

*“In what ways is the new focus on leaders an improvement on the leader-focused approaches taken during the 1960s and 1970s? In what ways might the older literature be relevant that the new approaches are missing?”*

Do leaders' preferences affect international outcomes? We can immediately think of two reasons why this might not be so – leaders might be institutionally constrained to behave in particular ways in spite of their preferences, or the process of leadership selection might be such that only an individual with particular preferences can rise to a leadership role in the first place. Waltz (1979:86-88) credits both of these explanations in his discussion of the type of politician able to become prime minister of Great Britain (apparently only the cautious, cabinet government type). However, Waltz is insufficiently alive to the possibilities both of preference falsification and structural change over time within the institution itself. Both mechanisms could allow individuals with preferences at variance with the institutions that selected them to rise to power. It seems that we have a good case for investigating the circumstances under which leaders' preferences matter, but we should keep Waltz's conjecture in mind – the preferences and idiosyncrasies of leaders may actually be the result of a selection process, and despite their apparent variation, actions taken by these leaders may converge on outcomes favorable to the state.

Fearon (1995) pointed out the ex-post inefficiency of war and the seeming paradox that states nevertheless engage in it. However, if we consider that a leader's incentives may come apart from those of the state as a whole, we can begin to explain a whole class of international behavior as serving the leader's preferences at the expense of the state. Downs and Rocke (1995) argue that imperfect information leads electorates to adopt suboptimal removal mechanisms. Because uncertainty requires electorates to evaluate conflict on its outcome rather than its ex-ante advisability, competent leaders who enter an (ex-ante advisable) conflict but nevertheless lose will be removed from office, at some cost to the public interest. More troublingly, leaders will be inspired to “gamble for resurrection” by perpetuating or escalating a conflict beyond the level at which a well-informed public would balk. Because they have nothing to lose by entering an (ex-ante inadvisable) conflict, the phenomenon of “gambling for resurrection” seems to be a direct ramification of the imperfect information environment. As for the identity of these gambling leaders, Horowitz and Stam (2014) find that the leaders most likely to initiate militarized disputes and wars are 1) former rebels and 2) those with prior military service but no combat experience (this effect was strongest in nondemocratic regimes).

Chiozza and Goemans (2011) find that initiating international conflicts lowers the risk of regular and forcible removals of leaders from office. They further find that victory in crises short of war lowers the risk of forcible removal from office, and defeats in crises and wars makes *forcible* removal more likely, but has no significant effect on regular removal. For the country as a whole, war is always ex-post

inefficient, but maybe not for the leaders as individuals because they can perpetuate themselves in office by continuing the conflict. This seems to be related to gambling for resurrection (or as they put it, “gambling for survival,” in the sense that it describes a case where the incentives of the individual leaders come apart from the incentives of the states they rule. Leaders at risk of forcible removal will be more likely to initiate a conflict when the prospects of removal with punishment are high, and if it is possible to send one’s political opponents to fight a war (“fighting for survival”), leaders will be more likely to initiate conflict. One missing piece of the explanation seems to be any informal constraint mechanisms by which society constrains the leader.

States might be more likely to achieve fully rational outcomes than individuals. Kahler (1998) introduces prospect theory to explain individual-level deviations from expected-utility theory. He predicts that the field will continue to relax rationalist assumptions to accurately model bounded rationality, deconstruct the unitary state, and account for cultural beliefs. Assumptions of rationality should not be held dogmatically, and institutional design must account for people as they are. Kahler is skeptical of attempts to aggregate upwards that do not make their assumptions explicit, and he argues that individual-level behavior may be swamped by group-level phenomena like culture. He suggests that we pay more attention to culture, norms and identity. McDermott (2004) argues that prospect theory usefully incorporates features of human cognition (such as loss-aversion and framing effects) that are typically absent in rational-choice models. While prospect theory provides a micro-foundational basis on which to understand decision-making, it has limited applicability to group behavior and does not account for emotion, rendering it of uncertain utility in political science.

Saunders (2009) argues that leaders can be usefully characterized on the basis of their causal beliefs about the origin of threats, and she further argues that this variation in the attributes of individual leaders must play a vital role in any explanations for how states use military force. Saunders identifies “internally focused” leaders who see a causal connection between threatening foreign and security policies and the internal organization of other states, and are thus more willing to undertake “transformative” interventions, whereas “externally focused” leaders diagnose threats directly from the foreign and security policies of other states, prompting them to pursue “nontransformative” strategies. She asserts that beliefs regarding internal vs. external focus are typically set in late adolescence and early adulthood. In a similar vein, Gross Stein (2017) argues that rational choice is a normative baseline for group decision-making, and that our goal should be to establish thresholds for when psychology affects rational outcomes.

Yarhi-Milo (2013) proposes that individual perceptual biases interact with organizational interests and practices to influence precisely which types of indications of foreign behavior are deemed credible.

As a result, the theory predicts disagreement between a state's political leaders and its intelligence community, with decisionmakers deemphasizing costly signals in favor of vivid events and personal connections with foreign leaders. Calling this the selective attention thesis ("individual biases influence credibility"), she distinguishes it from models of behavior based on capabilities and costly signals. Her fundamental argument is that leaders update disproportionately based on interactions with foreign leaders and significant events. It is unclear, however, whether we should view this as a fundamentally different process from mere Bayesian updating.

The background of leaders seems to matter. However, given the impossibility of observing a counterfactual it seems very difficult to demonstrate that any particular international outcome was a function either purely or primarily of leadership idiosyncrasy. On the long view, such perturbations may be mere noise, obscuring the elegant dance of national interests that Waltz described. To the extent that leaders deviate from expected utility theory because they are human, we should expect similar deviations to occur in the aggregate. Work on the formative effects of individual background seems to me to be incapable of overcoming Waltz's conjecture that the people who rise to power do so because their preferences are appealing to the state. Useful work in this area might examine historical examples of leaders maintaining preferences at variance with those of their selectorate over the long term. But we might expect preferences to change in the long term, particularly if a skillful leader were able to motivate a reinterpretation of the national interest. The concept of a preference seems unhelpful in answering this question. On balance, prospect theory seems less useful than a focus on the background and characteristics of particular leaders, but both approaches dramatically enrich the explanatory possibilities open to us. However, before introducing an explanation relying on the psychology of leaders, we should shoulder the burden of demonstrating that Waltz's conjecture does not apply in the case under consideration.

## References

- Chiozza, Giacomo and H.E. Goemans. 2011. *Leaders and International Conflict*. New York: Cambridge University Press, Chapters 1-4, pp.11-116.
- Downs, George W., and David M. Roake. 1995. *Optimal Imperfection? Domestic Uncertainty and Institutions in International Relations*. Princeton: Princeton University Press.
- Fearon, James. 1995. "Rationalist Explanations for War." *International Organization* 49 (3), pp. 379-414.
- Gross Stein, Janice. 2017. "The Micro-Foundation of International Relations Theory: Psychology and Behavioral Economics," *International Organization* , 71(S1): S249-63.
- Horowitz, Michael C., and Allan C. Stam. 2014 "How Prior Military Experience Influences the Future Militarized Behavior of Leaders," *International Organization*, 68(3): 527-59.
- Kahler, Miles. 1998. "Rationality in International Relations," *International Organization*, 52(4) 919-41.
- McDermott, Rose. 2004. "Prospect Theory in Political Science: Gains and Loses From the First Decade," *Political Psychology*, 25(2): 289-312.
- Saunders, Elizabeth N. 2009. "Transformative Choices: Leaders and the Origins of Intervention Strategy," *International Security*, 34(2): 119-61.
- Waltz, Kenneth. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley Publishing.
- Yarhi-Milo, Keren. 2013. "In the Eye of the Beholder: How Leaders and Intelligence Communities Assess the Intentions of Adversaries," *International Security*, 38(1): 7-51.