

## **It takes all types: social psychology, trust, and the international relations paradigm in our minds**

BRIAN C. RATHBUN\*

*School of International Relations, University of Southern California, Los Angeles, CA, USA*

The insights of social psychology are not thoroughly integrated into international relations theory, yet social psychology has much to offer. Social psychology provides a conceptualization of a number of varieties of trust – moralistic, strategic, and generalized – and their opposites that implicitly drive the logic of major works of international relations. It also reveals the empirical presence of a number of different types of trusters who make different assumptions about the trustworthiness of others and consequently show markedly different propensities towards cooperation. The rough correspondence between these different ‘social orientations’ and the logics of the three approaches of structural realism, neoliberal institutionalism, and constructivism suggest that individuals carry a crude paradigm in their minds. Metatheoretically, the implication for international relations theory is that scholars capture a part but not the totality of world politics, the behavior of those who trust (or do not trust) in a particular way that matches the logic of their paradigms. Theoretically it suggests a research agenda at multiple levels of analysis, utilizing all of the types of trust and trusters. I review the work of others that offers some preliminary evidence for its plausibility, suggest some hypotheses of my own, and address potential theoretical objections.

**Keywords:** social psychology; trust; international relations theory; cooperation; foreign policy

Dramatic innovations in international relations theory, perhaps more so than other subfields of political science, generally involve borrowing from outside the discipline. The two major revolutions in international relations theorizing in recent decades have involved the importation of models, analogies, and insights from outside. Rationalism draws inspiration from economics, constructivism from sociology. This borrowing has been so fecund because it led scholars to fundamentally different conclusions than realism about the nature of international politics. Robert Keohane’s (1984) ‘neoliberal

\* E-mail: brathbun@usc.edu

institutionalism' was ground-breaking because it pointed to the possibility of contingent cooperation even on the part of purely egoistic states in anarchy. Alexander Wendt's (1999) claim that international politics is socially constructed, was theoretically transformative because it highlighted the cultural basis of realist power politics and therefore the possibilities for sustained and deep cooperation through the formation of collective identity. Rationalism and constructivism tell us much more than this of course, but it is largely on this basis that they are regarded as powerfully new approaches that changed our understanding of world politics.

The importation of social psychology has not yielded similar success. Social psychology is sometimes called in to contribute a micro-foundation for realism or constructivism but does not provide an independent contribution. For instance, the fundamental attribution error, the tendency to attribute hostile actions to disposition rather than structural circumstances exacerbates the security dilemma (Jervis, 1976; Mercer, 1996; Tetlock, 1998). In-out group distinctions inhibit signaling intentions in situations of hardened 'images' of the other (Jervis and Harvard University, Center for International Affairs, 1970; Jervis, 1976). Constructivists point out that an expanded sense of group identity can lessen threat perception and lead to enhanced cooperation even in an anarchic setting (Cronin, 1999; Rousseau, 2006). Cronin notes the similarities to social identity theory, but it is unclear what social psychology's unique role might be in accounting for important outcomes in international politics.

Although it is not the only avenue, social psychology can have an impact on international relations theory akin to that of economics or sociology if it can tell us something new about trust. This is because many of the core differences between major theorists of international relations stem from often implicit assumptions about how states and statesmen respond to what social psychologists call 'social uncertainty', a lack of knowledge about others' intentions (Luhmann, 1979; Yamagishi, 2001). Social uncertainty is particularly salient in international relations, given the anarchic nature of world politics.

Trust is the belief that one's interests will not be harmed when placed in the hands of another. But within this broader definition, social psychologists highlight the importance of different varieties of trust that are present but underconceptualized in international relations theory. I argue that they are necessary for distinguishing the logic of key international relations theorists from one another. Eric Uslaner (2002), a political scientist, has developed a typology of trust largely on the basis of psychological findings. *Strategic trust* is based on the accumulation of information about specific others' intentions. *Moralistic trust* is an assessment of the trustworthiness of others based on a conclusion about the overall integrity and character of the potential partner. Moralistic

trust can either be generalized or particularized. *Generalized* trust is a general belief that most others can be trusted. *Particularized* trust is the belief that a specific other is inherently trustworthy.

Evidence of the presence of generalized trust in international relations theory might be somewhat rare, but the lack of generalized trust, what I call fear, is the motor that drives Waltz's structural realism. It was Keohane's replacement of fear with strategic trust that led him to conclude that realism's pessimistic picture of international relations was misplaced, even if states act purely egoistically. Contingent cooperation becomes possible on the basis of increased information. The possibility that moralistic trust might develop is necessary for the transition from an environment of contingent cooperation among egoists described in rationalism to a deeper and more sustained cooperation foreseen by Wendt's notion of constructivism.

Social psychology, therefore, provides conceptual tools to help better specify the different logics of these three different approaches. It identifies the micro-foundations of these macro-paradigms. It takes all types of trust to understand international relations theory. But it offers more than that. Social psychologists have identified a number of different 'social orientations', ways in which individuals approach their interactions with others that reduce to different levels and kinds of trust. Just as there are different types of trust, there are different types of trusters. In experiments, some subjects generally fear others' intentions, some make case-by-case decisions based on feedback from others, and others believe that others are generally trustworthy. Individuals vary dispositionally in their level of trust and therefore the kind of trust they practice.

This observation has both a potentially great metatheoretical and theoretical impact. In terms of the former, these social orientations show a remarkable similarity to the logic of the three works described above. If individuals employ those orientations as a general heuristic in decision-making, we might say that they carry an international relations paradigm in their minds. Given that important international relations theorists rely on implicit assumptions about how and whether decision-makers trust, they capture accurately but might be largely limited to explaining the actions of a particular subgroup of individuals who approach social uncertainty in a particular way. They might be mistaking structure for disposition. They draw broader macro-level conclusions from what is only the externalization of micro-processes. In the extreme form of the argument, since there are different types of trusters simultaneously operating, there is no overall macro-logic to international relations that applies everywhere or even most places. There is only the belief on the part of those with different social orientations in different macro-logics of the environment in which they are operating.

Theoretically, a focus on social orientation and dispositions to trust allows us to account for monadic, dyadic, and systemic features of foreign policy and international relations. Social psychology, the varieties of trust and social orientation are implicit in and essential for explaining key phenomena of interest in international relations such as the democratic peace and the higher conflict propensity in rivalries. They provide an alternative logic for the evolution of different zones of peace and war. They invert the conventional wisdom concerning the foundations of international cooperation.

In the sections that follow, I first explain Uslaner's typology of trust and tie it into international relations theory, claiming that differences among Waltz, Keohane, and Wendt emerge from implicit reliance on different types of trust. Contrary to Hoffman's (2006: 4) claim that trust is not a major element in the logic of international relations paradigms, I argue that is perhaps central. The *explicit* study of trust in international relations is relatively new (Larson, 1997; Hoffman, 2002) and has been dominated by a particular, rationalist variety of trust (Kydd, 2005) with some notable exceptions (Larson, 1997; Mercer, 2005). However, trust has been the implicit foundation of debates about the character of state interaction since discussions about the nature of the interstate system began. I focus predominantly on these three theorists, as their work has been central for creating paradigms that have greatly impacted the study of world politics. They might not have been the first to work in their paradigms, and others in those traditions might have subsequently refined them. There are variations within paradigms as well, but all must come to terms with those three theorists.

Second, I review the social-psychological literature on social orientation, much of which rests on the finding that individuals approach the same problems of cooperation differently, largely based on how and whether they trust others. I note the similarities between these three theorists' logic and the three social orientations revealed by social-psychological experiments. Realism, rationalism, and constructivism each have an analogue in a particular type of truster.

In the third section, I discuss the theoretical consequences for international relations theory if the world is in fact populated by individuals whose actions are driven, albeit crudely, by the particular international relations paradigm in their mind. I lay out a social-psychological research agenda with trust and social orientation at its core, reviewing some tentative empirical evidence consistent with the main premise so as to demonstrate its initial plausibility. I address potential theoretical objections concerning problems of external validity, units of analysis, ecological fallacy, and levels of analysis. I do not pretend that this review in any way firmly establishes the superiority of a social-psychological approach, but it lays out the basis for a discussion and a roadmap for assessing its utility *vis-à-vis* existing approaches.

## The varieties of trust and international relations theory

An uncontroversial and encompassing definition of trust is the belief that one will not be harmed when his or her fate is placed in the hands of others. Trust entails a combination of uncertainty and vulnerability (Larson, 1997: 19; Sztompka, 1999; Kramer *et al.*, 2004: 25; Hardin, 2006: 29; Hoffman, 2006: 17). The truster never knows for certain what the trustee will do; the trustee might have a desire to violate that trust. If the intentions of the other are known for sure, trust is not involved, but this is rarely the case in social settings. Trust leaves the truster exposed and vulnerable to potential opportunism.

Nevertheless, trust is often essential to cooperation and reaching outcomes that leave all sides better off. In the context of cooperation, trust is the belief that others will reciprocate. Kydd (2005) has put it in game theoretic terms. Trust is the belief that another has assurance game rather than prisoner's dilemma game preferences, that he or she prefers mutual cooperation to exploiting and suckering others (2005: 6–7). It is crucial in situations of 'mixed-motives', or 'collaboration games' in which there is an incentive to cooperate, but also to defect if others cooperate (Olson, 1971; Stein, 1982; Keohane, 1984).<sup>1</sup> If two actors, player A and player B are placed in the same strategic situation with the same utility function, but player A is more optimistic about the intentions of others, player A will be more willing to cooperate.

This definition of trust is agnostic on the basis on which trust is conferred and the extent to which trust extends. Trust or distrust might be generally applied to all our interactions, or it might be situation-specific. Trust or distrust might be based on information about the interests of others, or on an assessment of other's honesty and integrity. There are a number of varieties of trust, and we can see all of them in international relations theory. Indeed, we need to draw them out in order to get at the core differences between major theorists.

### *Strategic trust*

Arguably the most prominent notion of trust in political science and international relations is what Uslaner (2002) calls 'strategic trust'. In this understanding of the concept, actors trust on the basis of beliefs about others' interests, cooperating when they believe that others' interests 'encapsulate' their own (Hardin, 2006). Trust is a belief that potential

<sup>1</sup> Other factors, such as scarcity, also inhibit cooperation but through a different mechanism. Conflict in scarce situations might reflect a lack of resources rather than uncertainty about intentions.

partners have a self-interest in cooperation, generally an incentive in building or sustaining a long-term, mutually beneficial relationship. Tyler (2001: 287) calls this approach 'calculativeness' as it envisions a cold, sober evaluation of the advantages and disadvantages of trust. Whether to trust, argues Hardin (2006: 22), is a 'matter of prudential assessment, not moral choice'. Actors make such judgments on the basis of information. Hardin writes that the two central elements of the rationalist account of trust are the incentives of the trusted to honor the trust, and knowledge to justify the truster's trust (Yamagishi and Yamagishi, 1994: 139; Hardin, 2006: 44). There is of course also strategic distrust, a belief that specific others are trustworthy based on information, albeit incomplete, about their interests.

This is a highly structural account of trust. Strategic trust develops in situations in which actors have an incentive to honor their agreements. It has nothing to do with the attributes of the individuals cooperating. The implication of strategic trust is that even if there is an ongoing relationship of mutual self-interest in one area, those involved draw no general lesson about the moral character of the other into other domains. The relationship might produce incentives for potential partners to honor their commitments in other fields but primarily through the mechanism of a linkage with previously existing areas of cooperation. There is 'no moral force' in strategic trust (Uslaner, 2002: 22). As Uslaner writes, the etymology of strategic trust is 'A trusts B to do X' (2002: 21).

International relations scholars have recently begun utilizing this conception of trust explicitly (Kydd, 2005). However, it has a much longer pedigree. This is the kind of trust utilized, often only implicitly, in rationalist work in international relations. It provided the foundation of what was known as 'neoliberal institutionalism', an approach pioneered by Robert Keohane and others but which has now reached beyond studies of international cooperation into all elements of state interaction. States are unsure as to whether others will maintain their end of the bargain, uncertain whether others will keep their word. Cheating, which in collaboration games yields a 'sucker' payoff, is a key concern (Stein, 1982; Grieco, 1988; Mearsheimer, 1994). Incomplete information stymies potentially rewarding collaboration.

In keeping with this conception, rationalist work on trust focuses almost exclusively on changing the structure of the situation to generate more cooperation, such as by iterating the game or by reducing the number of players in the game to prevent free-riding. By institutionalizing cooperation, international organizations can facilitate this process. Through the provision of information, trust of particular others becomes possible. As shadows of future are lengthened, linkages are created, and

reputations put at a premium, the level of trust between egoists increases as their interests come to encapsulate one another. All of these mechanisms focus almost exclusively on creating incentives for trustworthy behavior through adjustments to the strategic environment (Axelrod, 1984; Keohane, 1984; Oye, 1985).

*Moralistic trust: generalized and particularized*

Trust, however, can also be based on beliefs about the honesty and integrity of potential partners. This is the other basis of trust, what is called *moralistic* trust. Messick and Kramer (2001: 91) are describing moralistic trust when they define trust as making decisions as if others will abide by ordinary ethical rules, most important of which are truth-telling and not harming others. This is a ‘social conception of trust’ in which trusters believe that intentions and behavior reflect traits of the trustee, rather than the situation (Tyler and Degoe, 2004: 332). Hoffman (2002: 20–22) calls it ‘fiduciary trust’. It is based on ‘an implicit theory of personality’, a belief that others have consistent personalities and traits that do not vary by situation (Sztompka, 1999: 75; Mercer, 2005: 95). When we trust, we are making judgments about the character and integrity of partners (Yamagishi and Yamagishi, 1994: 132; Larson, 1997: 22; Booth and Wheeler, 2008: 243–245). When moralistic trust is present, the truster does not need to restrict that trust of another to transactions or cooperation of a particular kind, because while interests vary across situations, character does not (Kramer *et al.*, 2004: 367). Moral attributions allow us to draw broader conclusions about the trustworthiness of others without the need to constantly collect information in every new situation or insist on specific reciprocity. It provides a deeper and more durable foundation for cooperation.

Moralistic trust varies in scope. It might apply to a specific other or most others. If moralistic trust extends more broadly so as to constitute a belief that not all, but most others can be trusted, we can speak of *generalized* trust. The etymology of generalized trust is ‘A trusts’ or ‘A is trusting’ (Uslaner, 2002: 4). Generalized trust rests on a general belief in the honesty and benevolent character of others. When Rotter defines trust as the ‘generalized expectancy held by an individual that the word, promise, oral or written statement of another individual or group can be relied on’, he is talking about generalized trust (Rotter, 1980: 1).<sup>2</sup> Generalized trust is a belief in human benevolence (Cook and Cooper, 2003: 215).

<sup>2</sup> Although related, generalized trust is not the same as risk propensity. A decision about whether to cooperate in a situation of strategic interdependence involves both an evaluation of the potential payoffs *and* an assessment of the intentions of others (Hardin, 2006: 44). *Ceteris*

Generalized trust must, by definition, be moralistic, as we cannot collect information about the interest of all prospective interaction partners. As Sztompka (1999: 70) writes, ‘Trustfulness may incline people to grant trust, and suspiciousness to withhold trust, quite independently of any estimate of trustworthiness. The origins of trustfulness or suspiciousness are not epistemological. They have nothing to do with knowledge about the partners of future engagements...’ Mercer (2005: 95) writes that it involves ‘certainty beyond observable evidence’. Hardin (2001: 5), somewhat pejoratively, calls it a ‘non-cognitive’ approach. Better stated, it is ‘non-calculative’ in the utilitarian sense. It is most certainly cognitive in that it is based on a set of beliefs about others.

The presence and therefore the relevance of generalized trust might not be immediately obvious or prevalent in international relations, but its absence is at the heart of Waltz’s version of structural realism. The opposite of generalized trust is fear, the presumption that most others cannot be trusted.<sup>3</sup> Fear has a long association with classical realism, but structural realism is generally understood as relocating realism’s grounds for pessimism about cooperation in international politics from the ‘first image’, what classical realists regarded as human nature, to the systemic level. However, Waltz needs fear, and sneaks it through the back door. As a number of analysts have noted, all major security dilemma theorists, going back to Butterfield and Herz, rely on fear to make their arguments go (Booth and Wheeler, 2008: Ch. 2; Tang, 2008). Not all self-described realists rely on the assumption of fear, but many do (Rathbun, 2007b; Booth and Wheeler, 2008; Tang, 2008).<sup>4</sup>

*paribus*, cooperation is more likely (1) the more players value the gains from cooperation relative to the losses of defection and (2) the more players believe that others will not exploit them. This first element captures the utility function. Only the second captures the level of trust. If placed in the same structural situation, some player A values the same payoff for mutual cooperation more than another player B, and player A will be more likely to take a chance on cooperation, even if the two have similar assessments of the probability of opportunistic behavior that would frustrate reaching that goal (i.e. their degree of trust). By virtue of player A’s higher evaluation of the payoffs for cooperation, he or she is more willing to take a chance on cooperation at higher levels of the risk of opportunism (Kim and Bueno de Mesquita, 1995; Bueno de Mesquita *et al.*, 1997). This is known as risk propensity and applies to a lower evaluation of the payoffs for mutual cooperation. Trust is risky, since given uncertainty, there is always some chance of defection. But trust is not the same as risk propensity (Luhmann, 1979; Hoffman, 2006: 21–22, 25).

<sup>3</sup> While it is possible to equate fear with a belief that particular others have malign intentions as does Tang (2008), I reserve the term to capture a general pessimism about others, preferring the term ‘strategic distrust’ to describe the former.

<sup>4</sup> Indeed fear plays a greater role in Waltz’s ‘defensive’ than in Mearsheimer’s ‘offensive’ realism. In the latter, conflict is driven not by uncertainty but on scarcity. In this version of realism, states know that others are out to get them; it is just a question of when. For a good review, see Taliaferro (2000).



In *Theory of International Politics*, Waltz (1979) maintains that the possibility of opportunism given uncertainty about intentions, combined with the lack of any external protector, is enough to induce recurrent patterns of balancing power, the security dilemma, and war. Anarchy and uncertainty preclude the creation of trust and international cooperation. States engage in self-help, manifested in the accumulation of power against potential adversaries, real or presumed, through military build-ups and alliance formation. Waltz (1979: 105) maintains that ‘the condition of insecurity – at least, the uncertainty of each about the other’s future intentions and actions – works against their cooperation’.

Yet anarchy and uncertainty alone themselves do not generate what Glaser (1994) calls the ‘competition bias’ in Waltz’s neorealism. Only if actors react to uncertainty in a particular way, by assuming the worst, do the dynamics that Waltz predicts ensue (Booth and Wheeler, 2008: 22–23; Tang, 2008). There is a presumption that states believe others have malign aims, or must be treated as if they do. Brooks (1997) calls this a ‘possibilistic’ logic. In Waltz’s realism, states distrust not specifically, but generally. They have a general inclination not to trust. This working assumption is insensitive, perhaps impervious, to revision through accumulation of information, and precludes objective, case-by-case assessments of others’ trustworthiness. The presence of fear is evident in the work of those neorealists who argue that states respond to the problem of uncertain intentions, that is whether others should be trusted, not by thoroughly and carefully parsing the evidence case-by-case in search of information about their objectives, but by balancing power (Grieco, 1988: 500). According to many neorealists, states do *not* carefully evaluate another state’s intentions because a state can never have enough certainty to fully let down its guard, and intentions can change (Glaser, 1994; Taliaferro, 2000: 144–145; Kydd, 2005: 15). The implicit mechanism of fear and the consequent focus on power over information explains why many structural realist scholars have dedicated little attention to issues of signaling and screening, as noted by Glaser (1994) and Booth and Wheeler (2008: 27). Power, not past action, is how states gauge intentions and judge credibility (Press, 2005).<sup>5</sup>

It was the implicit replacement of fear with strategic distrust that led Keohane (1984, 1989), in particular, to such different conclusions regarding the prospects for international cooperation than Waltz. By removing the

<sup>5</sup> Material features of the environment might ameliorate the security dilemma by making states less vulnerable, but not more trusting (Jervis, 1978; Walt, 1987; Van Evera, 1999; Taliaferro, 2000). Brooks (1997) calls these refinements of the conception of military capacity ‘probabilistic realism’, in which statesmen assess the likelihood of conflict based on the *ability* of states to inflict harm. But fear of intentions is still there; it is simply covered up.

presumption of a general fear of others' intentions that is not responsive to information (and generates the concern for 'relative gains'), states are able to collaborate and might reach mutually beneficial outcomes even in an anarchic environment. Nevertheless, it would be incorrect to label neoliberal institutionalists optimistic or idealistic about the character of international relations; rather they are agnostic. They believe in the contingent possibilities for international relations in particular situations when states have common interests. Trust is specific to particular situations and based on information. Neoliberal institutionalists envision states proceeding skeptically, careful not to be burned, but not to the degree that they miss opportunities for potentially rewarding cooperation. In keeping with their strategic conception of trust, they maintain that whether to trust or fear others is a function of the situation, not a general predisposition on the part of states.<sup>6</sup>

Generalized trust (or fear) is a belief held about the overall trustworthiness of others as a whole. It is a quality of the truster, not the relationship or even the characteristic of the specific target of distrust or trust (Sztompka, 1999: 97). In contrast, particularized trust characterizes a relationship in which one individual or group believes that a specific other or group is inherently trustworthy in a moralistic sense. The etymology of what Uslaner (2002: 28) calls *particularized* trust is 'A trusts B'. What we could call particularized fear would be the opposite. Unlike generalized trust, particularized trust is relational in nature, rather than dispositional, even though both are moralistic varieties of trust. Particularized trust is dyadic, perhaps created and certainly sustained by opposing identities. We trust others like us and fear those who are different. Nevertheless, these two types of trust are likely related. It is likely that relationships characterized by particularized fear (trust), for instance,

<sup>6</sup> A number of self-described realists have also stressed that approaching international politics fearfully might needlessly leave a state worse off by generating arms races and instability (Glaser, 1994; Kydd, 1997). Overestimating hostility is often just as dangerous as underestimating it. Case-by-case analysis of others' intentions is necessary. Drawing on the insights of Thomas Schelling (1966) and Robert Jervis (1978), Charles Glaser (1994) and Andrew Kydd (2005) have each attempted to modify Waltz's structural realism by incorporating the processes of signaling and screening made impossible with an assumption of fear. This is in keeping with what Lake and Powell (1999) call the 'strategic choice' approach, which identifies information as one of the two core features of the environment. If credible information indicates that another state might be willing, and there are possible mutual gains, it is often rational to cooperate. Glaser (1994) writes of 'contingent' realism. Trust is built strategically through the accumulation of evidence about others' intentions. Kydd (2005: 14) calls his approach 'Bayesian realism', which eliminates the premise of mistrust as a 'constant' or 'permanent state' that makes it pointless in Waltz's structural realism to distinguish between aggressor and peaceful states.

are more likely to develop among individuals with a more fearful (trusting) disposition.

Particularized trust is essential for Alexander Wendt's (1992, 1999) version of constructivism. Wendt's constructivism does not see conflict as an inevitable or general feature of international relations but does maintain that a 'culture' of anarchy marked by fear might arise. States are not doomed to repeat forever the cycles of conflict and acrimony as they are in realism. They can embark on a more cooperative path and change the dynamics of their interaction. Trust is the mechanism. It is what Wendt (1999: 358) calls the 'fundamental problem of collective identity formation'. While the core variable in constructivism is generally thought to be identity, trust is logically prior to the formation of a common identity (Adler and Barnett, 1998: 46). One cannot feel part of a community composed of others whom he or she believes might mean him harm. He or she must believe that his or her needs will be respected.

Trust for Wendt is created through interaction. Acts of concession or generosity can create a reinforcing spiral of cooperation. Trust-building might begin through a rationalist process of signaling and conveying information, but the process is more transformative, allowing for a redefinition of self and other from adversary to partner to friend. However, for this process to take hold, states must overcome their fear of 'being engulfed... by those with whom they would identify' (Wendt, 1999: 357). One way to provide this assurance is through some sort of institutional constraint, as would be the case in both structural realism and rationalism. But Wendt argues this is not sufficient. States must '*believe* that others have no desire to engulf them, nor would do so out of self-interested opportunism... [T]hen it will be easier to trust that in identifying with them, their own needs will be respected, even in the absence of external constraints' (1999: 359) (emphasis added). If states trust other states in an environment in which there is no institutional sanction, even if the latter might have a self-interest in defection, trust rests on a judgment about the other's integrity and honesty rather than the situational incentives of the time. Trust is not a function of the situational exigencies of the moment, but rather predicated on a belief that others have a particular character, that they are inherently trustworthy. The kind of trust operating in this instance must be based on a belief that others feel morally compelled to reciprocate cooperation – that is, moralistic trust.

Of course, Wendt's constructivism does not deterministically assume that this process will occur and is equally comfortable describing a fearful dynamic overtaking state interactions. It does, however, stand somewhat alone in offering an account for the potential development of deeper cooperation based on moralistic trust. Wendt's use of this moral variety of

trust leads it to the conclusion that the possibilities for cooperation in international relations are far more profound than those foreseen by rationalism, with its focus on strategic trust.

### **Social orientation and international relations theory**

The insights of social psychology give us a better sense of the logic of these three international relations paradigms through a better conceptualization of, and more variety in, trust. However, the impact on international relations theory is more than conceptual. Social psychology grounds these concepts in findings about how individuals behave in the real world. Research on 'social orientation' has revealed the presence of a number of different types of trusters who act very differently in similar structural settings on the basis of beliefs about the trustworthiness of others. Only non-structural factors or attributes of the players, what social psychologists call dispositional factors, can explain such variation within a particular incentive structure. The relevance for international relations theory is that these different social orientations correspond remarkably to the logic of the theorists reviewed above.

Social orientation research began with the observation that individuals transform objective decision matrices given by researchers into 'effective' decision matrices that reflect their own subjective weights of particular outcomes (Kelley and Thibaut, 1978; McClintock and Liebrand, 1988; Parks, 1994). Pioneered by Morton Deutsch (1960), Messick and McClintock (1968) and others have built upon the premise to identify a number of distinct social orientations (McClintock, 1972; Kuhlman *et al.*, 1986). Most commonly used to classify subjects in experiments are the competitive disposition, the social orientation that maximizes the difference between own and others' payoffs; the individualistic disposition, the orientation that simply maximizes individual payoffs; and the cooperative disposition, the orientation that maximizes the combination of own and other payoffs. In international relations terms, cooperators maximize joint gains; competitors maximize relative gains; and individualists, absolute gains.<sup>7</sup> The most common way of measuring orientation is through the method of 'decomposed games'. Before the experiment, subjects are presented with a variety of hypothetical own–other outcomes and asked to rank their preferences. Analysts have found significant

<sup>7</sup> Less commonly used are the aggressive social orientation, that which aims to minimize others' payoffs regardless of the foregone gains for oneself and altruism, the maximization of others' payoffs. Often competitors and individualists are collapsed into a common category of 'pro-selves' while altruism and cooperators are assembled together as 'pro-socials'.

numbers of competitors, cooperators, and individualists in their research samples, and consistent links between these orientations and subsequent behavior in experimental settings (Kuhlman and Marshello, 1975; Kuhlman and Wimberley, 1976; McClintock and Liebrand, 1988; Parks, 1994; De Cremer and Van Vugt, 1999). More cooperative subjects cooperate most, competitors the least, with individualists somewhere in between.

These game transformations occur largely on the basis of beliefs about the trustworthiness of others (Kuhlman and Wimberley, 1976; Kuhlman *et al.*, 1986). Researchers consistently find a strong link between the measurement of social orientation of subjects before experiments, the behavior of subjects during experiments, and their expectations before experiments begin about how others are likely to behave. Kuhlman and Wimberley (1976) report that cooperators are the most optimistic predictors of cooperation, individualists less so, and competitors the lowest. The more trusting individuals are, the more they cooperate in 'social dilemmas' such as commons and public goods problems, a finding extensively corroborated beyond the social orientation literature (Tyszka and Grzelak, 1976; Alcock and Mansell, 1977; Dawes *et al.*, 1977; Marwell and Ames, 1979; Messick and Brewer, 1983; Messick *et al.*, 1983; Yamagishi and Sato, 1986; Brann and Foddy, 1987; McClintock and Liebrand, 1988; Yamagishi and Cook, 1993).

Using the same logic as Kelley and Stahleski (1970), Kuhlman *et al.* (1986) hypothesize that competitors and cooperators have different views of the world. Competitors see a dangerous environment populated by other competitors. This amounts to a lack of generalized trust, or fear, and to a zero-sum characterization of social interaction based on expectations about others. Competitors compete because they believe they must. Their level of cooperation with both cooperators and other competitors is low and their behavior is not responsive to information or the behavioral cues of other players (Van Lange and Kuhlman, 1994), or at least is less so (Yamagishi, 2001). Because of their general fearfulness, competitors have been found to be the worst judges of actual intentions in experimental settings (Maki and McClintock, 1983). Empirically, competitors find it more difficult to use communication to build trust (Cook and Cooper, 2003: 212). In a landmark study, Kuhlman and Marshello (1975) found that in a repeated prisoner's dilemma game, competitors showed consistently high levels of defection regardless of the strategy of the other player, which varied among three pre-programmed strategies of 100% cooperation, 100% defection, and tit-for-tat. They do the same in assurance games as well, even though the game structure is more conducive to cooperation (McClintock and Liebrand, 1988). There are clear parallels between this tendency and the theoretical premises of Waltzian

realism reviewed above, which holds that decision-makers are not sensitive to information about intentions, but rather assume the worst. In both the competitive type and the Waltzian paradigm, decision-makers lack generalized trust.

Cooperators, in contrast, are cooperative because they have a more optimistic view of the world. They are generalized trusters. This leads them generally to begin an interaction with cooperation (McClintock and Liebrand, 1988). In the Kuhlman and Marshello study, cooperators cooperated at high levels, and much higher than individualists and competitors, *vis-à-vis* opponents playing a 100% cooperation or a tit-for-tat strategy. Cooperation on the part of more trusting individuals surges even further when they are given positive feedback, whereas that of competitors does not change (Alcock and Mansell, 1977; Yamagishi, 2001). Generalized trusters are relatively more open to signs of the intentions of others (Rotter, 1980). In experiments, cooperators are the best predictors of the intentions of others (Maki and McClintock, 1983). More trusting individuals have been shown to exhibit higher degrees of 'social intelligence', an unfortunate term that captures the effort made to judge the intentions of others (Yamagishi, 2001).

Research indicates that cooperators are operating on the basis of moral considerations, as would be expected from generalized trusters, as generalized trust is a kind of moralistic trust.<sup>8</sup> They do not exploit a 100% strategy of cooperation in a repeated prisoner's dilemma, indicating a degree of moral restraint. A number of studies reveal the presence of the 'might vs. morality effect'. Cooperators attribute behavior on the part of others to moral characteristics, holding cooperators to be honest and defectors to be dishonest. They expect that honesty will have a great effect on the level of cooperation of others, and their own level of cooperation increases much more sharply against players identified as honest than does that of competitors (Liebrand *et al.*, 1986; Van Lange and Kuhlman, 1994).

The belief that morality matters lead cooperators to believe that they can elicit cooperation, even in staged, one-shot games in which there is no long-term gain from cooperation (Hayashi *et al.*, 1999). In repeated games, they are more willing for some period of time to engage in 'compensatory trust', putting up with some defections in order to establish a cooperative relationship (Kramer *et al.*, 2004: 375). Cooperators' willingness to expose themselves to opportunism in the absence of any institutional constraint on the basis of an assumption of reciprocity that

<sup>8</sup> Recall that while generalized trust is always moralistic, not all moralistic trust is generalized but rather might apply to specific others.

must be moral in nature indicates that cooperators practice the kind of trust that Wendtian constructivists maintain could potentially transform the nature of international relations. They are, in Booth and Wheeler's (2008) terms, 'transcenders'.

Competitors, in contrast, attribute behavior to strength characteristics, viewing cooperators as weak and competitors as strong. It appears that competitors, believing that others are untrustworthy, come to see their surroundings as an amoral environment in which only strength matters. The only way for them to explain cooperation on the part of others is a lack of choice due to weakness. The parallel between realism and a competitive social orientation is again clear. The studies bring to mind Thucydides dictum about the weak doing as they must.

Kelley and Stahleski (1970) argue that competitors, by expecting the worse and not responding to informational cues, create a self-fulfilling prophecy. Due to their general pessimism and fear, competitors tend to begin games by not cooperating. Even cooperators eventually defect against competitors, who are in their eyes immoral. Their preference for joint gains is contingent on reciprocity. Cooperators are not altruists (Kelley and Stahleski, 1970; Kuhlman and Marshello, 1975; Kuhlman and Wimberley, 1976; McClintock and Liebrand, 1988). When competitors' defection eventually elicits defection on the part of even cooperators, it reinforces their view that the world around them is threatening and that others cannot be trusted (Kuhlman and Marshello, 1975; Kuhlman and Wimberley, 1976). They mistake cooperators for defectors. International relations scholars will undoubtedly recognize in this phenomenon the structural realist logic of the 'security dilemma' (Jervis, 1978). Competitors are what Booth and Wheeler (2008) call 'fatalists'. In foreign policy terms, they would be classified as Waltzian realists.

There are other parallels between the three social orientations and international relations paradigms. The only solution for competitors to the problem of social uncertainty is enforcement, a *Leviathan*. To cooperate, less trusting individuals require some form of coercive protection. Institutions are the primary mechanism to allow 'cooperation without trust' (Sztompka, 1999: 87; Ostrom, 2003: 19; Kramer *et al.*, 2004: 383; Cook *et al.*, 2005). Yamagishi (1988) found that those with lower levels of generalized trust only increase their cooperation sharply if there is a sanctioning mechanism against defection. The effect of such an enforcement body on cooperators' behavior is much less pronounced, as they are less pessimistic in the first place. Again, the link in outlook between competitors and Waltzian realism is palpable.

Individualists lie somewhere in between competitors and cooperators in all regards – their overall levels of cooperation, their expectations about

others, and their attributions of other's behavior to might or morality. This is not surprising since individualists are, in international relations terms, rational egoists. As strategic trusters, they do not have any particular tendency to trust or distrust independent of the structural situation and the feedback they are receiving from other players. McClintock and Liebrand (1988) find that individualists, and only individualists, respond consistently in terms of their level of cooperation to variations in game structure, cooperating more in assurance games and defecting more in prisoner's dilemma games. As their orientation resembles the asocial egoistic assumptions game theorists make, they are the most inclined to behave as these analysts would predict. Rationalists are contingent cooperators, or 'mitigators' of the trust problem (Booth and Wheeler, 2008). In the Kuhlman and Marshello study (1975), individualists largely cooperated with tit-for-tat, defected against 100% defection, but feeling no moral constraint (unlike cooperators), exploited 100% cooperation. Whereas the behavior of cooperators is almost identical in prisoner's dilemma and assurance games, individualists are much more likely to defect in the former but not in the latter against a strategy of 100% cooperation, as defection pays for a rational egoist in a prisoner's dilemma.

### **Social orientation at multiple levels of analysis: the social-psychological research profile**

The parallels between findings in social psychology and longstanding concepts in international relations theory strongly suggest the relevance of the former for the latter. Individuals seem to carry around one of the international relations paradigms in their mind, although they use it as a general heuristic in approaching social interaction in general. This insight has a potentially enormous metatheoretical impact. Given the close way in which the social orientations map roughly onto the three paradigms of structural realism, rationalism, and constructivism (or at least their leading theorists), it might be that international relations paradigms (while acknowledging the variation within and the fluidity between them) mistake structure for disposition. It could be, for instance, that those who approach the world as a competitive place in which others cannot be trusted are responding to objective circumstances captured by realism. However, it could also be that this is merely the manifestation of a particular type of trust by a particular type of truster, and taken as evidence for realism. What appears to be structure is actually disposition, the externalization of a social orientation defined by a belief that crudely resembles an international relations paradigm. That disposition of course might consist of a belief in the presence of a systemic and structural



character of international relations specifically, and human relations generally. In other words, realists mistake a macro-logic for a micro-logic. The same could be true of rationalists and constructivists as well.

This is the maximalist position, that it is ‘dispositions all the way down’, to turn Wendt’s phrase. The epistemological goal of international relations theory shifts away from looking for an eternal nature of international affairs – be it antisocial, asocial, or social – as international relations would only have a nature if the distribution of types skews in one direction or another. The key question is not what paradigm best explains world politics, but which paradigms are particularly likely to guide the decision-making of particular types of individuals? Who are the neorealists, the rationalists, and the constructivists in the real world? Stated another way, who are the moralistic trusters, the strategic trusters, and the fearful?<sup>9</sup>

Others in the psychological tradition have gotten close to this conclusion. Operational code analysts noted the importance of core beliefs such as the nature of the international system, noting for instance the similarities between realism and those whose worldviews rest on pessimistic assumptions about human nature. And they demonstrated how these codes were significant for foreign policy choices (George, 1969; Walker, 1983; Larson, 1994). However, they never drew out the significance for the study of international relations, and as a result were never taken seriously enough by systemic and structural international relations theorists, whether they be constructivist, realist, or rationalist.

The fact that all three approaches can simultaneously point, at any point in time, to a significant body of empirical evidence, indicates that such a claim is plausible. However, pushing the argument this far would require a book (perhaps several!), not an article, so this suggestion must remain suggestive. And there is no *a priori theoretical* reason to believe it to be true. Ultimately, the question as to the relative weight of socio-structural and social dispositional factors is empirical. More likely is that each is only capturing a part of the puzzle.

The hurdles for a social-psychological approach to international relations based on types are severe. Questions remain as to (1) whether foreign policy decision-making environment is populated by the same variety and types of trusters found in social psychology (the external validity

<sup>9</sup> This is related, but not equivalent, to claiming that there are different types of states, the familiar distinction being between revisionist and status quo (Morgenthau, 1972; Schweller, 1998). The distinction between different types of states might rest in different levels of trust, but they also might reflect different goals altogether. That is, the differences might be in the utility function, not in trust.

question); (2) whether individual actor's dispositions to trust or fear manifest themselves in the decisions they make or the positions they support on foreign policy questions and in international relations in a way that is not reducible to or trumped by situational imperatives (the level of analysis problem); (3) how we can draw inferences about state behavior and macro-structural features of the international system using an approach that builds on the individual level (the ecological fallacy problem); and (4) whether theories describing the behavior of individuals and small groups are relevant for describing the behavior of states, as the latter are much larger collectivities (the units of analysis problem). The presence of different type of trusters in experimental settings only matters if individuals in the population of interest of those studying international relations demonstrate significant variation in dispositions to trust; those dispositions to trust serve as general belief systems that affect decision-making in foreign policy; those individuals have access to institutions that translate their dispositions into action; those actions are part of creating macro-structural dynamics in international relations that go beyond foreign policy; and the effect of dispositions to trust are not trumped by other structural factors.

Below, I offer a research agenda based on social orientation, arguing how we might demonstrate the presence, explain the origins, and trace the impact of different types of trusters at both the micro and macro level. Along the way, I use findings from already existing studies, some explicitly social psychological and others only implicitly so, to confirm the plausibility of a social-psychological agenda and to rebut some likely criticisms from various other approaches centered on the three obstacles mentioned above. I also draw out some original implications of my own.

### *Monadic effects of social orientation: projecting fear and trust*

There is evidence that dispositions to trust serve as a core feature underlying political ideology and link domestic and foreign policy positions. Jost *et al.* (2003), on the basis of a remarkable review and synthesis of studies on the source of differences between left and right, conclude that a significant share, if not most, of the variation in political ideology is explained by the right's greater level of fear. The left, in contrast, is more trusting. The logical response to a belief that others are untrustworthy is of course to accumulate power to use against those who would do harm (Jost *et al.*, 2007). In domestic affairs, these tendencies exhibit themselves in a tendency for the right to adopt more authoritarian positions on civil liberties. In foreign policy, the right endorses greater military spending and less cooperation in international organizations where others might

take advantage. The left puts greater faith in diplomacy and multilateralism (Hurwitz and Peffley, 1987; Chittick *et al.*, 1995; Rathbun, 2007a). Differences on trust appear to structure ideological cleavages and link domestic and foreign policy beliefs, suggesting that individuals' social orientations serve as core values anchoring their political program in general.

If domestic political ideology bleeds into foreign policy because both are manifestations of social orientation and propensities to trust, social psychology expects and can explain variation and internal contestation over foreign policy based on differences in core beliefs about others. The connection also allows us to observe and measure social orientation at the domestic level independently of its behavioral manifestation at the foreign policy level and avoid tautological reasoning. As certain individuals with particular social orientations or groups of like-minded individuals enter or even dominate a state apparatus, their views are translated into policy.

Even while this is only a monadic effect of social orientation on foreign policy, it still has implications for structural international relations paradigms, such as rationalism and structural realism. Theorists in these traditions might mistake the foreign policy behavior of a particular type of truster as evidence for the nature of the international system, when in fact what they are capturing is dispositions predicated on certain beliefs about that system emerging from a social orientation that resembles the logic of the paradigm in which they work. A competitor might claim the system made him do it, but that does not mean it did. Only careful analysis will tell.

Of course, in these instances, realists are drawing conclusions about the nature of the system as a consequence of the actions of individuals who represent the state. Because individuals do not have tremendous choice in their eyes, they are able to talk meaningfully of the state as something more than the individuals who occupy its key posts. Given its focus on disposition, social psychology is less inclined to think of states in anthropomorphic terms, and more as a shorthand for those directly and heavily involved in the policy-making process. The presence of variation in social orientation among foreign policy makers in the same structural setting would be evidence for the social-psychological view.

Domestic political contestation based on different social orientations also provides leverage to evaluate a constructivist critique of a social-psychological approach, that the phenomenon I am describing is based on culture rather than beliefs. Wendt argues that there are three potential 'cultures of anarchy' marked by dynamics that are roughly similar to the beliefs of the three different social orientations. Wendt distinguishes between 'role-identities' and 'roles', the former being 'subjective self-understandings', the latter being 'objective, collectively constituted positions

that give meanings to those understandings'. Roles are macro-level phenomena that transcend the individual level of analysis and are critical for creating culture. Beliefs 'come and go as individuals take on or discard beliefs', whereas roles 'persist as long as someone fills them' (Wendt, 1999: 259). Roles take on a life of their own, independent of individuals. Intersubjectivity trumps subjectivity. It is this stress on roles that allows some constructivists, like realists, albeit in a different way, to think of states as ontological units in and of themselves. Again, given its stress on variation in type, a social orientation approach expects more contestation than consensus. To the extent that we witness significant domestic contestation over foreign policy, we are more inclined to think in terms of ideology rather than culture.

However, domestic ideological cleavages are not the only place to observe different types of trust and trusters. Particular societies might be composed primarily of a particular type of truster. In that sense it would be possible to say that 'states' have a social orientation. For instance, liberal democratic societies are premised on a more benign view of human nature (Doyle, 1986, 1997; Owen, 1997; Rousseau, 2006). In other words, we might say that mature liberal societies have a predominance of individuals with cooperative social orientations. Uslander (2002: Ch. 8) cites evidence that democratic societies are marked by higher degrees of generalized trust than authoritarian societies.<sup>10</sup> While some argue that distrust drives liberal democracy, as evident in the importance placed on accountability of elected officials and the necessity of a free press (Levi, 1998), recall that generalized trust is the belief that most others can be trusted, but not everyone. Liberalism is distrust of power centralized in the hands of a few individuals, but trust in the broader general public. The king does not get the same benefit of the doubt.

Thinking of liberal democratic states as more cooperative and trusting in social orientation provides a micro-logic to some findings about liberal societies and conflict. Democracies are found to be less inclined to *initiate* armed hostilities, even against non-democracies (Rousseau *et al.*, 1996). In other words, there is a monadic effect of democracy in the instigation of conflict. If democracies are indeed marked by generalized trust, this behavior makes sense, as they are less likely to feel that they will be taken advantage of if they enter into diplomatic negotiations. *Ceteris paribus*,

<sup>10</sup> The institutions, however, are more the effect than the cause of generalized trust. Simply creating new institutions, as is the case in transitioning countries, does not affect overall levels of generalized trust. For this reason there is substantial variance in levels of generalized trust among democracies. More mature democracies have much greater levels of generalized trust. See Uslander (2002: Ch. 8).

democracies are better able and more likely to make the first conciliatory gesture in an effort to elicit trust. However, unlike a constructivist account that might consider this democratic culture, a social-psychological approach is also comfortable explaining the presence of non-liberals in liberal societies, who are always present (Owen, 1997).

A focus on dispositions to trust also leads to a potentially unique social-psychological approach to the study of international cooperation and the design of international organizations.

As reviewed above, rationalists understand trust in situational terms. It has nothing to do with the attributes or the character of the individuals cooperating. Therefore, rationalist work on cooperation has focused almost exclusively on how changing the structure of the situation can create incentives for trustworthy behavior. Institutions and organizations can help.

The rationalist literature on cooperation, because it relies on strategic trust, specifies a particular cause and effect relationship between international organizations and trust. Distrust drives the creation of international organizations, which are the producers of strategic trust. And where distrust is greater, states create stronger institutions. The argument that states pool or delegate sovereignty or make firmer commitments to circumvent defection is a common one in the rationalist literature (Snyder, 1984; Garrett, 1992; Pollack, 1997; Abbott and Snidal, 1998; Moravcsik, 1998, 2000; Lake, 1999; Weber, 2000).

This is a plausible argument, but social psychologists would suggest a different mechanism. Significant international cooperation, in which states delegate or pool sovereignty, involves placing one's fate in the hands of others. Particularly in highly uncertain situations, in which numerous issues are potentially at stake among a large number of partners, it becomes difficult, if not impossible, to predict one's interests and the likely alignments of others. Surrendering some degree of control over policy in these instances would be an indication of generalized trust, not distrust.

Social psychology inverts the relationship seen by rationalism. Trust (better stated, generalized trust) precedes and is necessary for the creation of international organizations. Generalized trust is a 'springboard' in uncertain situations to leap into the 'outside world', 'emancipating' us from the secure confines of stable relationships and allowing us to seek other cooperative partners with whom there might be greater gains (Messick and Kramer, 2001: 108; Yamagishi, 2001: 141). Experimental research shows that in prisoner's dilemma games in which subjects are given the option to enter or leave the game, it is trusters who tend to play and keep playing (Orbell *et al.*, 1984; Orbell and Dawes, 1993). Non-trusters, in contrast, tend to leave games or form committed relationships among specific others with whom they are familiar (Yamagishi *et al.*, 1998).

By this logic, individuals with different social orientations will demonstrate different propensities to build international organizations and have different preferences over institutional design. Cooperators will be less isolationist and more inclined to pool or delegate sovereignty. As compared to cooperators, competitors are likely to insist on a veto in institutions or flexibility clauses that allow members to withdraw or limit the term of an agreement. They will want to restrict membership so as to lower the number of members and allow for easier monitoring and enforcement. These dispositions to trust are best seen by looking for variation amongst the preferences of similarly situated actors, for instance in domestic political settings, something that rational design research cannot explain (Olson, 1971; Snidal, 1985; Kahler, 1992; Abbott and Snidal, 2000; Koremenos *et al.*, 2004: 24). Interestingly, even where social-psychological type affects foreign policy in a monadic way, in terms of approaches to international cooperation, it might produce an outcome – an international organization – that helps govern the international system.

*Dyadic effects of particularized trust and fear: rivalries and the democratic peace*

Trust and fear, more specifically particularized trust and particularized fear, are implicit in accounts of arguably the two most important dyadic phenomena in international relations – rivalries and the democratic peace. The former concerns itself with spirals of fear, the latter with spirals of trust. The logics of both apparently rest on a social-psychological foundation.

The international relations literature on rivalries proceeds from a startling observation – that the majority, perhaps even the predominance of international conflicts, whether they be crises, or wars, take place between a small number of dyads (Goertz and Diehl, 1993; Thompson, 1995). These rivals have relations that at first glance appear to be the type expected by structural realism and security dilemma theorists. Defensive actions on the part of state A are understood as potentially offensive by state B and are met in kind. This causes concern on the part of state B and the process can begin a spiral of escalating hostility (Thompson, 2001: 561–562). However, the central premise of rivalry theory is that conflicts between rivals taken own a ‘life of their own’ (Thompson, 1995: 196) due to the ‘psychological baggage’ (Colaresi and Thompson, 2002b: 1182) that comes with repeated crisis and conflict. Vasquez (1996: 532) defines rivalries as relationships ‘characterized by extreme competition and... psychological hostility...’. Thompson (2001: 561) understands it as ‘very much a social-psychological process’.

Although not always explicit, the psychological baggage that accumulates in rivalries seems to be what I would call particularized fear, an expectation on the part of participants that a specific other is inherently untrustworthy. Colaresi and Thompson (2002a: 271) write, 'In a rivalry context, as opposed to more moderate crisis situations, uncertainty is likely to take on a different meaning... Within rivalry, this may increase the fear of a dangerous environment where escalation is inevitable and thus encourage hyper-vigilance. Thus, under increased insecurity, rivals are likely to prepare for the worst.' This is not strategic distrust, in which states are simply objectively updating their conception of others' interests based on prior experiences. Rivals have categorized others as enemies who are intent on harm. They have made attributions about others character, not their interests. Rivals have 'had time to develop images of their adversaries as threatening opponents with persistent aims to thwart their own objectives' (Colaresi and Thompson, 2002a: 263–264).

The psychological nature of distrust and fear in rivalries is evident in how it affects both rivals' information processing and their conception of their interests. The non-rationalist premise on which rivalry theory is built is evident in the fact that rivalry theorists understand their subjects as prone to cognitive biases based on prior experience, biases that distract them from their objective interests. Colaresi and Thompson (2002a: 269) write that 'a rivalry context with its increased mistrust and expectation of conflict will affect the way objective information is processed, all else being equal... Information passed between rivals must be processed within the conflictual setting of that relationship.' The difficulty of transcending a relationship marked by an untrustworthy image of an adversary has been demonstrated by attribution theorists. Social psychology tells us that cooperative and peaceful demonstrations of intent on the part of an adversary are attributed to situational imperatives, whereas hostile actions are attributed to an untrustworthy disposition (Jervis, 1976; Mercer, 1996). Once in place, these images are difficult to break free from. 'The expectations become more rigid, less sensitive to changes in adversary behavior, and less in need of continued reinforcement', writes Thompson (2001: 561–562).

Rivalry research is also non-rationalist in that the nature of conflict in rivalry is different than in other types of conflictual relationships. Vasquez writes that rivals are 'governed primarily by their attitude toward each other rather than by the stakes at hand'. Vasquez (1996: 532) calls this 'negative affect'. This would include not only fear – the belief that others are untrustworthy and intent on harm. It would generate positive utility for any harm to the other regardless of the costs to the self. We might call this hate, or what social psychologists have labeled as an 'aggressive'

social orientation. Whereas ‘normal conflict’ is marked by cost–benefit analysis, rivalries ‘get out of hand’ and can lead to what would normally be considered self-defeating behavior, according to Vasquez. He writes, ‘The relationship is difficult to change because each side is involved in a vicious circle in which hostility makes actors define issues in ways that are intractable and threatening and actors become hostile, in part, because of the way they have defined the issues that divide them’ (1996: 532–533). In other words, rivals transform the game objectively given by the strategic circumstances.

Rivalries have a beginning and grow. They are not a product of objective, structural features of the strategic setting. Rivalry research is based on the premise that the past crises and wars between a particular set of states have an effect on future ones, independent of the structural features of the interaction such as relative capabilities or the persistence of the issues at stake. Crises and wars between the same pair of states cannot be considered independent of one another. In the language of trust, rivalries are marked by *particularized* fear that arises from the history of conflict. This kind of fear is moralistic in that participants have made a judgment about inherent traits of the other side. However, it is not dispositional in that it is not applied beyond that particular other. It is a quality of the relationship not the objective strategic situation or the actors involved. Continuing conflict reinforces and buttresses fear of the other side.

Given that particularized trust is a quality of a relationship and not reflective of the disposition of those in the relationship, constructivism is of considerable help in explaining the persistence of rivalries once constituted. Rivalries are, in Wendt’s terms, roles that require an ‘other’ in order to be taken up. Once constituted, a rivalry is an identity relationship based on an image of the other that is hard to break regardless of the different preferences of state leaders. Colaresi (2004) finds that efforts to initiate rapprochement in rivalries fail at greater rates than in other protracted conflicts, and that unreciprocated efforts result in higher likelihood of losing office. The same is not true in non-rivalry settings, where there is no effect of failure on leadership tenure. Rivalries create an intersubjective cultural context that is irresponsive to the subjective beliefs of individuals.

However, given that rivalries are created through a process of interaction, the features that sustain rivalries cannot explain their origins or their end. The latter must be exogenous. On this basis, rationalists have concluded that there is no reason to believe that rivalries have psychological dynamics. As rivalries only develop over time, after somewhat persistent conflict, the beginning of conflicts between would-be rivals must have other sources. As claimed by Gartzke and Simon (1999), those



same factors likely explain their persistence. They also argue that given the relatively small number of rivalries, the higher level of conflict among rivals can just as easily be explained stochastically as a statistical aberration.

Rivalry theorists have met this theoretical challenge empirically. Colaresi and Thompson (2002b) show that while crises tend to beget a higher likelihood of future conflicts among the same pair of states, self-identified rivals have a much higher likelihood still of another crisis. Rivalry increases the odds of war by a factor of almost three, even while controlling for other factors. Crises between rivals, as opposed to non-rivals, tend to escalate more quickly, to higher stages in the escalation ladder, and often over non-violent, even symbolic cues. While power disparity between non-rivals decreases the likelihood of conflict, there is no such effect on rivalry dyads. The psychological baggage trumps rational calculation.

Nevertheless, rationalists do raise an important theoretical question, even if they do not have a satisfactory empirical answer. To the extent that rivalries are a function of particularized fear, which is a relational factor that transcends the attributes and dispositions of those involved, something exogenous to the relationship must explain changes in the direction of this process. Rivalries have been found to have an arc (Colaresi, 2005). They escalate and de-escalate. Types and social orientation might play an important role.

Different types are very present in the work of those who have tried to specify a micro-logic to the creation of rivalries, even if they are underconceptualized. Absent any preexisting image of the other as enemy, it is likely that the instigators of rivalries and subsequently their greatest defenders are competitors. Theorists of rivalries generally claim that there is a distribution of hawks and doves (in social orientation parlance, competitors and cooperators). Hawks are more inclined to initiate conflict in general, and the conflicts that their policies produce can produce shifts in the overall level of particularized fear *vis-à-vis* their adversary in their societies, which adds to their stature and puts them in position to sustain and escalate the rivalry (Vasquez, 1993; Senese and Vasquez, 2008). Conversely, de-escalation might begin with gestures of conciliation more likely to come from dispositionally more trusting individuals and groups. Particularized and generalized trusts are likely linked empirically.

Social psychology also helps explain interactive aspects of the democratic peace. The democratic peace is a dyadic phenomenon in which democracies are thought to behave peacefully towards other democracies, but not towards autocratic regimes. The incidence of large-scale war and even crisis between democracies has been found to be exceedingly rare,

whereas democracies are no less likely to become engaged in warfare with autocratic regimes than the latter are amongst one another. The peaceful nature of democratic dyad relations does not require significant elaboration. I discussed the monadic effect of the democratic peace above. Trusting gestures are likely to be matched by regimes marked by a similar cooperative social orientation. Hence, democracies rarely fight wars with one another (Maoz and Russett, 1993).

More puzzling is why democracies, if they in fact embrace non-violent methods of conflict resolution, do not exhibit a generically more peaceful foreign and defense policy as a whole toward all other states. A social-psychological focus on trust helps explain this puzzle. Kuhlman and Wimberley (1976) and Kuhlman and Marshello (1975) find that while cooperators tend to begin with cooperation in social dilemmas, they *do not continue* to cooperate against a 100% competitive strategy, opting to defect against defection (although individualists resort to defection much more quickly). Cooperative social orientations are based on an expectation of reciprocity, not altruism. Generalized trust is a default tendency that adjusts in light of the circumstances. In the face of consistent information revealing a breach of trust, cooperation declines sharply on the part of these generalized trusters.

The combination of 'behavioral assimilation' by cooperators and competitors' consistent resistance to cooperation leads to what Kelley and Stahleski (1970) call the 'triangle effect', in which the behavior of individuals with higher predispositions to cooperate based on a greater disposition to trust eventually converges on non-cooperation over time in the face of defection. It might be that even if liberal democracies are predominantly composed of individuals with a cooperative social orientation, when they find reciprocity unforthcoming, they are forced to behave like competitors. Therefore they are no less likely to fight non-democracies than non-democracies are to fight each other, even if they are less inclined to initiate conflict (see above).

Liberal democracy is based both on shared beliefs and a set of institutions. It rests on both a benign and trusting view of human nature and a set of procedures to select political leaders. The advantage of a social-psychological approach is that it can explain both the short-term variation in foreign policy in democracies that is the natural result of electing different types into positions of leadership, as well as the long-term tendency of democracies to behave differently in foreign affairs as a result of the culture of generalized trust that marks most but not all of its citizens. Owen (1997) notes that even liberal democracies have illiberal leaders from time to time, but that the general inclination of democracies in foreign affairs is nevertheless distinct. The dynamics of illiberal

democracies, those with democratic institutions but without a trusting democratic culture, would be, indeed have been found to be, different (Mansfield and Snyder, 2002).

*Systemic effects of social orientation: selection mechanisms allowing cooperation*

A social-psychological approach to international relations based on social orientation is on toughest terrain at the systemic level, where the levels of analysis, the external validity and the ecological fallacy critiques are most applicable. Structural realists might assert that the nature of international relations, the lack of any institutional enforcement of cooperation, means that the applicability of social-psychological findings to foreign policy and international relations is highly questionable. Structure is much more powerful when the stakes are high and protections are low. Cooperating in an experiment with a couple of dollars is one thing, playing the high stakes game of international politics, where one's very survival is in question, is another. Vulnerability is simply too high to trust and will overwhelm any variation in dispositions to trust. In this view, the material nature of international politics acts as a selection mechanism for only competitive orientations. In this vein, Tang (2008) makes the case for the evolutionary superiority of a fearful human nature. This view is not necessarily antithetical to an emphasis on types. Individual states and statesmen are free to trust. They will simply be weeded out of the system. However, it would lead to doubts that there are significant numbers of individualists, much less cooperators.

Neorealists miss that moralistic trust, or even strategic trust, can yield greater payoffs for cooperators and individualists that makes them stronger and better able to take the risk of trusting again in the future. Obviously the risks of cooperation in international politics are higher relative to other arenas, but so might be the rewards. Moralistic trust, both particularized and generalized, serves as a form of 'social capital', allowing for higher levels of cooperation without the need for costly enforcement and the collection of information (Putnam, 1993; Sztompka, 1999: 62; Tyler, 2001: 285; Cook and Cooper, 2003: 209). It allows states to seek potentially greater gains with unfamiliar cooperative partners. It frees its holders from isolation or the familiar but restrictive confines of those with whom one regularly interacts. As stated above, moralistic trust 'emancipates' individuals (Yamagishi, 2001). It might facilitate cooperation even in the most dangerous, anarchical environments. In prisoner's dilemma experiments in which subjects can enter and exit the game, those willing to trust generally fare better than more fearful

individuals do (Orbell and Dawes, 1993). Cooperators do better than competitors.

The observation that cooperators might do better than cooperators also suggests a mechanism by which the distribution of cooperators and competitors might change over time, changing with it the overall character of the system in which they operate. A social-psychological approach to international relations predicated on the importance of types would explain systemic features of the environment in this way, as a consequence of the aggregation of the traits of the units that make it up. Systemic qualities are 'emergent macro-phenomena', to use Cederman's term, created by a multitude of micro-level interactions. Ahn and Esarey (2008) suggest such a game theoretic model consistent with this notion of what biologists call 'reciprocal altruism' (Trivers, 1971). As mentioned above, individualists or cooperators, if they can successfully identify one another, might fare better based on their cooperation than competitors. If this cooperation yields material benefits, more pro-social individuals (or states) are in a better position to defend themselves. They can carve out a space for themselves, replicate, and thrive. Their relative distribution in the overall population might even grow over time. We can imagine this process in biological terms, if dispositions to trust are passed genetically, or in social terms, as a cultural phenomenon. There is evidence for both (Uslaner, 2002; Oxley *et al.*, 2008). This process of selection could provide the link between the individual level of analysis and systemic features of the environment and helps address and redress the ecological fallacy critique leveled at social psychology.

Lars-Erik Cederman (2001b) provides a very similar, if not identical, selection argument to that of Ahn and Esarey (2008) to explain the democratic peace. Cederman frames his argument similarly, asking how more benign democratic polities manage to attempt cooperation if it exposes them to potential extinction. Implied is the danger of trust. However, for the reasons identified above, non-democratic, competitive orientations are not necessarily dominant. Cooperators can thrive, even amongst competitors, if they are able to identify other cooperators. While Ahn and Esarey write of 'credentialing mechanisms', Cederman calls them 'strategic tags'. The most obvious is democracy itself. Democracy might serve to provide evidence of trustworthiness through a strategic trust mechanism, its very transparency, or through a moralistic trust mechanism, democracies' culture of non-violent conflict resolution (Maoz and Russett, 1993; Owen, 1997; Schultz, 1999; Ikenberry, 2001). As a result, democracies thrive on the basis of mutual exchange and their more peaceful relations. If the benefits gained by cooperation among democracies leads to the proliferation of democracy itself, as Kant predicted it could, the very nature of international politics changes.

Cederman (2001a) offers some empirical evidence that the more trusting democracies, over time, created such a transformation.

This description evokes Wendt's stylized account of the origins of a different culture of anarchy. As more and more members of a collectivity come to believe others are enemies, a tipping point is reached at which these representations take over (1999: 264). An individual level phenomenon becomes a social structural one. This is the movement from agency to structure in 'structuration'.

Wendt's understanding of this process is cultural. But the beliefs of individuals come first and they likely vary. The process is driven at the beginning by psychological processes. Fear comes before identity. The motor is at the individual level. The social-psychological hunch would be that there would always be different types, even if the overall distribution varies over time and place. A social orientation approach to international relations would be skeptical that any particular type predominates so much so that there is a systemic 'culture of anarchy'. Indeed the existence of different types is necessary to begin any change in a culture of anarchy. Anarchy can only be what states make of it if there are some types who think a little differently. The interaction process foreseen by Wendt's constructivism leading to the creation of moralistic trust is most likely confined to particular relationships, and not broad enough to create an overall nature of the system. It is mostly a dyadic approach. For a culture of trust or fear to form, states (or individuals) would have to draw general conclusions from their more local experiences and these are unlikely to be homogenous.

Regardless of which mechanism is correct, demonstrating the systemic manifestations of micro-level processes is extremely difficult, something which might be thought to give a certain edge to structural realism. However, a social-psychological approach based on type alerts us to the possibility that even neorealism's systemic qualifications are dubious. Structural realism and realism in general expect international relations to be largely conflictual and premised on self-help, lacking in trust and cooperation except in situations of mutual threat. Yet rivalry theory notes that conflict in the international system is relatively rare empirically and confined to a small number of dyads. As discussed, hostility amongst them is driven less by their objective structural and material circumstances than the accumulated psychological baggage of fear. Rivalry theory would tell us that realism oversamples on conflict, drawing general conclusions about systemic qualities of the international environment that are in fact psychological and confined in space and time to dyads. Rivals 'create and sustain a *relationship of atypical hostility*', writes Thompson (2001: 558) (emphasis added). In other words, neorealism is reductionist.

### It takes all types (of theorists)

Of course the idea that it is dispositions all the way down is just one possibility. More likely is that social orientation works in tandem and interacts with structural forces to generate both micro- and macro-level outcomes. *Ceteris paribus*, cooperation is less likely in situations where defection is more costly, as neorealists maintain. Rationalism draws attention to the informational environment, which can affect the level of cooperation independent of dispositions to trust. Constructivism provides an explanation to how particular dispositions accumulate and aggregate in societies and how they sustain themselves independent of the individuals who make them up. This would still leave room for social psychology to offer a micro-logic to how those social constructions come to be and how they come to change; how individuals vary in their perception of the relative costs and benefits of cooperation in the same structural circumstances; and how they perceive their informational environment differently. In short, it might well take all types of theory and theorists to build better theories, but social psychology should be in the mix as having something unique to offer.

Structure and disposition are often difficult to disentangle. However, the problem is oftentimes the lack of consideration of a dispositional explanation in a largely macro-oriented field. One sure-fire method of judging the presence of a true structural cause for a phenomenon is to evaluate whether there is significant variation in response to the stimuli on the part of similarly situated actors. For instance, state X might balance power against state Y, from which some analyst might draw broader lessons about the nature of the international system. However, this might simply have been the response of leaders with a competitive social orientation. If we find that policy was criticized by other more cooperatively-oriented political actors in the same arena, the broader conclusions reached about the nature of the international system become more suspect as others in the same situation might have responded differently. Conversely, if there are no significant differences on the part of those in the same structural situation, a structural argument is on safer ground (although the correct explanation might still be cultural).

Ironically, some of the resistance to a focus on types might come from the consumers of social psychology itself. Drawing on social identity theory, Mercer (1995) stresses that the inclination of individuals to join groups in order to bolster self-esteem comes at the price of difficult relations between in-groups and out-groups. He draws the conclusion that the nature of international relations, as it amounts to interactions among groups, is likely to be conflictual, not because of anarchy but

human nature. However, this argument focuses too narrowly on one strand of social-psychological work to the detriment of others. The overwhelming consensus of the literature reviewed above is that the findings of social identity theory, as understood by Mercer, capture just one of the many social orientations observed empirically. In other words, he is overstating the frequency of a competitive social orientation. It will take all types to bring social psychology to the fore in international relations theory.

## Acknowledgements

The author thanks Jacques Hymans, Shiping Tang, Ned Lebow, Jonathan Mercer participants in the Working Paper series in the School of International Relations at USC, students in the Advanced International Relations Theory course of the POIR program, the editors and anonymous reviewers of *International Theory*, and especially Nicholas Onuf for helpful guidance and comments on this article.

## References

- Abbott, K.W. and D. Snidal (1998), 'Why states act through formal international organizations', *The Journal of Conflict Resolution* 42(1): 3–32.
- (2000), 'Hard and soft law in international governance', *International Organization* 54(3): 421–456.
- Adler, E. and M.N. Barnett (eds) (1998), 'A framework for the study of security communities', in *Security Communities*, Cambridge: Cambridge University Press, pp. 29–66.
- Ahn, T.K. and J. Esarey (2008), 'A dynamic model of generalized social trust', *Journal of Theoretical Politics* 20(2): 151–180.
- Alcock, J.E. and D. Mansell (1977), 'Predisposition and behaviour in a collective dilemma', *Journal of Conflict Resolution* 21(3): 443–457.
- Axelrod, R.M. (1984), *The Evolution of Cooperation*, New York: Basic Books.
- Booth, K. and N.J. Wheeler (2008), *The Security Dilemma: Fear, Cooperation and Trust in World Politics*, Basingstoke, England; New York: Palgrave Macmillan.
- Brann, P. and M. Foddy (1987), 'Trust and the consumption of a deteriorating common resource', *The Journal of Conflict Resolution* 31(4): 615–630.
- Brooks, S.G. (1997), 'Dueling realisms', *International Organization* 51(3): 445–477.
- Bueno de Mesquita, B., J.D. Morrow and E.R. Zorick (1997), 'Capabilities, perception, and escalation', *The American Political Science Review* 91(1): 15–27.
- Cederman, L.-E. (2001a), 'Back to Kant: reinterpreting the democratic peace as a macro-historical learning process', *The American Political Science Review* 95(1): 15–31.
- (2001b), 'Modeling the democratic peace as a Kantian selection process', *Journal of Conflict Resolution* 45(4): 470–502.
- Chittick, W.O., K.R. Billingsley and R. Travis (1995), 'A three-dimensional model of American foreign policy beliefs', *International Studies Quarterly* 39(3): 313–331.
- Colaresi, M. (2004), 'When doves cry: international rivalry, unreciprocated cooperation, and leadership turnover', *American Journal of Political Science* 48(3): 555–570.

- Colaresi, M.P. (2005), *Scare Tactics: The Politics of International Rivalry*, 1st edn., Syracuse: Syracuse University Press.
- Colaresi, M. and W.R. Thompson (2002a), 'Strategic rivalries, protracted conflict, and crisis escalation', *Journal of Peace Research* 39(3): 263–287.
- Colaresi, M.P. and W.R. Thompson (2002b), 'Hot spots or hot hands? Serial crisis behavior, escalating risks, and rivalry', *The Journal of Politics* 64(4): 1175–1198.
- Cook, K.S. and R.M. Cooper (2003), 'Experimental studies of cooperation, trust and social exchange', in E. Ostrom and J. Walker (eds), *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research*, Vol. 6. New York: Russell Sage Foundation, pp. 209–244.
- Cook, K.S., R. Hardin and M. Levi (2005), *Cooperation Without Trust?*, New York: Russell Sage Foundation.
- Cronin, B. (1999), *Community Under Anarchy: Transnational Identity and the Evolution of Cooperation*, New York: Columbia University Press.
- Dawes, R.M., J. McTavish and H. Shaklee (1977), 'Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation', *Journal of Personality and Social Psychology* 35(1): 1–11.
- De Cremer, D. and M. Van Vugt (1999), 'Social identification effects in social dilemmas: a transformation of motives', *European Journal of Social Psychology* 29(7): 871–893.
- Deutsch, M. (1960), 'The effect of motivation orientation upon threat and suspicion', *Human Relations* 13: 123–139.
- Doyle, M.W. (1986), 'Liberalism and world politics', *The American Political Science Review* 80(4): 1151–1169.
- (1997), *Ways of War and Peace: Realism, Liberalism, and Socialism*, 1st edn., New York: Norton.
- Garrett, G. (1992), 'International cooperation and institutional choice: the European community's internal market', *International Organization* 46(2): 533–560.
- Gartzke, E. and M.W. Simon (1999), '"Hot hand": a critical analysis of enduring rivalries', *The Journal of Politics* 61(3): 777–798.
- George, A.L. (1969), 'The "operational code": a neglected approach to the study of political leaders and decision-making', *International Studies Quarterly* 13(2): 190–222.
- Glaser, C.L. (1994), 'Realists as optimists: cooperation as self-help', *International Security* 19(3): 50–90.
- Goertz, G. and P.F. Diehl (1993), 'Enduring rivalries: theoretical constructs and empirical patterns', *International Studies Quarterly* 37(2): 147–171.
- Grieco, J.M. (1988), 'Anarchy and the limits of cooperation: a realist critique of the newest liberal institutionalism', *International Organization* 42(3): 485–507.
- Hardin, R. (2001), 'Conceptions and explanations of trust', in K.S. Cook (ed.), *Trust in Society*, Vol. 2. New York: Russell Sage Foundation, pp. 3–39.
- (2006), *Trust*, Cambridge, UK: Polity.
- Hayashi, N., E. Ostrom, J. Walker and T. Yamagishi (1999), 'Reciprocity, trust and the sense of control: a cross-societal study', *Rationality and Society* 11(1): 27–46.
- Hoffman, A.M. (2002), 'A conceptualization of trust in international relations', *European Journal of International Relations* 8(3): 375–401.
- (2006), *Building Trust: Overcoming Suspicion in International Conflict*, Albany: State University of New York Press.
- Hurwitz, J. and M. Peffley (1987), 'How are foreign policy attitudes structured? A hierarchical model', *The American Political Science Review* 81(4): 1099–1120.
- Ikenberry, G.J. (2001), *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars*, Princeton, NJ: Princeton University Press.



- Jervis, R. and Harvard University, Center for International Affairs (1970), *The Logic of Images in International Relations*, Princeton, NJ: Princeton University Press.
- (1976), *Perception and Misperception in International Politics*, Princeton, NJ: Princeton University Press.
- (1978), 'Cooperation under the security dilemma', *World Politics* 30(2): 167–214.
- Jost, J.T., J. Glaser, A.W. Kruglanski and F.J. Sulloway (2003), 'Political conservatism as motivated social cognition', *Psychological Bulletin* 129(3): 339–375.
- Jost, J.T., J.L. Napier, H. Thorisdottir, S.D. Gosling, T.P. Palfai and B. Ostafin (2007), 'Are needs to manage uncertainty and threat associated with political conservatism or ideological extremity?', *Personality and Social Psychology Bulletin* 33(7): 989–1007.
- Kahler, M. (1992), 'Multilateralism with small and large numbers', *International Organization* 46(3): 681–708.
- Kelley, H.H. and A.J. Stahleski (1970), 'Social interaction basis of cooperators' and competitors' beliefs about others', *Journal of Personality and Social Psychology* 16(1): 66–91.
- Kelley, H.H. and J.W. Thibaut (1978), *Interpersonal Relations: A Theory of Interdependence*, New York: Wiley.
- Keohane, R.O. (1984), *After Hegemony: Cooperation and Discord in the World Political Economy*, Princeton, NJ: Princeton University Press.
- (1989), 'Neoliberal institutionalism: a perspective on world politics', in *International Institutions and State Power*, Boulder: Westview, pp. 1–20.
- Kim, W. and B. Bueno de Mesquita (1995), 'How perceptions influence the risk of war', *International Studies Quarterly* 39(1): 51–65.
- Koremenos, B. (2001), 'Loosening the ties that bind: a learning model of agreement flexibility', *International Organization* 55(2): 289–325.
- Koremenos, B., C. Lipson and D. Snidal (2004), *The Rational Design of International Institutions*, Cambridge, UK; New York: Cambridge University Press.
- Kramer, R.M., M.B. Brewer and B.A. Hanna (2004), 'Collective trust and collective action: the decision to trust as a social decision', in R.M. Kramer and K.S. Cook (eds), *Trust and Distrust in Organizations: Dilemmas and Approaches*, Vol. 7. New York: Russell Sage Foundation, pp. 357–389.
- Kuhlman, D.M., C.R. Camac and D.A. Cunha (1986), 'Individual differences in social orientation', in H.A.M. Wilke, D.M. Messick and C.G. Rutte (eds), *Experimental Social Dilemmas*, Frankfurt am Main: Verlag Peter Lang, pp. 151–176.
- Kuhlman, D.M. and A.F. Marshello (1975), 'Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma', *Journal of Personality and Social Psychology* 32(5): 922–931.
- Kuhlman, D.M. and D.L. Wimberly (1976), 'Expectations of choice behavior held by cooperators, competitors, and individualists across four classes of experimental games', *Journal of Personality and Social Psychology* 34(1): 69–81.
- Kydd, A.H. (1997), 'Game theory and the spiral model', *World Politics* 49(3): 325–357.
- (2005), *Trust and Mistrust in International Relations*, Princeton, NJ: Princeton University Press.
- Lake, D.A. (1999), *Entangling Relations: American Foreign Policy in its Century*, Princeton, NJ: Princeton University Press.
- Lake, D.A. and R. Powell (1999), *Strategic Choice and International Relations*, Princeton, NJ: Princeton University Press.
- Larson, D.W. (1994), 'The role of belief systems and schemas in foreign policy decision-making', *Political Psychology* 15(1): 17–33.
- (1997), *Anatomy of Mistrust: US–Soviet Relations During the Cold War*, Ithaca, NY: Cornell University Press.

- Levi, M. (1998), 'A state of trust', in M. Levi and V. Braithwaite (eds), *Trust and Governance*, New York: Russell Sage Foundation.
- Liebrand, W.B., R.W.T.L. Jansen, V.M. Rijken and C.J.M. Suhre (1986), 'Might over morality: social values and the perception of other players in experimental games', *Journal of Experimental Social Psychology* 22: 203–215.
- Luhmann, N. (1979), *Trust; and, Power: Two Works*, Chichester; New York: J. Wiley.
- Maki, J.E. and C.G. McClintock (1983), 'The accuracy of social value prediction: actor and observer influences', *Journal of Personality and Social Psychology* 45(4): 829–838.
- Mansfield, E.D. and J. Snyder (2002), 'Incomplete democratization and the outbreak of military disputes', *International Studies Quarterly* 46(4): 529–549.
- Maoz, Z. and B. Russett (1993), 'Normative and structural causes of democratic peace, 1946–1986', *The American Political Science Review* 87(3): 624–638.
- Marwell, G. and R.E. Ames (1979), 'Experiments on the provision of public goods. I. resources, interest, group size, and the free-rider problem', *The American Journal of Sociology* 84(6): 1335–1360.
- McClintock, C.G. (1972), 'Social motivation: a set of hypotheses', *Behavioral Science* 17(5): 438–454.
- McClintock, C.G. and W.B. Liebrand (1988), 'Role of interdependence structure, individual value orientation, and another's strategy in social decision making: a transformational analysis', *Journal of Personality and Social Psychology* 55(3): 396–409.
- Mearsheimer, J.J. (1994), 'The false promise of international institutions', *International Security* 19(3): 5–49.
- Mercer, J. (1995), 'Anarchy and identity', *International Organization* 49(2): 229–252.
- (1996), *Reputation and International Politics*, Ithaca, NY: Cornell University Press.
- (2005), 'Rationality and psychology in international politics', *International Organization* 59(1): 77–106.
- Messick, D.M. and M.B. Brewer (1983), 'Solving social dilemmas: a review', *Review of Personality and Social Psychology* 4: 11–44.
- Messick, D.M. and R.M. Kramer (2001), 'Trust as a form of shallow morality', in K.S. Cook (ed.), *Trust in Society*, Vol. 2. New York: Russell Sage Foundation, pp. 89–118.
- Messick, D.M. and C.G. McClintock (1968), 'Motivational bases of choice in experimental games', *Journal of Experimental Social Psychology* 4(1): 1–25.
- Messick, D.M., H. Wilke, M.B. Brewer, R.M. Kramer, P.E. Zemke and L. Lui (1983), 'Individual adaptations and structural change as solutions to social dilemmas', *Journal of Personality and Social Psychology* 44(2): 294–309.
- Morgenthau, H.J. (1972), *Politics Among Nations: The Struggle for Power and Peace*, 5th edn., New York: Knopf; distributed by Random House.
- Moravcsik, A. (1998), *The Choice for Europe: Social Purpose and State Power from Messina to Maastricht*, Ithaca, NY: Cornell University Press.
- (2000), 'The origins of human rights regimes: democratic delegation in postwar Europe', *International Organization* 54(2): 217–252.
- Olson, M. (1971), *The Logic of Collective Action; Public Goods and the Theory of Groups* (Revised edition), New York: Schocken Books.
- Orbell, J.M. and R.M. Dawes (1993), 'Social welfare, cooperators' advantage, and the option of not playing the game', *American Sociological Review* 58(6): 787–800.
- Orbell, J.M., P. Schwartz-Shea and R.T. Simmons (1984), 'Do cooperators exit more readily than defectors?', *The American Political Science Review* 78(1): 147–162.
- Ostrom, E. (2003), 'Toward a behavioral theory linking trust, reciprocity and reputation', in E. Ostrom and J. Walker (eds), *Trust and Reciprocity: Interdisciplinary*

- Lessons from Experimental Research*, Vol. 6. New York: Russell Sage Foundation, pp. 19–79.
- Owen, J.M. (1997), *Liberal Peace, Liberal War: American Politics and International Security*, Ithaca, NY: Cornell University Press.
- Oxley, D.R., K.B. Smith, J.R. Alford, M.V. Hibbing, J.L. Miller and M. Scalora, *et al.* (2008), 'Political attitudes vary with physiological traits', *Science* 321(5896): 1667–1670.
- Oye, K.A. (1985), 'Explaining cooperation under anarchy: hypotheses and strategies', *World Politics* 38(1): 1–24.
- Parks, C.D. (1994), 'The predictive ability of social values in resource dilemmas and public goods games', *Personality and Social Psychology Bulletin* 20(4): 431–438.
- Pollack, M.A. (1997), 'Delegation, agency, and agenda setting in the European community', *International Organization* 51(1): 99–134.
- Press, D.G. (2005), *Calculating Credibility: How Leaders Assess Military Threats*, Ithaca, NY: Cornell University Press.
- Putnam, R.D. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, NJ: Princeton University Press.
- Rathbun, B. (2007a), 'Hierarchy and community at home and abroad: evidence of a common structure of domestic and foreign policy beliefs in American elites', *Journal of Conflict Resolution* 51(3): 379–407.
- (2007b), 'Uncertain about uncertainty: understanding the multiple meanings of a crucial concept in international relations theory', *International Studies Quarterly* 51(3): 533–557.
- Rotter, J.B. (1980), 'Interpersonal trust, trustworthiness and gullibility', *American Psychologist* 35(1): 1–7.
- Rousseau, D.L. (2006), *Identifying Threats and Threatening Identities: The Social Construction of Realism and Liberalism*, Stanford, CA: Stanford University Press.
- Rousseau, D.L., C. Gelpi, D. Reiter and P.K. Huth (1996), 'Assessing the dyadic nature of the democratic peace', *The American Political Science Review* 90(3): 512–533.
- Schelling, T.C. (1966), *Arms and Influence*, New Haven: Yale University Press.
- Schultz, K.A. (1999), 'Do democratic institutions constrain or inform? contrasting two institutional perspectives on democracy and war', *International Organization* 53(2): 233–266.
- Schweller, R.L. (1998), *Deadly Imbalances: Tripolarity and Hitler's Strategy of World Conquest*, New York: Columbia University Press.
- Senese, P.D. and J.A. Vasquez (2008), *The Steps to War: An Empirical Study*, Princeton: Princeton University Press.
- Snidal, D. (1985), 'The limits of hegemonic stability theory', *International Organization* 39(4): 579–614.
- Snyder, G.H. (1984), 'The security dilemma in alliance politics', *World Politics* 36(4): 461–495.
- Stein, A.A. (1982), 'Coordination and collaboration: regimes in an anarchic world', *International Organization* 36(2): 299–324.
- Sztompka, P. (1999), *Trust: A Sociological Theory*, Cambridge, UK; New York, NY: Cambridge University Press.
- Taliaferro, J.W. (2000), 'Security seeking under anarchy: defensive realism revisited', *International Security* 25(3): 128–161.
- Tang, S. (2008), 'Fear in international politics: two positions', *The International Studies Review* 10: 451–471.
- Tetlock, P.E. (1998), 'Social psychology and world politics', in D.T. Gilbert, S.T. Fiske and G. Lindzey (eds), *Handbook of Social Psychology*, New York: McGraw Hill, pp. 868–914.
- Thompson, W.R. (1995), 'Principal rivalries', *Journal of Conflict Resolution* 39(2): 195–223.

- (2001), 'Identifying rivals and rivalries in world politics', *International Studies Quarterly* 45(4): 557–586.
- Trivers, R.L. (1971), 'The evolution of reciprocal altruism', *The Quarterly Review of Biology* 46(1): 35–57.
- Tyler, T.R. (2001), 'Why do people rely on others? Social identity and social aspects of trust', in K.S. Cook (ed.), *Trust in society*, Vol. 2. New York: Russell Sage Foundation, pp. 285–306.
- Tyler, T.R. and P. DeGoey (2004), 'Trust in organizational authorities: the influence of motive attributions on willingness to accept decisions', in R.M. Kramer and K.S. Cook (eds), *Trust and Distrust in Organizations: Dilemmas and Approaches*, Vol. 7. New York: Russell Sage Foundation, pp. 331–356.
- Tyszka, T. and J.L. Grzelak (1976), 'Criteria of choice in non-constant-sum games', *The Journal of Conflict Resolution* 20(2): 357–376.
- Uslaner, E.M. (2002), *The Moral Foundations of Trust*, Cambridge; New York: Cambridge University Press.
- Van Evera, S. (1999), *Causes of War: Power and the Roots of Conflict*, Ithaca: Cornell University Press.
- Van Lange, P.A.M. and D.M. Kuhlman (1994), 'Social value orientations and impressions of partner's honesty and intelligence: a test of the might versus morality effect', *Journal of Personality and Social Psychology* 67(1): 126–141.
- Vasquez, J.A. (1993), *The War Puzzle*, Cambridge; New York: Cambridge University Press.
- (1996), 'Distinguishing rivals that go to war from those that do not: a quantitative comparative case study of the two paths to war', *International Studies Quarterly* 40(4): 531–558.
- Weber, K. (2000), *Hierarchy Amidst Anarchy: Transaction Costs and Institutional Choice*, Albany, NY: State University of New York Press.
- Walker, S.G. (1983), 'The motivational foundations of political belief systems: a re-analysis of the operational code construct', *International Studies Quarterly* 27(2): 179–202.
- Walt, S.M. (1987), *The Origins of Alliances*, Ithaca: Cornell University Press.
- Waltz, K.N. (1979), *Theory of International Politics*, 1st edn., Boston, MA: McGraw-Hill.
- Wendt, A. (1992), 'Anarchy is what states make of it: the social construction of power politics', *International Organization* 46(2): 391–425.
- (1999), *Social Theory of International Politics*, Cambridge, UK; New York: Cambridge University Press.
- Yamagishi, T. (1988), 'The provision of a sanctioning system in the United States and Japan', *Social Psychology Quarterly* 51(3): 265–271.
- (2001), 'Trust as a form of social intelligence', in K.S. Cook (ed.), *Trust in Society*, Vol. 2. New York: Russell Sage Foundation, pp. 121–147.
- Yamagishi, T. and K.S. Cook (1993), 'Generalized exchange and social dilemmas', *Social Psychology Quarterly* 56(4): 235–248.
- Yamagishi, T. and K. Sato (1986), 'Motivational bases of the public goods problem', *Journal of Personality and Social Psychology* 50(1): 67–73.
- Yamagishi, T. and M. Yamagishi (1994), 'Trust and commitment in the United States and Japan', *Motivation and Emotion* 18(2): 129–166.
- Yamagishi, T., K.S. Cook and M. Watabe (1998), 'Uncertainty, trust, and commitment formation in the United States and Japan', *American Journal of Sociology* 104(1): 165–194.