menacing and the concomitant belief that the other’s hostility can only be explained by its aggressiveness’ (Jervis, *Perception and misperception*, p. 75).

opening moves in ending a conflict. Consider, for example, Strobe Talbott's recollection of a meeting between himself, as Deputy Secretary of State in the Clinton administration, the then US National Security Advisor Sandy Berger, and India's then Minister of External Affairs, Jaswant Singh, at which Berger lambasted Singh for his lack of empathy towards Pakistan. Berger, recalling his experiences with peacemaking in Northern Ireland and the Middle East, explained how he repeatedly found himself dealing with antagonists who expected the other party to make the first conciliatory move. He said:

Gerry Adams wants to talk about David Trimble and what the other guy has got to do to make peace; Netanyahu wants to talk about Arafat and what he's got to do, and vice versa. And now I find that all you [Singh] want to do is talk about the Pakistanis. Let's talk about what *India* can do to make the situation better.²⁷

But if both sides in a conflict each expect the other to make the first conciliatory move, can there ever be an escape from unending competition?

One possible pathway out of this dead end is *reciprocation*, in which both state parties enter into a negotiating process, which leads them to agree a series of reciprocal conciliatory moves. In principle, this circumvents the problem of one side losing face—or, worse, being exploited—if it makes a unilateral move when there is no guarantee of reciprocation; but it crucially assumes that both state parties can agree on what count as reciprocal steps. If the assumption is made, as Berger lamented in the examples he gave, that neither party has exercised security dilemma sensibility, then it is very difficult to see how even a limited conciliatory move of de-escalation might come about. But even if adversaries could take such a step, in the absence of trust (or perhaps compelling means of verification) it is hard to see how this could lead to future rounds of reciprocated cooperation, let alone a frame-breaking move—Kenneth Waltz recognized this problem when he wrote that 'first steps toward agreement do not lead to second and third steps'.²⁸ But where he explained the lack of cooperation in terms of the self-help logic of international anarchy, I want to highlight the failure of policy-makers to develop the interpersonal trust that is necessary before actors will make the kind of moves that are required if a process of diplomatic transformation is to take place.

Despite appreciating that mutual confidence was the moral condition for diplomacy, Martin Wight also fell into the way of thinking which Berger was lamenting, considering that 'the opposite number must prove his readiness to negotiate; he must create confidence on the other side'.²⁹ Nevertheless, Wight was emphatic that if an opponent made a conciliatory move, however limited this might be, then responsible statecraft required that one show a readiness to negotiate in good faith, stepping forward, as he put it, 'with a gleaming smile and an outstretched hand'.³⁰ Indeed, he cautioned against the Machiavellian temptation to push for yet greater concessions in the belief that the adversary was on

²⁷ Strobe Talbott, *Engaging India: diplomacy, democracy and the bomb* (Washington DC: Brookings, 2004), p. 122.

²⁸ Waltz, *Theory of international politics*, p. 175.

²⁹ Wight, *International theory*, p. 185.

³⁰ Wight, *International theory*, p. 186.

the ropes, and that if sufficient leverage was exercised an even more advantageous bargain could be struck.³¹

The second possible pathway to diplomatic transformation, and potentially the most fruitful, is *unilateralism*: rather than waiting for an opponent to make the first move, one side seizes the initiative and acts unilaterally. Writing nine months before the Cuban Missile Crisis, the US social psychologist Charles Osgood had proposed in his book *An alternative to war or surrender* a strategy for reducing tension between the superpowers, which he called GRIT (Graduated Reciprocation in Tension Reduction). Osgood's approach was predicated on one side making a limited conciliatory move and publicly inviting reciprocation on the part of the other. If reciprocity was forthcoming, Osgood recommended that the initiating state follow up with bolder cooperative moves. If, on the other hand, one's opponent did not reciprocate, then Osgood did not consider that the initiating state should switch to an alternative, perhaps even confrontational strategy. Rather, he argued that the state pursuing GRIT should have the determination to carry on making limited unilateral conciliatory moves in the expectation that this would eventually trigger reciprocation.³²

Osgood's notion that the only way to break out of the conflictual logic of Jervis's spiral model was for one side to act unilaterally had been suggested by Wight a few years earlier. Wight appreciated the difficulties of establishing the 'confidence' that others really want to negotiate, and this led him to suggest that one way of cutting through the Gordian knot was for one side to 'make a gesture'.³³ Although Wight used the term 'gesture', it is clear from the example he gave of one side dramatically reducing armaments³⁴ that his notion of a gesture was closer to the idea of what others have called a 'costly signal'.³⁵ The latter is a move that is designed, in Andrew Kydd's words, 'to persuade the other side that one is trustworthy by virtue of the fact that [such signals] are so costly that one would hesitate to send them if one were untrustworthy'.³⁶ Costly signals arguably overcome one of the major weaknesses of GRIT, which is that the cooperative moves it recommends are likely to be too limited to convince decision-makers in the adversary state who are mistrustful and suspicious of an adversary's motives and intentions.

This problem of the signal being too weak is exacerbated if decision-makers in the state to which the conciliatory move is addressed are operating within a mindset of ideological fundamentalism. This has been defined as a 'mindset

³¹ Wight, *International theory*, pp. 183, 187.

³² Charles E. Osgood, *An alternative to war or surrender*, 2nd edn (Chicago, IL: University of Illinois Press, 1970; first publ. 1962), pp. 85–9, 94–8. See also Larson, *Anatomy of mistrust*, pp. 27, 31.

³³ Wight, *International theory*, p. 198 (emphasis in original).

³⁴ Wight, *International theory*, p. 198.

³⁵ For the concept of costly signalling, see James Fearon, 'Threats to use force: costly signals and bargaining in international crises', PhD diss., University of California, Berkeley, 1992; James Fearon, 'Signalling foreign policy interests: tying hands versus sinking costs', *Journal of Conflict Resolution* 41: 1, 1997, pp. 68–90; Andrew Kydd, 'Trust, reassurance, and cooperation', *International Organization* 54: 2, 2000, pp. 325–57; Kydd, *Trust and mistrust in international relations*; Charles L. Glaser, *Rational theory of international politics: the logic of competition and cooperation* (Princeton, NJ: Princeton University Press, 2010), pp. 64–8.

³⁶ Kydd, 'Trust, reassurance, and cooperation', p. 326.

which assigns enemy status because of what the other *is*.³⁷ A particular actor's behaviour is perceived as threatening and untrustworthy as a result of certain inherent characteristics (e.g. its political identity, fundamental values, etc.). Ideological fundamentalism gives rise to what Ole Holsti has called (building on an idea first discussed by Henry Kissinger)³⁸ an 'inherent bad faith model'.³⁹ The implication of this way of thinking is that situational explanations of actions are unimportant, relative to the drivers of policy that are rooted in the dispositional characteristics of actors. Consequently, what one side's decision-makers view as a genuine conciliatory move might be discounted by decision-makers in an adversary state as tokenistic or, even worse, a trick aimed at lulling them into a false sense of security. The importance of costly signalling—and one does not have to commit to Kydd's methodological postulates of rational choice and game theory to recognize the value of this—is that it promises decision-makers the possibility of fashioning signals that can overcome a bad faith model because these are *exactly* the signals that actors would not send if they had aggressive designs.⁴⁰

Kydd's use of the language of trustworthiness is significant because he is the first theorist of security cooperation working within a rationalist framework to make the explicit link between the costliness of the signal and '*the level of trust*'.⁴¹ He defined trust 'as a belief that the other side is trustworthy, that is, willing to reciprocate cooperation, and mistrust as a belief that the other side is untrustworthy, or prefers to exploit one's cooperation'.⁴² Kydd's argument rests on the premise that states will update their beliefs about the other's trustworthiness as initial rounds of cooperation take place on less important issues. Deborah Welch Larson makes a similar argument about trust developing out of initial rounds of cooperation, and appreciates that trust is the key to achieving cooperation that goes beyond 'agreements of small scope' where the penalties for defection are less severe.⁴³ Nevertheless, neither she nor Kydd adequately explains how cooperation becomes possible in the first place if actors are holding peaceful/defensive self-images and/or are operating with a mindset of ideological fundamentalism and its logical corollary, bad faith thinking. In short, we are still left with the question as

³⁷ Booth and Wheeler, *The security dilemma*, p. 65.

³⁸ Henry A. Kissinger, *Necessity for choice: prospects for American foreign policy* (London: Chatto & Windus, 1960), p. 194.

³⁹ Holsti's pioneering contribution is developed in David J. Finlay, Richard R. Fagen and Ole R. Holsti, *Enemies in politics* (Chicago, IL: Rand McNally, 1967), p. 26. See also Larson, *Anatomy of mistrust*, p. 22.

⁴⁰ For a sceptical view of the role that costly signals can play in signalling a state's peaceful motives and intentions, see Robert Jervis, 'Dilemmas about security dilemmas', *Security Studies* 20: 3, 2011, p. 420.

⁴¹ Kydd, *Trust and mistrust in international relations*, p. 198 (emphasis in original). Kydd also argued that 'strategic situation' was a variable affecting the size of the signal, considering that the more powerful a state is, the weaker the signal it is prepared to send. Conversely, weaker states, he argued, 'will make more costly gestures'. At the same time, he also suggested that if the 'costs of conflict' were perceived to be high, then 'the stronger the reassuring signals states will be willing to send' (Kydd, *Trust and mistrust in international relations*, p. 200).

⁴² Kydd, *Trust and mistrust in international relations*, p. 30. For a critique of Kydd's approach to conceptualizing trust in international politics which emphasizes the emotional dimension of trust, see Booth and Wheeler, *The security dilemma*, pp. 232–4; Torsten Michel, 'Time to get emotional: phronetic reflections on the concept of trust in International Relations', *European Journal of International Relations*, April 2012, pp. 1–22; Toros, 'Terrorism, talking and transformation'; Head, 'Transforming conflict'. See also Guido Möllering, *Trust: reason, routine and reflexivity* (Oxford: Elsevier, 2006); Rathbun, *Trust in international cooperation*, pp. 1–57.

⁴³ Larson, *Anatomy of mistrust*, p. 21.

to how the circle of cooperation gets going; and without this, Kydd (like Larson) has no mechanism for generating the changed perceptions of trustworthiness that lead decision-makers to send costly signals.⁴⁴

Kydd has applied his theory to explain Gorbachev's unilateral conciliatory moves in the mid- to late 1980s which were so decisive in transforming the Cold War conflict. Although the Soviet Union made a series of unilateral concessions shortly after Gorbachev came to power in March 1985 (e.g. a moratorium on nuclear testing in July 1985),⁴⁵ these were dismissed by US policy-makers as little more than propaganda gestures, or, in Kydd's terms, 'cheap talk'.⁴⁶ As such, they did little to change the prevailing belief of US policy-makers that the Soviet Union had malign motives and intentions. Kydd argued that what changed US perceptions of the Soviet Union's trustworthiness were the costly signals that Moscow sent in 1987–8, crucially the Soviet concessions that made possible the signing of the Intermediate Nuclear Forces Agreement in December 1987 (in which the Soviet Union accepted on-site verification, and agreed as part of the treaty to the elimination of an entire class of Soviet missiles) and Gorbachev's commitment made before the UN General Assembly in December 1988 to remove from the Central Front in Germany those short-warning surprise attack forces that so worried NATO's strategic planners. If we apply Kydd's theory, the logical inference to draw from these costly signals is that the Soviet Union had developed an increased level of trust in the United States.

There are a number of problems with Kydd's application of his theory to this case. First, as noted above, there is the question whether it is meaningful to impute, as Kydd does, perceptions of trustworthiness to collectives. Kydd seems unaware of the analytical issues at stake here, and moves between the levels of individual decision-makers and state entities without justifying these shifts. Thus, when Kydd argues that the Soviet Union had gained a new sense of US trustworthiness, thereby making possible the costly signalling of 1987–8, it is not clear whether he has in mind the Soviet Union as a collective entity, the new Soviet leadership or Gorbachev himself.

The next difficulty concerns Kydd's mechanism for increasing the level of trust between two adversaries. In the US–Soviet case, there is no evidence of prior rounds of reciprocated cooperation leading to the growth of trust. Indeed, one of the striking features of this case is that despite the United States not positively reciprocating any of the Soviet Union's limited moves in 1985 (e.g. the testing moratorium),⁴⁷ the Soviet Union responded not only by continuing to

⁴⁴ This point is explored in Rathbun, *Trust in international cooperation*, pp. 12–23, esp. p. 17.

⁴⁵ Jonathon Dean, 'Gorbachev's arms control moves', *Bulletin of the Atomic Scientists* 4: 8, 1987, pp. 34–40; Federation of American Scientists, 'Comprehensive Test Ban Treaty chronology', <http://www.fas.org/nuke/control/ctbt/chron1.htm>, accessed 20 Feb. 2013.

⁴⁶ Kydd, *Trust and mistrust in international relations*, p. 225; Larson, *Anatomy of mistrust*, p. 202.

⁴⁷ The puzzle of the lack of US reciprocation is discussed in Andrew O. Bennett, 'Trust bursting out all over', in William C. Wohlforth, ed., *Cold War endgame: oral history, analysis, debates* (Philadelphia: Pennsylvania State University Press, 2003), pp. 175–207. Although the Reagan administration did not positively reciprocate the Soviet Union's moves, it did not do anything to exploit them. Towards the end of the Cold War Ken Booth made an important distinction between 'positive reciprocation' and 'negative reciprocation'. The former refers to actions which are taken in direct response to a conciliatory move and which reward the initiating state

make conciliatory moves, as in Osgood's GRIT schema, but taking progressively bolder steps—*contra* the GRIT model. The logic of Kydd's theory is that these Soviet moves must have been made possible by an increase in trust; but he has not supplied a convincing theory of how this level of trust was achieved.⁴⁸ The argument I will develop in the next section of the article is that the Soviet Union's conciliatory frame-breaking moves became possible because of the relationship of trust that developed out of the communicative encounters between Gorbachev and Reagan, and their key advisers.

Diplomatic encounters of the empathetic kind

In arguing that a critical and under-explored mechanism in the growth of cooperation between adversaries is the trust that can be built between policy-makers as a result of their communicative encounters, I am not claiming that all that is necessary for a conflict to be transformed is for enemies to meet and talk. Putting enemy policy-makers in the same room may simply have the effect of heightening their awareness of what is at stake in the conflict, how much they fear and distrust each other, and how determined they are not to make concessions. This was certainly the outcome of the disastrous summit meeting between US President John F. Kennedy and his Soviet counterpart Nikita Khrushchev in Vienna in June 1961. Neither leader was able to empathize with—let alone trust—his counterpart over the two days, during which Kennedy and Khrushchev hammered away at each other on the ideological failings of the other's system—a textbook case of the ideological fundamentalism discussed above.⁴⁹ Nevertheless, summitry of this kind does hold out the possibility that leaders will develop trust in the motives and intentions of their enemy counterparts. But before this outcome becomes possible, it is necessary to get policy-makers on one side, but preferably both, to exercise security dilemma sensibility; and the Reagan–Gorbachev case suggests that there are a number of pathways—cognitive and emotional—by which this might come about.

The US and Soviet leaders were able at their first summit in Geneva in November 1985 to avoid the dynamics of the ill-fated meeting between Kennedy and Khrush-

with an equivalent concession. By contrast, negative reciprocation occurs if the state that is rewarded with the concession(s) does not seek to take advantage of the initiating state's move(s) by taking steps which make it less secure. See Ken Booth and John Baylis, *Britain, NATO and nuclear weapons: alternative defence versus alliance reform* (Basingstoke: Macmillan, 1989), p. 205. See also Alan Collins, *The security dilemma and the end of the Cold War* (Edinburgh: Edinburgh University Press, 1997), p. 193. For a discussion of the Reagan administration's negative reciprocation of the Soviet Union's conciliatory moves, see Abraham Sofaer, 'A legacy of Reykjavik: negotiating with enemies', in Sidney Drell and George Shultz, eds, *Implications of the Reykjavik summit on its twentieth anniversary* (Stanford, CA: Hoover Institution Press, 2007), pp. 132–3. This discussion of different approaches to reciprocation and its application to US–Soviet security cooperation in the second half of the 1980s draws on material in Wheeler and Considine, 'Reagan may have had trust in Gorbachev'.

⁴⁸ The same criticism can be levelled at Larson, who accepts that 'Soviet–American cooperation did not arise through a tit-for-tat process' (referring to Robert Axelrod's 1984 work, *The evolution of cooperation*, and his model of how cooperation can begin in adversarial contexts). Indeed, she points out that Gorbachev made 'more than twenty-five concessions' that were not positively reciprocated. At the same time, Larson appreciates that Gorbachev made 'bold moves' despite the lack of US positive reciprocation. But she offers no convincing explanation of why Gorbachev chose to make these frame-breaking moves in 1987–8, which she credits as decisive to the transformation of US–Soviet relations (Larson, *Anatomy of mistrust*, pp. 233, 240).

⁴⁹ Reynolds, *Summits*, pp. 151–207.

chev in Vienna because each had already begun the process of empathizing with his adversary.⁵⁰ In the case of Reagan, it was the emotion of fear—fear of nuclear war—that led him to realize the importance of exercising security dilemma sensibility.⁵¹ A key event in what Beth Fischer has called Reagan's 'nuclear learning' was the Able Archer episode of November 1983.⁵² In an annual exercise, NATO simulated its nuclear release procedures, but in the climate of heightened fear and suspicion which characterized superpower relations in the early 1980s, US decision-makers subsequently believed that Soviet military and political authorities might have seen this as the countdown to a US nuclear attack against the Soviet Union.⁵³ Even if later scholarship has questioned how far Soviet authorities really believed that they were at risk from US nuclear attack,⁵⁴ what matters is the impact that the episode had on Reagan's thinking. The US President came for the first time to understand the role that fear might be playing in Soviet attitudes and behaviour, and to appreciate that Moscow did not accept the White House's peaceful/defensive self-image.⁵⁵ This was evident in Reagan's remarks after reading a US intelligence report which documented US perceptions of Soviet fears during the Able Archer exercise. According to a number of sources, the President commented to his National Security Advisor, Robert McFarlane: 'Do you suppose they really believe that? I don't see how they would believe that. But it's something to think about.'⁵⁶ Reagan reflected further on this in his memoirs, in a passage that showed an important awareness of how security dilemma dynamics might underpin the US–Soviet conflict. He wrote that 'there were myths and misconceptions that had contributed to misunderstandings and our potentially fatal mistrust of each other'.⁵⁷ Having realized that the Soviet Union could so misunderstand US intentions as to believe that Washington was preparing to launch a nuclear first strike, Reagan was anxious to get a Russian leader in a room so as to reassure him as to America's peaceful intentions.⁵⁸

Mikhail Gorbachev's accession to the leadership of the Soviet Union in March 1985 provided Reagan with an interlocutor who shared his desire for

⁵⁰ This discussion draws on material in Booth and Wheeler, *The security dilemma*; Wheeler, 'To put oneself in the other fellow's place'; and Wheeler and Considine, 'Reagan may have had trust in Gorbachev'.

⁵¹ Booth and Wheeler, *The security dilemma*, p. 150.

⁵² Beth A. Fischer, *The Reagan reversal* (Columbia: University of Missouri Press, 2006), p. 153.

⁵³ For treatments which focus on Soviet fears of a US nuclear attack and the impact on Reagan's thinking, see Christopher Andrew and Oleg Gordievsky, *KGB: the inside story of its foreign operations from Lenin to Gorbachev* (New York: HarperCollins, 1990), p. 502; Benjamin B. Fischer, 'The Soviet–American war scare of the 1980s', *International Journal of Intelligence and Counterintelligence* 19: 3, 2006, pp. 480–518; Fischer, *The Reagan reversal*, pp. 109–10, 122–31; Reynolds, *Summits*, pp. 323–4; Richard Rhodes, *Arsenals of folly: the making of the nuclear arms race* (New York: Vintage, 2008), pp. 163–7.

⁵⁴ See Vojtech Mastny, 'How able was "Able Archer"? Nuclear trigger and intelligence in perspective', *Journal of Cold War Studies* 11: 1, 2009, pp. 108–23; Len Scott, 'Intelligence and the risk of nuclear war: Able Archer-83 revisited', *Intelligence and National Security* 26: 6, 2011, pp. 759–77.

⁵⁵ Fischer, *The Reagan reversal*, pp. 133–43; Booth and Wheeler, *The security dilemma*, p. 53; Reynolds, *Summits*, p. 324; Rhodes, *Arsenals of folly*, pp. 180–81.

⁵⁶ See Ronald Reagan, 'American experience', broadcast on BBC2, 27 Dec. 2004; Martin Walker, *The Cold War: a history* (New York: Henry Holt, 1995), p. 286; Fischer, *The Reagan reversal*, pp. 131–40; Reynolds, *Summits*, p. 324; Rhodes, *Arsenals of folly*, p. 167.

⁵⁷ Quoted in Fischer, *The Reagan reversal*, p. 140.

⁵⁸ Ronald R. Reagan, *An American life: the autobiography* (New York: Simon & Schuster, 1990), pp. 11, 588–9, 603, 636; Walker, *The Cold War*, p. 286; Fischer, *The Reagan reversal*, pp. 48, 139.

better relations. There had been summits between US and Soviet leaders in the 1970s during the period of detente, most notably the 1972 Moscow summit at which President Richard Nixon and General Secretary Leonid Brezhnev signed the first treaties limiting US and Soviet offensive and defensive arms. Nevertheless, detente eventually failed because neither superpower could jettison the ideological fundamentalism which made each side continue to view the relationship in zero-sum terms. The significance of Gorbachev's 'new thinking' in Soviet security policy was that he and his key advisers came to challenge the ideological fundamentalism of the Khrushchev and Brezhnev years that had depicted the United States and its allies as inherently aggressive by virtue of their capitalist values and interests. Instead, the new Soviet leader was open to those in his inner circle who argued that the United States and the Soviet Union might be ensnared in a nuclear version of Jervis's spiral model. These advisers had been strongly influenced by ideas of common or mutual security that had importantly developed out of regular East–West contacts between peace movements inside the NATO and Soviet blocs.⁵⁹ The new Soviet leadership came to understand that the Soviet Union might profess defensive motives and intent, but its enemies were not so easily reassured in the face of Soviet conventional capabilities that were configured for offensive operations against NATO forces.⁶⁰

Consequently, both leaders arrived in Geneva open to the possibility that their countries were trapped in a vicious spiral of reciprocal fears. Nevertheless, it was one thing for US and Soviet leaders to frame the conflict as an example of the spiral model; it was quite another to begin making the unilateral moves that might transform the conflict. The question that guides this enquiry is how important trust was in making possible the Soviet Union's conciliatory frame-breaking moves, and, above all, what role diplomacy at all levels—but especially the highest level—played in building this trust.

There is a clear correlation between the Soviet Union's costly signalling in this period and the summit meetings that Reagan and Gorbachev held in Geneva in November 1985, Reykjavik in October 1986, Washington in December 1987 and Moscow in March 1988, and the related but critical meetings of US Secretary of State George Shultz and his counterpart Eduard Shevardnadze. Moreover, all the players who were present at these face-to-face meetings, including the two heads of state themselves, have emphasized how important these particular communicative encounters were in building trust, and how this trust deepened as they gained greater familiarity through increased contact and connection with one another.⁶¹ For example, after the Geneva summit Gorbachev told Anatoly Chernyaev, Deputy Head of the International Department of the CPSU, that 'a spark of

⁵⁹ These developments are discussed in Fierke, *Changing games, changing strategies*; Matthew Evangelista, *Unarmed forces: the transnational movement to end the Cold War* (Ithaca, NY: Cornell University Press, 1999).

⁶⁰ Mikhail S. Gorbachev, *For a nuclear free world* (Moscow: Novosti, 1987), p. 39; Geoffrey Wiseman, *Concepts of non-provocative defence: ideas and practices in international security* (London: Palgrave Macmillan, 2002), pp. 87–91; Rhodes, *Arsenals of folly*, pp. 191–3; Larson, *Anatomy of mistrust*, pp. 208–10; Glaser, *Rational theory of international politics*, p. 209.

⁶¹ Harmonie Toros talks about the role that familiarization can play in transforming conflict in *Terrorism, talking and transformation*, pp. 78–83. See also Fierke, *Diplomatic interventions*.

electric mutual trust ... ignited between us, like a voltaic arc between two electric poles'.⁶² Such sentiments were shared by Reagan, who wrote in his memoirs: 'Looking back now, it's clear there was a chemistry between Gorbachev and me that produced something very close to a friendship.'⁶³ Similarly, Gorbachev remarked to Nancy Reagan when she and her husband visited the country Reagan had once called the 'evil empire' that he and the US President had 'a certain chemistry' that he described as 'very rare'.⁶⁴ What this shows, as Jack Matlock, a former US diplomat who was present at all four summits, has reflected, is that 'face-to-face meetings between the Soviet and American leaders and their policy makers were essential to move the US–Soviet dialogue in a constructive direction. Under these conditions, the overwhelming suspicion characteristic of the Cold War was gradually replaced by trust.'⁶⁵ Matlock captured in this reflection the transformational potential of diplomacy at the highest level, an aspect of, and means towards, transforming adversarial relationships that Wight had neglected.

What Wight did capture in his rich writings on diplomacy was the creative imagination and empathy expected of diplomats. He wrote that diplomacy involves 'thinking the best of people, and trying to share their point of view, understand their interests ... honesty, moderation, courtesy, sympathy for others ... a reconciliation of interests, a composing of differences, even, perhaps, creating a common interest'.⁶⁶ Yet there is no evidence that US and Soviet diplomatic emissaries—the level on which Wight focused his attention—played any significant role in bringing the leaders together in Geneva. By contrast, the two foreign ministers played an important part in making the summit possible. Their meetings, first in Helsinki in July 1985, and then more productively in Moscow a month before the Geneva summit, cut through the bureaucratic obstacles that stood in the way of getting the two leaders in the same room.⁶⁷ David Reynolds, in his account of the Reagan–Gorbachev summitry, has gone so far as to argue that 'without his Secretary of State, Reagan would never have left a lasting legacy as a peacemaker'.⁶⁸ Nevertheless, the future transformation that would unfold could not have been predicted in November 1985, when each side went to Geneva with little agreed and minimal expectations of progress. The fact, then, that the summit *was* such a success, leading to an agreement on the core principle that neither side 'would seek military superiority',⁶⁹ and a commitment by both men to meet again, lends support to Matlock's claim that it was the face-to-face communicative encounters between Reagan and Gorbachev that were so decisive in the building of trust between them.

⁶² Quoted in David Reynolds, 'Summitry as intercultural communication', *International Affairs* 85: 1, Jan. 2009, p. 123.

⁶³ Quoted in Reynolds, *Summits*, p. 366.

⁶⁴ Quoted in Reynolds, *Summits*, p. 366.

⁶⁵ Jack F. Matlock, Jr, *Reagan and Gorbachev: how the Cold War ended* (New York: Random House, 2004), p. 319.

⁶⁶ Wight, *International theory*, pp. 187–8.

⁶⁷ Reynolds, *Summits*, pp. 329–32.

⁶⁸ Reynolds, *Summits*, p. 367.

⁶⁹ Quoted in Raymond L. Garthoff, *The great transition: American–Soviet relations and the end of the Cold War* (Washington DC: Brookings, 1994), p. 240.

Even if it is accepted that mutual trust developed between Reagan and Gorbachev through their summitry, it is still necessary to show that this was a key explanatory variable in leading the Soviet Union to engage in costly signalling. In so doing, it is important in particular to address two materialist arguments which seek to deny trust a major role in explaining the transformation of relations that occurred between the United States and the Soviet Union. The first of these arguments is premised on the claim that trust played little or no role in Soviet costly signalling because the Soviet Union was in the luxurious and rare position of having a 'margin of safety' that enabled it to send costly signals that increased US security without this decreasing the security of the Soviet Union.⁷⁰ The margin of safety is argued to have been provided by Soviet nuclear capabilities.⁷¹ On the basis of this argument, Evan Braden Montgomery claims that this is an exceptional case because it reverses the usual dilemma that faces decision-makers contemplating taking actions that might reassure an adversary (but one who is potentially open to cooperation), but which entail accepting at the same time an increase in vulnerability which could endanger the cooperating state if it turns out the other side does have aggressive motives and intent.⁷²

There is a debate among trust researchers in the social sciences and humanities as to whether trust requires only the perception of another actor's trustworthiness,⁷³ or whether it also requires the taking of deliberate actions based on perceptions of trustworthiness.⁷⁴ This article is firmly located in the latter camp, and the key action which can be seen as the litmus test of whether actors believe that another actor can be trusted, and the extent of that trust, is their willingness to take on new vulnerabilities or live with existing ones. As Annette Baier has put it: 'Trust is acceptance of vulnerability to harm that others could inflict, but which we judge that they will not in fact inflict.'⁷⁵ Consequently, if it could be shown that what purportedly was Soviet costly signalling in the period 1987–8 actually entailed no increase in the Soviet Union's vulnerability, then the claim that trust played an important role in making possible these frame-breaking moves would be an untenable one. In arriving at a judgement as to whether the Gorbachev leadership's decision to send the costly signals increased Soviet vulnerability, it is important to recognize that what counts as an acceptable level of vulnerability is not fixed, but is the subject of internal political contestation and bureaucratic battles, and also that changing perceptions of another leader and their key advisers'

⁷⁰ This idea is discussed further in Booth and Wheeler, *The security dilemma*, pp. 91, 241, 284.

⁷¹ Kydd, 'Trust, reassurance and cooperation', p. 343; Kydd, *Trust and mistrust in international relations*, p. 222; Evan Braden Montgomery, 'Breaking out of the security dilemma: Realism, reassurance and the problem of uncertainty', *International Security* 31: 2, 2006, pp. 181, 183; Glaser, *Rational theory of international politics*, p. 208.

⁷² Montgomery, 'Breaking out of the security dilemma', pp. 153–4. See also Glaser, *Rational theory of international politics*.

⁷³ See J. David Lewis and Andrew Weigert, 'Trust as social reality', *Social Forces* 63: 4, 1985, pp. 967–87; J. Pearce, 'Straining for shared meaning in organization science: problems of trust and distrust', *Academy of Management Review* 23: 3, 1998, pp. 405–21; Denise M. Rousseau, Sim B. Sitkin, Ronald S. Burt and Colin Camerer, 'Not so different after all: a cross-discipline view of trust', *Academy of Management Review* 23: 3, 1998, pp. 393–404.

⁷⁴ For examples of these opposing views, see Russell Hardin, *Trust and trustworthiness* (New York: Russell Sage Foundation, 2002); Möllering, *Trust*.

⁷⁵ Annette Baier, *Moral prejudices: essays on ethics* (Cambridge, MA: Harvard University Press, 1995), p. 152.

trustworthiness will alter this calculation in ways that open up new possibilities of cooperation.

Despite the contention of Kydd, Montgomery and Glaser that Moscow had a nuclear safety blanket, Gorbachev's unilateral conciliatory moves were perceived by others within the Soviet leadership, especially the military, as increasing the vulnerability of the Soviet state to potential US exploitation. Hence Gorbachev's historic decision in March 1986, at the deadlocked Stockholm Conference on Confidence and Security Building Measures and Disarmament in Europe, to become the first Soviet leader to agree to on-site inspections was strongly opposed by 'the General Staff, the Ministry of Defence, and the KGB'.⁷⁶ But according to Yevgeny Velikhov, the Chernobyl disaster a month earlier had reinforced Gorbachev's growing sense that 'a great instinctive leap to break the old cycle' of mistrust, suspicion and secrecy in East–West relations was required.⁷⁷ Indeed, Gorbachev had displayed frustration with the lack of progress being made in lower-level meetings between US and Soviet officials following the Geneva summit: 'We should', he said, 'give a shake-up to all this old cloth.'⁷⁸ This provided the impetus for him to push for another summit meeting with Reagan so that they could give momentum to the stalled negotiations at the lower levels of the bureaucracy.⁷⁹

The essential point to grasp here is that Gorbachev proceeded in this manner not—as Kydd, Montgomery and Glaser would have us believe—because he thought that the Soviet Union's nuclear arsenal made it secure from US exploitation. Indeed, he fully understood that making concessions entailed increasing Soviet vulnerability if the United States interpreted this as a sign of weakness, and chose to exploit it as Soviet hardliners feared it would. Gorbachev sought further talks not out of a sense of invulnerability, but rather out of trust; he had sensed in Geneva that he could trust Reagan and his inner circle, especially Shultz, not to exploit any conciliatory moves. This sense of trust in Reagan's *bona fides* (and Gorbachev was careful to distinguish between his trust in Reagan and his confidence in the motives and intentions of the United States, which he saw as still dominated by a military–industrial complex)⁸⁰ deepened after Reykjavik. At the same time, Gorbachev reasoned that any vulnerabilities entailed by his policy of cooperation had to be set against the risks and costs of the alternative path of confronting the United States. What is more, returning to Baier's core contention, it is in the nature of trust that the vulnerabilities of trusting are always discounted, because actors only act on trust when they believe that it will not prove misplaced (though there are no guarantees here).

⁷⁶ See Lynn Hansen and Oleg Grinevsky, 'Negotiating CSCE', in Shai Feldman, ed., *Confidence building and verification: prospects in the Middle East* (Tel Aviv: Jaffee Center, 1994), p. 59.

⁷⁷ Velikhov, a renowned Soviet physicist, was a member of the Soviet delegation to the nuclear and space talks in Geneva. Quoted in Robert D. English, *Russia and the idea of the West* (New York: Columbia University Press, 2000), p. 217.

⁷⁸ Quoted in Reynolds, *Summits*, p. 355.

⁷⁹ Reynolds, *Summits*, pp. 355–8.

⁸⁰ This theme guides the argument in Wheeler and Considine, 'Reagan may have had trust in Gorbachev'. See also Larson, *Anatomy of mistrust*, p. 216.

The key proposition advanced in this article, that Gorbachev's unilateral frame-breaking moves depended upon the relationship of trust that developed out of his summitry with Reagan, does not, however, negate the second possible materialist argument alluded to above: namely, that Gorbachev had no choice but to conciliate given the Soviet Union's desperate economic situation; that, in view of the weaknesses of the Soviet economy, any Soviet leader at this time would have been compelled to make concessions in the way Gorbachev did.⁸¹ But in response to this argument it is crucial to note that, unlike the first materialist argument (that the Soviet nuclear safety blanket ensured that Gorbachev's frame-breaking moves did not require an increase in Soviet vulnerability), this proposition is actually compatible with the trust argument proffered here. Indeed, it can be agreed with the materialist and Realist scholars who espouse this position that the material pressures exerted by a declining economy were a critical enabling condition of Gorbachev's actions. Certainly, it is very difficult to imagine that he could have secured domestic support for his dramatic accommodative strategy towards the second Reagan administration had the Soviet Union not been so materially weak *vis-à-vis* the United States. Nevertheless, accepting that material and power political factors played a key role in the transformation of US–Soviet relations does not negate the importance that should be attached in any explanation to the enabling ideational conditions of reciprocal security dilemma sensibility and crucially the interpersonal trust built between Gorbachev and Reagan.

The central point here is that the road of conciliation and cooperation was not the only one that the Soviet Union could have taken at this time; Gorbachev's actions were not an inevitable response to the material pressures that the Soviet Union faced. Such a view finds support from Chernyaev, who asserted that arguments premised on the influence of domestic economic pressures, though important, do 'not definitively capture Gorbachev's motives'. These, he argued, drew on a number of elements including his awareness of the potential devastation of nuclear war, his personal moral principles and his belief that no state, and in particular not the United States, would attack the Soviet Union.⁸² Chernyaev does not explicitly mention the role that communicative encounters with Reagan—and the trust that developed out of these—played in leading to a belief that the Soviet Union would not be harmed if it accommodated the United States while Reagan was president. But he does recount in his memoirs Gorbachev telling the Politburo on his return from the Washington summit in December 1987—at which the INF

⁸¹ See William C. Wohlforth, 'Realism and the end of the Cold War', *International Security* 19: 2, 1994, pp. 91–129; Stephen G. Brooks and William C. Wohlforth, 'Power, globalization and the end of the Cold War: reevaluating a landmark case for ideas', *International Security* 25: 3, 2000/2001, pp. 5–53; Stephen G. Brooks and William C. Wohlforth, 'Economic constraints and the end of the Cold War', in Wohlforth, ed., *Cold War endgame*, pp. 273–311.

⁸² Anatoly Chernyaev, oral history transcript, published in Wohlforth, ed., *Cold War endgame*, p. 20. See also Robert D. English, 'The road(s) not taken: causality and contingency in analysis of the Cold War's end', in Wohlforth, ed., *Cold War endgame*, pp. 243–72; George W. Breslauer and Richard Ned Lebow, 'Leadership and the end of the Cold War: a counterfactual thought experiment', in R. K. Hermann and R. N. Lebow, eds, *Ending the Cold War: interpretations, causation, and the study of International Relations* (Basingstoke: Palgrave, 2004), pp. 161–88.

Agreement was signed—that the summit showed ‘how much the human factor means in international politics ... In our age, it turns out, this has the biggest impact on political decisions.’⁸³ It is inconceivable that Gorbachev would have expressed such sentiments had his face-to-face encounters with Reagan resembled the communicative dynamics that his predecessor had experienced with Kennedy in their meeting in Vienna. Rather, both leaders were able to search out and build upon the common interests and values that their countries shared in avoiding nuclear ruin; and this sense of shared responsibility for the security of humankind allowed them to forge a trusting relationship across the East–West divide.

Conclusion

Martin Wight reflected that ‘the diplomatic system is the master-institution of international relations’.⁸⁴ There is no doubt that Wight expanded greatly our understanding of the diplomatic system, and of the qualities necessary for the conduct of responsible statecraft. Nevertheless, what is missing from Wight’s account is an awareness of the possibilities of diplomacy, when conducted at the highest levels of government, in building trust between the leaders of adversary states. Sadly, Wight did not live to see the end of the Cold War, and so it can never be known whether this diplomatic transformation would have led Wight to revise his verdict that diplomacy could at best mitigate the recourse to force in international politics. The end of the Cold War dramatically reduced the fear of nuclear war between the United States and Russia; but since then new nuclear threats have emerged in adversarial contexts (e.g. the conflicts between the United States and a nuclear-armed North Korea, and between the United States and Iran, which is seen in Washington as developing a nuclear weapons break-out capability, as well as between the nuclear-armed states of India and Pakistan) which pose tests of leaders every bit as demanding as those that confronted Ronald Reagan and Mikhail Gorbachev.

The question, then, is whether today’s top-level policy-makers in the nuclear-armed and arming states can learn lessons from the Reagan–Gorbachev experience of trust-building that will assist in de-escalating and transforming their conflicts. The most important lesson from the Reagan–Gorbachev case is that trust begins with communication at the highest level of diplomacy. This lesson is worth pondering in the context of the decade-long stand-off between Washington and Tehran over the latter’s nuclear ambitions, since during that period there has been no high-level bilateral dialogue between US and Iranian policy-makers, let alone any kind of summitry such as Reagan and Gorbachev achieved. As the 1961 Vienna summit shows, direct personal encounters—such as the one between Kennedy and Khrushchev—can be disastrous; but neither of those leaders was able to exercise, let alone operationalize, security dilemma sensibility.

⁸³ Gorbachev’s notion of the ‘human factor’ is discussed in Booth and Wheeler, *The security dilemma*, p. 156, 166–67, 237–41, 251–57.

⁸⁴ Wight, *Power politics*, p. 113.

Empathy of this kind is rare among decision-makers, who are often burdened with mindsets of ideological fundamentalism and hold peaceful/defensive self-images; but, as Reagan and Gorbachev along with their closest advisers showed, these blocks to empathy can be overcome. Nevertheless, even if decision-makers can develop increased security dilemma sensibility, this is not a sufficient condition for the frame-breaking moves that are necessary to transform a conflict. Security dilemma sensibility is necessary if a meaningful dialogue is to take place, and it might lead, perhaps through face-to-face encounters that deepen empathy of this kind, to first, second, and maybe even third steps which begin a process of limited de-escalation. However, if these communicative encounters do not at the same time build trust between leaders (and their advisers), then it cannot be expected that this step-by-step approach will lead political leaders to send the potentially game-changing costly signals that were so transformational in the US–Soviet case.

This case inevitably raises questions about how important power asymmetries are in creating the conditions under which one state decides to accommodate an adversary, as well as the importance of the stronger state reciprocating (ideally positively, but at least negatively) and not exploiting the weaker state's accommodative strategy.⁸⁵ At the same time, the evidence from this case suggests that more research is needed on the types of communicative encounter that can promote trust between policy-makers in adversarial relationships. Such research also needs to explore the conditions—ideational and material—under which interpersonal trust of this kind can be a key enabling condition of policy-makers making the frame-breaking moves that are critical to the process of diplomatic transformation. The challenge then is to better understand how the qualities of empathy, mutuality and mutual respect—which Wight saw as the essence of a sophisticated conception of diplomacy, and which were so crucial to the trust that grew between Reagan and Gorbachev—can be promoted between today's leaders grappling with nuclear security dilemmas.

⁸⁵ The importance of power asymmetries in explaining why states pursue strategies of accommodation is the central argument in Charles A. Kupchan, *How enemies become friends: the sources of stable peace* (Princeton, NJ: Princeton University Press, 2010).