Review of International Studies (2025), page 1 of 19
doi:10.1017/S0260210525000026

RIS

RESEARCH ARTICLE

The Rousseauian roots of neorealism

Joseph Clarkson

Department of Political Science, University of Notre Dame, Notre Dame, IN, USA
Email: jclarks2@nd.edu

(Received 11 June 2024; revised 11 December 2024; accepted 13 December 2024)

Abstract

Does neorealism contain a conception of human nature? Although neorealists usually claim to sidestep the question of human nature altogether, scholars frequently trace the theory back to the work of Thomas Hobbes, a philosopher who in fact defended a robust account of human motivations. As a result, some scholars have concluded that neorealism contains a Hobbesian view of human nature. Against the conventional wisdom, this article argues that neorealism contains a Rousseauian philosophical anthropology. Whereas Hobbes provides a deeply pessimistic account of human motivations, Rousseau combines optimism about human motivations with pessimism about social structures. Rousseau's critique of Hobbes profoundly influenced neorealism's founding father, Kenneth Waltz, a political theorist who later gravitated towards International Relations. After exploring Waltz's reading of Rousseau and demonstrating Rousseau's influence on Waltz's theory, the article investigates what is gained by reading neorealism as Rousseauian. Returning to the Rousseauian roots of neorealism reveals the true character of the tragic heuristic employed in neorealist theory, sheds new light on the role of pity in neorealist foreign policy, and clarifies the logic of the theory itself.

Keywords: classical realism; human nature; International Relations theory; neorealism; structural realism; tragedy

Does neorealism contain a conception of human nature?[1] If it does, what is its conception? The conventional wisdom holds that neorealism does not have a conception of human nature,[2] though a few scholars have proposed arguments to the contrary.[3] More often than not, neorealists describe themselves as sidestepping the question of human nature altogether. For instance, John Mearsheimer – probably the most well known living neorealist – contrasts neorealism with the 'human nature realism' of classical realists such as Hans Morgenthau.[4] In so doing, he follows the lead of neorealism's founding father, Kenneth Waltz, who rejected Morgenthau's theory on similar grounds.[5]

[1]Neorealism is a school of International Relations theories that privileges structural factors. Joseph Parent and Sebastian Rosato, 'Balancing in neorealism', *International Security*, 40:2 (2015), pp. 51–86 (p. 51).

[2]Chris Brown, 'Structural realism, classical realism and human nature', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 143–157 (p. 150).

[3]Neta Crawford, 'Human nature and world politics: Rethinking "man"', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 158–176 (p. 158); Jean Bethke Elshtain, 'Woman, the state, and war', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 177–192 (p. 185); Annette Freyberg-Inan, *What Moves Man: The Realist Theory of International Relations and Its Judgment of Human Nature* (New York: SUNY Press, 2004), p. 73.

[4]John J. Mearsheimer, *The Tragedy of Great Power Politics* (New York: W.W. Norton, 2014), pp. 19–22.

[5]Cf. Kenneth N. Waltz, *Theory of International Politics* (Reading, MA: Addison-Wesley, 1979), pp. 60–7. On Waltz as the father of neorealism, see Daniel Bessner and Nicolas Guilhot, 'How realism Waltzed off: Liberalism and decisionmaking in Kenneth Waltz's neorealism', *International Security*, 40:2 (2015), pp. 87–118 (p. 87); Freyberg-Inan, *What Moves Man*, p. 73.

Nevertheless, when neorealists recount their theory's intellectual genealogy, they frequently trace it back to thinkers such as Thomas Hobbes who, in fact, articulated a robust conception of human nature. Mearsheimer himself claims that his 'theory can also be used to explain the behavior of individuals' and that the 'most important work in this regard is Thomas Hobbes'.[6] Others looking for neorealism's roots also frequently trace it back to Hobbes or characterise the theory as Hobbesian.[7]

Against the conventional wisdom, this article argues that neorealism implies a Rousseauian conception of human nature. In other words, neorealism's core causal logic (the security dilemma), when applied to individuals dispersed in a pre-political state, predicts behaviour that more closely resembles Rousseau's state of nature rather than Hobbes's. That neorealism's implied view of human nature is Rousseauian should come as little surprise because neorealism's founding father, Kenneth Waltz, a political theorist who later gravitated towards International Relations (IR), was deeply influenced by Rousseau's critique of Hobbes. Rousseau had rejected Hobbes's anthropological pessimism on the grounds that people behave aggressively in some contexts but not others.[8] Like Rousseau, Waltz rejected Hobbes's philosophical anthropology because it assumes that human beings have motives which would make them warlike even in the absence of situational constraints.[9] If this argument is sound, scholars should describe neorealism's philosophical anthropology as Rousseauian rather than Hobbesian.[10]

Although Robert Gilpin developed an alternative neorealism simultaneously, Ken Booth points out that 'intellectually speaking, we are all Waltz's subjects', because 'the discipline defines itself in relation to the authority of his work'. Cf. Ken Booth, 'Introduction to the Special Issue "The King of Thought: Theory, the Subject, and Waltz"', *International Relations*, 23:3 (2009), pp. 179–181 (p. 179). Indeed, Waltz's rivals took him as their point of departure, e.g., Robert Keohane, *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton, NJ: Princeton University Press, 1984); Alexander Wendt, *Social Theory of International Politics* (Cambridge: Cambridge University Press, 1999); cf. Ken Booth, 'Realism redux: Contexts, concepts, contests', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 1–14 (p. 5).

[6]Mearsheimer, *The Tragedy of Great Power Politics*, p. 424.

[7]Jack Donnelly, 'The discourse of anarchy in IR', *International Theory*, 7:3 (2015), pp. 393-425 (p. 417); Freyberg-Inan, *What Moves Man*, p. 61; David Polansky, 'Drawing out the Leviathan: Kenneth Waltz, Hobbes, and the neorealist theory of the state', *International Studies Review*, 18:2 (2016), pp. 268–289.

[8]Jean-Jacques Rousseau, 'Principles of the right of war', in Christopher Bertram (ed.), *Of the Social Contract and Other Political Writings*, trans. Quentin Hoare (New York: Penguin, 2012), pp. 149–168 (p. 157).

[9]Kenneth N. Waltz, 'Realist thought and neorealist theory', *Journal of International Affairs*, 44:1 (1990), pp. 21–37 (p. 35).

[10]Although Keith Shimko and Deborah Boucoyannis have argued that Waltz belongs to the classical liberal tradition, and Daniel Deudney places him in the republican security tradition, this article recovers the specifically Rousseauian roots of Waltz's theory. Keith Shimko, 'Realism, neorealism, and American liberalism', *The Review of Politics*, 54:2 (1992), pp. 281–301; Deborah Boucoyannis, 'The international wanderings of a liberal idea, or why liberals can learn to stop worrying and love the balance of power', *Perspectives on Politics*, 5:4 (2007) pp. 703–727; Daniel H. Deudney, *Bounding Power: Republican Security Theory from the Polis to the Global Village* (Princeton, NJ: Princeton University Press, 2008), pp. 76–9, 85–8. In recovering Waltz's debt to Rousseau, this article makes a significant contribution to the growing literature on Waltz: Bessner and Guilhot, 'How realism Waltzed off'; Jack Donnelly, 'Systems, levels, and structural theory: Waltz's theory is not a systemic theory (and why that matters for International Relations today)', *European Journal of International Relations*, 25:3 (2019), pp. 904–930; Stacie Goddard and Daniel H. Nexon, 'Paradigm lost? Reassessing theory of international politics', *European Journal of International Relations*, 11:1 (2005), pp. 9–61; David LaRoche and Simon Pratt, 'Kenneth Waltz is not a neorealist (and why that matters)', *European Journal of International Relations*, 24:1 (2018), pp. 9–61; Joseph MacKay, 'Kenneth Waltz's approach to reading classic political theory and why it matters', *International Theory*, 14:2 (2022), pp. 338–357; Joseph MacKay, '"Making democracy safe for the world": Kenneth Waltz on realism, democracy, and war', *International Studies Quarterly*, 68:3 (2024), pp. 1–12; Alex Prichard, 'Kenneth Waltz's Kantian moral philosophy: "The virtues of anarchy" reconsidered', *International Theory* 16:3 (2024), pp. 410–437; Paul R. Viotti, *Kenneth Waltz: An Intellectual Biography* (New York: Columbia University Press, 2024). It is also worth mentioning here that Waltz's work has sparked enduring interest in Rousseau 's views on IR: Blaise Bachofen, 'The paradox of "just war" in Rousseau's theory of interstate relations', *American Political Science Review*, 109:2 (2015), pp. 314–325; Ian Clark, *The Hierarchy of States: Reform and Resistance in the International Order* (Cambridge: Cambridge University Press, 1989), pp. 73–8; Stanley Hoffmann, 'Rousseau on war and peace', *American Political Science Review*, 57:2 (1963), pp. 317–333; Stanley Hoffmann and David P. Fidler (eds), *Rousseau on International Relations* (Oxford: Oxford University Press, 1991); Torbjørn L. Knutsen, 'Re-reading Rousseau in the post–Cold War world', *Journal of Peace Research*, 31:3 (1994), pp. 247–262; Shmuel Nili, 'Democratic disengagement: Toward Rousseauian global reform', *International Theory*, 3:3 (2011), pp. 355–389; Grace G. Roosevelt, *Reading Rousseau in the Nuclear Age* (Philadelphia: Temple University

IR scholars should take this argument seriously. After all, any IR theory requires assumptions about what motivates people.[11] There are, however, other benefits to reading neorealist anthropology as Rousseauian. Doing so brings clarity and coherence to three separate strands of neorealist theorising.

First, as we will see, Rousseau differed from Hobbes in viewing conflict's origins as tragic, with the resultant suffering contradicting the benign motives of the agents. Like Rousseau, neorealists such as John Mearsheimer tend to combine optimism about agent's motives for action with pessimism about the context or structure within which they act.[12] Put simply, Rousseau and neorealists are united in seeing spirals of security competition – sometimes leading to war – as tragedies of innocent suffering.

Second, returning to Rousseau sheds new light on the role of pity in neorealism. Although pity is one of the two basic human motivations described by Rousseau and by John Herz, the theorist who coined the 'security dilemma', this emotion or faculty has played a less significant role in subsequent realist thought. Only recently have IR theorists begun to reevaluate its place, whether in the form of a 'security dilemma sensibility' or of 'strategic empathy'.[13] Returning to Rousseau indicates not only why pity was destined to make a comeback in neorealism but also why its role has so often remained obscure.

Finally, a Rousseauian interpretation of neorealist anthropology clarifies the logic of the security dilemma itself. Rousseau's account of conflict's origins shows that anarchy is not sufficient to cause competition for power among units who seek survival and who can harm one another. For the anxiety about others' future intentions that leads agents in anarchy to compete for security to occur, those agents need first to have become conscious of time, or of the self as extending through time. Whereas Hobbes assumes that language and, with it, self-consciousness are natural to human beings, Rousseau recognises that language cannot arise among solitary individuals. Rather than being innate to us, both time-consciousness and the anxiety about others' future intentions that it enables belong to the social state, not to human beings considered individually. Returning to Rousseau, then, shifts the emphasis away from a monocausal logic of anarchy as a driver of security competition and towards a multicausal logic driven both by anarchy and what neorealists call 'the problem of the future'.[14]

This article explores neorealism's Rousseauian anthropology in five sections. The first section defines the terms and scope of the paper. The second section defends Waltz's reading of Rousseau in *Man, the State, and War* (hereafter *MSW*) by showing that Rousseau uses structural variables such as proximity and geography to explain conflict's origins, vindicating human nature. To show that Waltz remained under Rousseau's influence later on, the next section turns to Waltz's mature theory. It shows that Waltz levels a Rousseauian objection against Hobbes's motivational pessimism late in his career; that Waltz's concept of structure in *Theory of International Politics* (hereafter *TIP*)

Press, 1990); Mark B. Salter, 'Anarchy, scarcity, nature: Rousseau's stag hunt and the Arctic walrus hunt compared', *American Political Science Review*, 118:3 (2024), pp. 1145–1157; Céline Spector, 'Who is the author of the abstract of Monsieur l'Abbé de Saint-Pierre's "Plan for Perpetual Peace"? From Saint-Pierre to Rousseau', *History of European Ideas*, 39:3 (2013), pp. 371–393; Michael C. Williams, *The Realist Tradition and the Limits of International Relations* (Cambridge: Cambridge University Press, 2005), pp. 52–81.

[11] Annette Freyberg-Inan, 'Rational paranoia and enlightened machismo: The strange psychological foundations of realism', *Journal of International Relations and Development*, 9:3 (2006), pp. 247–268; Daniel Jacobi and Annette Freyberg-Inan, 'The forum: Human being(s) in International Relations', *International Studies Review*, 14:4 (2012), pp. 645–665; Waltz, *TIP*, pp. 91–2.

[12] Mearsheimer, *The Tragedy of Great Power Politics*, pp. 3, 32; John J. Mearsheimer and Sebastian Rosato, *How States Think: The Rationality of Foreign Policy* (New Haven, CT: Yale University Press, 2023), p. 101.

[13] Ken Booth and Nicholas Wheeler, *Security Dilemma: Fear, Cooperation, and Trust in World Politics* (Basingstoke: Palgrave MacMillan, 2008), p. 7; Joshua Baker, 'The empathic foundations of security dilemma de-escalation', *Political Psychology*, 40:6 (2019), pp. 1251–1266; Stephen M. Walt, 'The geopolitics of empathy', *Foreign Policy* (2021), available at: {https://foreignpolicy.com/2021/06/27/the-geopolitics-of-empathy/}.

[14] Dale C. Copeland, 'Rationalist theories of international politics and the problem of the future', *Security Studies*, 20:3 (2011), pp. 441–450; Sebastian Rosato, *Intentions in Great Power Politics: Uncertainty and the Roots of Conflict* (New Haven, CT: Yale University Press, 2021), pp. 71–2.

follows Rousseau's positional model of structure rather than Kant's transformational model; and that Waltz's argument for the functional equivalence of units in anarchy implies a Rousseauian anthropology when applied to individuals. The fourth section then examines what is gained by reading neorealism as Rousseauian, even beyond Waltz. The final section sums up the argument and its implications.

Terms and scope

This section lays the groundwork for the paper by defining key terms and delimiting the scope of the argument. I begin by examining the security dilemma, the tragic logic at the heart of neorealism. To explain what makes it tragic, I then define tragedy before explaining why I read Rousseau as a tragic thinker but not Hobbes. Throughout the paper, I emphasise tragedy because it provides a helpful lens for seeing commonalities among – and differences between – various strands of realism.

First, what is the security dilemma? By most accounts, the security dilemma refers to the tendency of security measures taken by one state under anarchy to decrease others' security, leading to self-defeating spirals of security competition.[15] Although offensive and defensive realists disagree about its severity and pervasiveness,[16] they generally agree that four conditions must be present for its occurrence. They are (1) anarchy, (2) offensive capabilities, (3) the desire to survive, and (4) uncertainty about others' intentions. Moreover, IR theorists of all stripes agree that anarchy is a situation in which no overarching authority has enforcement power over politically autonomous actors.[17]

Since John Herz first coined the term, the security dilemma has almost always been described as tragic.[18] As Bob Jervis once put it, 'when the security dilemma is at work, international politics can be seen as tragic in the sense that states may desire – or at least be willing to settle for – mutual security, but their own behavior puts this very goal further from their reach'.[19] John Mearsheimer goes even further. In his estimation, the security dilemma represents the 'basic logic' of the tragedy of great power politics.[20]

But what is tragedy? A review of the literature on tragedy in IR is beyond this paper's scope, though scholars usually start from Aristotle's definition.[21] Rousseau himself was familiar with

[15]John H. Herz, 'Idealist internationalism and the security dilemma', *World Politics*, 2:2 (1950), pp. 157–180 (p. 157); Robert Jervis, 'Cooperation under the security dilemma', *World Politics*, 30:2 (1978), pp. 167–214 (p. 169); Charles L. Glaser, 'The security dilemma revisited', *World Politics*, 50:1 (1997), pp. 171–201 (p. 171).

[16]On offence–defence theory and its critics, see Richard K. Betts, 'Must war find a way? A review essay', *International Security*, 24:2 (1999), pp. 166–198; Robert Jervis, *The Meaning of the Nuclear Revolution* (Ithaca, NY: Cornell University Press, 1989); Kier A. Lieber and Daryl G. Press, *The Myth of the Nuclear Revolution* (Ithaca, NY: Cornell University Press, 2020); Kier A. Lieber, *War and the Engineers* (Ithaca, NY: Cornell University Press, 2008); Rosato, *Intentions in Great Power Politics*; Stephen van Evera, 'Offense, defense, and the causes of war', *International Security*, 22 (1998), pp. 5–43. Because Waltz saw his defensive neorealism as compatible with Mearsheimer's offensive neorealism, and because Mearsheimer introduces a defence-dominant variable (the stopping power of water) into his theory, I do not distinguish sharply between offensive and defensive realism in this paper. Cf. Anna Cornelia Beyer, *Kenneth Waltz's Life and Thought: An Interview* (Morrisville, NC: Lulu Press, 2015), pp. 55–6.

[17]Jack Snyder, 'Anarchy and culture: Insights from the anthropology of war', *International Organization*, 56:1 (2002), pp. 7–45 (p. 7 n. 2).

[18]Herz, 'Idealist internationalism and the security dilemma'; John H. Herz, *Political Realism and Political Idealism: A Study in Theories and Realities* (Chicago: University of Chicago Press, 1951); cf. Shiping Tang, 'The security dilemma: A conceptual analysis', 18:3 (2009), pp. 587–623 (pp. 593–5).

[19]Robert Jervis, 'Was the Cold War a security dilemma?', *Journal of Cold War Studies*, 3:1 (2001), pp. 36–60 (p. 36).

[20]Mearsheimer, *The Tragedy of Great Power Politics*, p. 35.

[21]Toni Erskine and Richard Ned Lebow, 'Learning from tragedy and refocusing International Relations', in Toni Erskine and Richard Ned Lebow (eds), *Tragedy and International Relations* (New York: Palgrave MacMillan, 2012), pp. 185–217 (pp. 186–7); Mervyn Frost, 'Tragedy, ethics and International Relations', in Toni Erskine and Richard Ned Lebow (eds), *Tragedy and International Relations* (New York: Palgrave MacMillan, 2012), pp. 21–43 (pp. 25–30); *Aristotle's Poetics*, trans. S. H. Butcher (New York: Hill and Wang, 1961), pp. 61–3, 76–7.

Aristotle's definition and references it in his *Letter to d'Alembert*.[22] For Aristotle, there are two essential elements of any tragedy. They are plot, i.e. 'the arrangement of the incidents', and character, i.e. 'that in virtue of which we ascribe certain qualities to the agents'.[23] Whereas comedies depict 'characters of a lower type', Aristotle observes that tragedies imitate 'characters of a higher type'.[24] Despite the good qualities of the tragic hero, the structure of the plot leads to an outcome that produces 'pity and fear' on the hero's behalf.[25] This is because tragedy involves a reversal of fortune that contradicts the protagonist's good qualities.[26] As Aristotle puts it, 'pity is aroused by unmerited misfortune, fear by the misfortunes of a man like us'.[27]

Accordingly, we can define tragedy as a discordant narrative wherein the structure of events leads to a bad outcome (usually death), contradicting the protagonist's good qualities and prompting pity and fear in the spectator. *Oedipus Rex* may serve as a paradigmatic example because, when its protagonist kills his father and sleeps with his mother, he harms others and transgresses social norms, not due to transgressive desires, but due to the situation in which he acts.[28] In like manner, Alan Collins argues that the 'very character of the security dilemma is one of tragedy, since … the participants' intent was benign'.[29]

Similarly, I interpret Rousseau's *Second Discourse* (hereafter *SD*) as a tragedy because the situation in which natural man is embedded causes an unmerited reversal in humanity's fortune.[30] Although some Rousseau scholars criticise him for succumbing to tragedy,[31] I follow others who affirm tragedy as a constitutive element of Rousseau's thought.[32] In the *SD*, humanity's disfigurement occurs through the advent of a situation 'in which everyone profits from the others' misfortune' (*SD*, 203). Early on, Rousseau warns his reader, 'You will look for the age at which you would wish your Species had stopped. Discontented, you might wish perhaps to go backward' (*SD*, 136). The story's arc is one of tragic reversal with a naturally good protagonist coming to a bad end due to a change in the environment.

By contrast, Hobbes's narrative does not fit tragedy's narrative arc. In Hobbes's telling, life in man's natural condition is 'solitary, poore, nasty, brutish, and short'.[33] Nevertheless, establishing states puts an end to the war of all against all caused by 'the nature of man'.[34] Although Hobbes thinks that states remain in a 'posture of war' towards one another once established, he concludes that, 'because they uphold thereby, the Industry of their Subjects; there does not follow from it, that misery, which accompanies the Liberty of particular men'.[35] Whereas the protagonists of tragedies

[22] Jean-Jacques Rousseau, *Politics and the Arts: Letter to M. D'Alembert on the Theatre*, trans. Allan Bloom (Ithaca, NY: Cornell University Press, 1968), p. 28.

[23] *Aristotle's Poetics*, p. 62.

[24] *Aristotle's Poetics*, pp. 59–60.

[25] *Aristotle's Poetics*, pp. 61, 70.

[26] *Aristotle's Poetics*, pp. 63, 72–3, 76.

[27] *Aristotle's Poetics*, p. 76.

[28] Rousseau, *Politics and the Arts*, pp. 32–3; Michael Spirtas, 'A house divided: Tragedy and evil in realist theory', *Security Studies*, 5:3 (1996), pp. 385–423 (p. 387 n. 5).

[29] Allan Collins, 'State-induced security dilemma: Maintaining the tragedy', *Cooperation and Conflict*, 39:1 (2004), pp. 27–44 (p. 29).

[30] Jean-Jacques Rousseau, 'Discourse on the Origin and Foundation of Inequality among Men, or the Second Discourse', in Victor Gourevitch (ed.), *The Discourses and Other Early Political Writings* (Cambridge: Cambridge University Press, 2019), pp. 115–229.

[31] Steven Johnston, *Encountering Tragedy: Rousseau and the Project of Democratic Order* (Ithaca, NY: Cornell University Press, 1999); Benjamin Storey and Jenna Silber Storey, *Why We Are Restless* (Princeton, NJ: Princeton University Press, 2021), pp. 99–139.

[32] Mark S. Cladis, 'Tragedy and theodicy: A meditation on Rousseau and moral evil', *Journal of Religion*, 75:2 (1995), pp. 181–199; Jason Neidleman, 'Politics and tragedy: The case of Rousseau', *Political Research Quarterly*, 73:2 (2020), pp. 464–475; Alice Ormiston, 'A tragic desire: Rousseau and the modern democratic project', *Telos*, 154 (2011), pp. 8–28.

[33] Thomas Hobbes, *Leviathan*, ed. Richard Tuck (Cambridge: Cambridge University Press, 1996), p. 89.

[34] Hobbes, *Leviathan*, p. 88.

[35] Hobbes, *Leviathan*, p. 90.

come to a bad end, in Hobbes's narrative, naturally competitive and warlike individuals end up faring better than their natures would suggest. Hence, the narrative arc of Hobbes's story does not fit the tragic mould.

Finally, the scope of this paper is limited to questions of philosophical anthropology. I do not claim that all of neorealism is Rousseauian, nor that Waltz takes no inspiration from other thinkers such Immanuel Kant, G. W. F. Hegel, William Graham Sumner, etc. The thesis is simply that neorealism implies a Rousseauian speculative anthropology.

The Waltzian reading of Rousseau

This thesis must remain unintelligible until we understand how neorealists read Rousseau and what Rousseau's conception of human nature is. Accordingly, this section proceeds by recovering Rousseau's image of human nature as described in Part I of the *SD*, a text upon which Kenneth Waltz drew heavily. Dan Deudney argues that Waltz finds in it not only the 'stag hunt' but also reasons why anarchy 'is simultaneously undesirable, difficult to overcome, and an independent cause of conflict among states'.[36] Indeed, Waltz concluded his first book with the claim that 'Rousseau's explanation of war's origin among states is, in broad outline, the final one'.[37]

To show why Waltz reached this conclusion, this section then reads Part II of the *SD* through a Waltzian lens. It concludes that Waltz was a better reader of Rousseau than is generally acknowledged and that he was correct to see Rousseau as giving a structural account of conflict's origins, combining optimism about human motivations with pessimism about relations between groups of human beings.

Rousseau's first-image optimism in SD *Part I*

In his dissertation and first book, Waltz identifies three 'images' used to explain war. The first is human nature, the second the state, and the third the system of states.[38] Because Hobbes argued 'during those times men live without a common Power to keep them all in awe, they are in that condition which is called Warre',[39] Hobbes's theory appears to fit the third image. Yet Hobbes also claims to have based his state of nature on an 'Inference, made from the Passions', locating the 'principall causes of quarrel' in 'the nature of man'.[40]

In part because of this, Waltz picks Rousseau over Hobbes as the prime exemplar of the third image. Waltz devotes great attention to Rousseau's critique of Hobbes, while acknowledging that 'from one point of view this criticism of Hobbes is mere quibbling'.[41] That Rousseau's critique of Hobbes may appear 'mere quibbling' is owed to the fact that Rousseau arrives at his conclusion 'simply by starting one step further back'.[42] Waltz, then, viewed Rousseau's analysis of conflict's origins as more radical (in the original sense of the word) than Hobbes's account.

[36]Daniel H. Deudney, 'Anarchy and violence interdependence', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 17–34 (p. 23).

[37]Kenneth Waltz, *Man, the State, and War: A Theoretical Analysis* (New York: Columbia University Press, 2018 [1959]), p. 231.

[38]Waltz later dubbed the first and second images 'reductionist' since they use agents' attributes to explain outcomes, whereas the third image makes structure an independent cause. See Waltz, *TIP*, p. 18. For a critique, see Alexander Wendt, 'The agent–structure problem in International Relations theory', *International Organization*, 41:3 (1987), pp. 335–370. On the ethical implications, see Catherine Lu, 'Agents, structures and evil in world politics', *International Relations*, 18:4 (2004), pp. 498–509. Although scholars agree structure undergirds neorealism, they disagree about how Waltz conceptualised it: Goddard and Nexon, 'Paradigm lost?'; Jeffrey W. Legro and Andrew Moravcsik, 'Is anybody still a realist?', *International Security*, 24:2 (1999), pp. 5–55; Ole Wæver, 'Waltz's theory of theory: The pictorial challenge to mainstream IR', in Ken Booth (ed.), *Realism and World Politics* (New York: Routledge, 2011), pp. 67–88.

[39]Hobbes, *Leviathan*, p. 88.

[40]Hobbes, *Leviathan*, pp. 88–9.

[41]Waltz, *Man, the State, and War*, p. 166.

[42]Ibid.

But how does Rousseau go 'further back' than Hobbes? Waltz observes that Rousseau concurred with Montesquieu's judgement that Hobbes had erred in 'assuming that men in nature possess all of the characteristics and habits they acquire in society but without the constraints imposed by society'.[43] Ultimately, Hobbes's definition of human nature comes up short because it 'reflects both man's nature and the influence of his environment'.[44] 'Theoretically at least', Waltz continues, 'one can strip away environmentally acquired characteristics and arrive at a view of human nature itself. Rousseau himself has advanced "certain arguments, and risked some conjectures", to this end'.[45] Thus, Waltz thinks that Rousseau's method of controlling for environmental influences is at least potentially able to provide a clear view of human nature, while Hobbes's method lacks this potential.

There is much truth in Waltz's reading of Rousseau. To see the differences between Rousseau's account of human motivations and Hobbes's, the *SD* is the best place to start. Although Rousseau briefly mentions Hobbes in the *First Discourse*,[46] the *SD* 'confronts Hobbes's view of human nature head-on'.[47] Rousseau writes that in the *SD* he most clearly developed 'the fundamental principle of all morality about which I have reasoned in all my Writings'.[48] Rousseau's fundamental principle is that 'man is a naturally good being'.[49]

As Waltz perceived, Rousseau's method for disclosing fundamental human nature, which he takes to be fundamentally good or peaceable, is to remove everything artificial from humanity. Rousseau does this by altering the proximity of individuals until the spread of language becomes impossible (*SD*, 148–9), while holding both human biology and psychology constant (*SD*, 129, 137). Rousseau compares his method to the 'hypothetical and conditional reasonings' that 'our Physicists daily make regarding the formation of the World' (*SD*, 135). But whereas physicists form hypotheses about matter and motion, Rousseau claims to form 'conjectures based solely on the nature of man and of the Beings around him' (*SD*, 136).

While Hobbes points to Indigenous Americans and 'Kings, and Persons of Soveraigne authority' as examples of people living in the state of nature,[50] Rousseau says expressly that his own state of nature never existed historically.[51] Rousseau's state of nature is meant to be a 'true fiction', as it were.[52] In Rousseau's own words, it is 'better suited to elucidate the nature of things than to show their genuine origin' (*SD*, 135). It provides an image that we need to know if we are to know our ourselves (*SD*, 136) but does not represent actual history or an ideal to which we should aspire to return.[53]

What is this image? Rousseau identifies two psychological principles originally at work in human nature. As he puts it, 'one interests us intensely in our well-being and our self-preservation, and the other inspires in us a natural repugnance to seeing any sentient Being, and especially any being like ourselves, perish or suffer' (*SD*, 129). Because the desire for self-preservation

[43] Waltz, *Man, the State, and War*, p. 165.

[44] Waltz, *Man, the State, and War*, p. 166.

[45] Ibid.

[46] Jean-Jacques Rousseau, 'Discourse on the sciences and arts, or the First Discourse', in Victor Gourevitch (ed.), *The Discourses and Other Early Political Writings* (Cambridge: Cambridge University Press, 2019), pp. 1–28 (p. 26).

[47] Roosevelt, *Reading Rousseau in the Nuclear Age*, p. 27.

[48] Jean-Jacques Rousseau, 'Letter to Beaumont', in Christopher Kelly and Eve Grace (eds), *Letter to Beaumont, Letters Written from the Mountain, and Related Writings*, trans. Judith R. Bush and Christopher Kelly (Hanover, NH: Dartmouth College Press, 2012), p. 28.

[49] Ibid.

[50] Hobbes, *Leviathan*, pp. 89–90.

[51] For the colonialist implications of Hobbes's view, see Pat Moloney, 'Hobbes, savagery, and international anarchy', *American Political Science Review*, 105:1 (2011), pp. 189–204. For a defence of Rousseau's, see Charles W. Mills, 'Rousseau, the master's tools, and anti-contractarian contractarianism', *The CLR James Journal*, 15:1 (2009), pp. 92–112.

[52] Emma Planinc, 'The figurative foundations of Rousseau's politics', *Modern Intellectual History* (2023), available at: {https://doi.org/10.1017/S1479244321000688}.

[53] Waltz, *Man, the State, and War*, pp. 5, 166.

(*amour de soi*) is a benign motive, and pity (*pitié*) precludes natural man from committing unnecessary acts of violence, Rousseau thinks that man in the state of nature is 'naturally peaceable'.[54] Accordingly, Rousseau rejects Hobbes's depiction of mankind as inherently warlike or aggressive.

By going one step further back than Hobbes to pre-linguistic human beings, Rousseau can argue:

> Above all, let us not conclude with Hobbes that because he has no idea of goodness man is naturally wicked … this Author should have said that, since the state of Nature is the state in which the care for our own preservation is the least prejudicial to the self-preservation of others, it follows that this state was the most conducive to Peace and the best suited to Mankind (*SD*, 154–5).

Prior to the Hobbesian war of all against all, Rousseau posits a more fundamental 'pure' state of nature.[55] In this pure state of nature, *amour de soi* had not yet turned into *amour-propre*.[56]

Rousseau's third-image pessimism in SD *Part II*

To Waltz's mind, Rousseau's account raises three important questions. Specifically, Waltz asks, '(1) Why, if the original state of nature was one of relative peace and quiet, did man ever leave it? (2) Why does conflict arise in social situations? (3) How is the control of conflict related to its cause?'[57] Let us see how Rousseau answers these questions.

In Rousseau's telling, man leaves the state of nature through a change in his environment, through 'new circumstances' (*SD*, 145). Since Rousseau never describes these circumstances in Part I, we must look for them in Part II. Here, Rousseau imagines several possible events, including floods and earthquakes, bringing an end to the isolation of the pure state of nature. Each possibility involves a natural catastrophe constricting the space where previously isolated individuals were living, bringing them into closer proximity (*SD*, 169). Rousseau thus hypothesises that a change in the positionality of individuals vis-à-vis one another, not a change in the characteristics, motives, or other attributes of the agents, could suffice to explain the deepest roots of conflict.

After Rousseau's natural man leaves the state of nature, 'everything begins to change' (*SD*, 169). As Rousseau puts it elsewhere, 'the first society formed necessarily leads to formation of all the rest. People have to belong to it, or unite to resist it. People have to imitate it, or let themselves be swallowed up by it.'[58] In the *SD*, these changes begin to occur because proximity, a structural rather than agential variable, introduces two conditions absent in the pure state of nature.

First, it introduces interaction under anarchy. Without regular interactions between people, natural inequalities could have no regular effects in the state of nature. Prior to the advent of society, people 'might perhaps meet no more than twice in their life, without recognizing and speaking with one another' (*SD*, 147–8). Accordingly, Rousseau asks, 'even if Nature displayed as much partiality in the distribution of its gifts as is claimed, what advantage would the more favored enjoy at the expense of the others in a state of things that allowed for almost no relations of any sort between them?' (*SD*, 162). Ultimately, the state of war, as opposed to the state of nature, 'is a permanent state presupposing constant relations, and such relations very rarely obtain between one man and

[54]Rousseau, 'Principles of the right of war', p. 159.

[55]The concept of the 'pure' state of nature originated in theological debates about humanity's prelapsarian condition: Victor Gourevitch, 'Rousseau's pure state of nature', *Interpretation*, 16:1 (1988), pp. 23–59 (pp. 25–33); John T. Scott, 'The theodicy of the Second Discourse: The "pure state of nature" and Rousseau's political thought', *American Political Science Review*, 86:3 (1992), pp. 696–711 (pp. 696–704). On state-of-nature arguments in IR generally, see Jan Niklas Rolf, 'The sate of nature analogy in International Relations', *International Relations*, 28:2 (2014), pp. 159–182; Zichao Tong, 'State of nature versus states as firms: Reassessing the Waltzian analogy of structural realism', *International Relations*, 38:4 (2024), pp. 615–634.

[56]Rousseau describes *amour-propre* as 'a relative sentiment, factitious and born in society, that inclines every individual to set greater stock by himself than by anyone else' (*SD*, 224).

[57]Waltz, *Man, the State, and War*, p. 166.

[58]Rousseau, 'Principles of the right of war', p. 161.

another, where among individuals everything is in a continual flux'.[59] Once regular interaction is introduced, however, 'natural inequality imperceptibly unfolds together with unequal associations, and the differences between men, developed by their different circumstances, become more perceptible, more permanent in their effects, and begin to exercise a corresponding influence on the fate of individuals' (*SD*, 174). Thus, interaction is important because anarchy cannot affect relations between individuals, groups, or states that have no relations, as Waltz also points out.[60] By regularising interactions, Rousseau's natural catastrophe enables physical inequalities to have systemic effects they could not have in the pure state of nature.

Second, the closer proximity introduced by the constriction of man's environment enables conflicts to emerge because it creates the need and opportunity for time-consciousness to develop. Rousseau tells us that 'knowledge of death and of its terrors was one of man's first acquisitions on moving away from the animal condition' (*SD*, 145). Yet consciousness of death could not have been the first step in moving from the state of nature because other conceptions not yet available to man must have preceded it. In the pure state of nature, man knows only 'the sentiment of present existence'. To become conscious of death, he must first become conscious of himself as a being extended not only in space corporally, as he could always see with his own eyes, but also in time imaginatively, a much more difficult discovery to make insofar as temporal extension can only be observed indirectly via projection of the self into the future.

Whereas man in the pure state of nature would have no sense of vulnerability except in rare moments when facing an immediate threat, time-conscious man in an anarchic setting must wonder whether he will face threats tomorrow, the next day, the day after, *ad mortem*. Rousseau suggests that, however long it took man to realise it, 'as soon as it was found to be useful for one to have provisions for two, equality disappeared, property appeared, work became necessary, and … slavery and misery were soon seen to sprout and grow together with the harvests' (*SD*, 172). To preserve himself, time-conscious man faces a fundamental dilemma: provide for tomorrow at today's expense or enjoy today at tomorrow's loss (*SD*, 173). Like Herz's security dilemma, Rousseau's dilemma is tragic. For Rousseau, the dilemma not only drives a wedge between the demands of self-preservation and those of pity, but also between satisfaction and the means to satisfaction. Thus emerges the unhappy busyness of Rousseau's civilised man, who 'works to death, even rushes toward it in order to be in a position to live' (*SD*, 192). In Rousseau's narrative, the emergence of time-consciousness, and thereby uncertainty about the future, marks the beginning of competition for power among individuals and eventually between groups.

As Waltz was fond of pointing out, Rousseau's account makes clear that the roots of conflict lie 'not only in the actors but also in the situation they face'.[61] In the age-old debate about the relationship between agents and social structures, or nature versus nurture, Rousseau places emphasis on the efficacy of the latter over that of the former. 'In the third image', writes Waltz, 'the old problem of political philosophy – do men create the societies and states in which they live or do those societies and states, so to speak, remake the men who live in them? – here appears in a different form. Rousseau has argued that the sources of conflict are not so much in men as they are in society.'[62] Waltz thus took the famous 'stag hunt' from the *SD* as just one example of a broader tendency in Rousseau's philosophy.[63]

Although Michael Williams criticises Waltz's emphasis on the stag hunt because it occurs at 'an immature point in the development of humanity',[64] this criticism forgets that the *SD* centres around

[59] Rousseau, 'Principles of the right of war', p. 160.
[60] Waltz, *TIP*, p. 79.
[61] Waltz, *Man, the State, and War*, p. 170.
[62] Kenneth Waltz, 'Political philosophy and the study of International Relations', in William T. R. Fox (ed.), *Theoretical Aspects of International Relations* (Notre Dame, IN: University of Notre Dame Press, 1959), pp. 51–68 (pp. 64–5).
[63] Waltz, *Man, the State, and War*, p. 167.
[64] Williams, *The Realist Tradition and the Limits of International Relations*, p. 56.

a narrative of progressive worsening from economic inequality, to political inequality, to absolute despotism.[65] More importantly, by focusing too much on one example that Waltz gives of a general tendency in Rousseau's thought, it fails to ask whether this general tendency might in fact be better demonstrated by other passages in Rousseau's writings. As this section has shown, other passages bear out the tendency Waltz identified in Rousseau's thought.

The Rousseauian reading of Waltz

In the previous section, we saw that Rousseau's tragic account of conflict's origins deeply influenced Kenneth Waltz's early thought. Still, one may wonder whether Rousseau's political philosophy had lasting influence on Waltz beyond those early years. After all, Waltz published a piece on Immanuel Kant in the American Political Science Review in 1962, and recent studies emphasise the importance of Kant for Waltz.[66] The conventional wisdom, however, remains that Waltz was Hobbesian.[67]

This section argues in three steps that Waltz's mature writings retain Rousseau's influence. First, it recovers an overlooked objection that Waltz makes to Hobbes's anthropological pessimism. Second, it demonstrates that Waltz's conception of structure in *TIP* owes more to Rousseau than to Kant. Third, it shows that Waltz discovered in Rousseau the insight that structure itself can shape the motives of actors in a system, whether they be people or states. Because of the 'sameness effect', which Waltz found pre-eminently in Rousseau, units in anarchic systems are pushed to pursue security above all other ends. Because self-preservation is a benign motive, Waltz follows Rousseau in thinking that the wars that result from security competition are for the most part tragic.

Waltz's Rousseauian critique of Hobbes

To understand Waltz's objection to Hobbes's anthropological pessimism, the best place to begin is his critique of classical realism. Waltz claims that classical realists explained the recurrence of power maximisation by inferring an *animus dominandi*. This view, based largely on Morgenthau's statements,[68] has been significantly qualified or rejected, even by structural realists.[69] Nevertheless, it gained great traction within the field. Convention thus maintains that classical realism says that human beings are egoistic power-maximisers from birth. If true, classical realists need only a single assumption to account for competition for relative power – a remarkably parsimonious theory.

Since parsimony and explanatory power provide the standards by which Waltz judges theories,[70] one would expect him to endorse classical realism. After all, there is little in the substance of classical realism with which 'Waltz would disagree', although in 'the terminology of *Theory of International Politics* [classical realism] is metaphysics and he, Waltz, is engaged in science'.[71] In truth, Waltz's rejection of classical realism is not merely a matter of rhetorical difference or of scientific posturing. Waltz objects to classical realism by referring to observable variations in human aggressiveness. His argument just happens to mirror Rousseau's objection to Hobbes's anthropological pessimism.

[65] Ryan Hanley, 'Rousseau's three revolutions', *European Journal of Philosophy*, 29:1 (2021), pp. 105–119 (p. 106).

[66] Prichard, 'Kenneth Waltz's Kantian moral philosophy'; Viotti, *Kenneth Waltz*, pp. 21–5, 109–13.

[67] Donnelly, 'The discourse of anarchy in IR', p. 417; Freyberg-Inan, *What Moves Man*, p. 74; Polansky, 'Drawing out the Leviathan'.

[68] Hans Morgenthau, *Scientific Man versus Power Politics* (Chicago: University of Chicago Press, 1946), pp. 192–5.

[69] For instance, Joseph Parent and Joshua Baron, 'Elder abuse: How the moderns mistreat classical realism', *International Studies Review*, 13:2 (2011), pp. 193–213.

[70] Waltz, *TIP*, pp. 5–13.

[71] Brown, 'Structural realism, classical realism and human nature', p. 151.

In the *Leviathan*, Hobbes had claimed that he could deduce his warlike state of nature from his theory of the passions.[72] He thought that the motives of gain, safety, and reputation in particular lead men to invade one another.[73] At other times, however, Hobbes identifies the 'art of words, by which some men can represent to others, that which is Good, in the likenesse of Evill; and Evill, in the likenesse of Good' as a major source of conflict.[74]

Despite the complexity of parsing Hobbes's views, Rousseau's reply was simple. Alluding to Hobbes's description of natural equality as the ability of even the weakest human being to kill the strongest human being while sleeping,[75] Rousseau argues that:

> If that mutual and destructive enmity was part of our constitution, then it would still make itself felt, and would thrust us back despite ourselves through every social chain. The dreadful hatred of humanity would gnaw at man's heart. He would be anguished by the birth of his own children, would rejoice at the death of his brethren – and when he found one of them asleep, his first impulse would be to kill him.[76]

Rousseau thus rejects Hobbes's anthropological pessimism because people behave aggressively in some contexts but not others. As a result, aggressiveness must be a variable not a constant. It cannot, then, be caused by invariant human nature.

Similarly, for the young Waltz, the problem with first-image pessimism was that it could not explain variation in human behaviour under varying conditions.[77] The mature Waltz recognised that this put him in disagreement with Hobbes. In an article published 31 years after *MSW*, Waltz acknowledged that 'both Hobbes and Morgenthau see that conflict is in part situationally explained'.[78] Nevertheless, Waltz concluded that both Hobbes and Morgenthau err in assuming 'that even were it not so, *pride, lust*, and *the quest for glory* would cause the war of all against all to continue indefinitely'.[79]

Put simply, then, Waltz's problem with Hobbes's theory is that it assumes that human beings have motives which would make them warlike even in the absence of situational constraints. Since Waltz found this critique of Hobbes in Rousseau, Waltz's neorealism owes an intellectual debt to Rousseau that it does not owe to Hobbes.

Waltz's Rousseauian conception of structure

Although Waltz made a Rousseauian argument against Hobbes's human-nature pessimism, one might still object that Waltz thought of himself more as a Kantian than a Hobbesian. According to Waltz's biographer, for instance, Waltz saw himself as kind of a Kantian transcendental idealist rather than a Hobbesian materialist.[80] Be that as it may, since this paper's scope is limited to questions of philosophical anthropology, we must ask whether this Kantian influence extends beyond Waltz's ontology.

In this subsection, I argue that, at least with regard to neorealism's key concept, structure, Waltz follows Rousseau rather than Kant. In the subsequent subsection, I tie Waltz's Rousseauian conception of structure to his Rousseauian conception of human nature. But first we must determine whether Waltz's concept of structure is closer to Kant's or to Rousseau's.

[72]Hobbes, *Leviathan*, p. 89.
[73]Hobbes, *Leviathan*, p. 88.
[74]Hobbes, *Leviathan*, pp. 119, 74; Arash Abizadeh, 'Hobbes on the causes of war: A disagreement theory', *American Political Science Review*, 105:2 (2011), pp. 298–315.
[75]Hobbes, *Leviathan*, pp. 86–7.
[76]Rousseau, 'Principles of the right of war', p. 157.
[77]Waltz, *Man, the State, and War*, pp. 32–3.
[78]Waltz, 'Realist thought and neorealist theory', p. 35.
[79]Ibid., emphasis added.
[80]Viotti, *Kenneth Waltz*, p. 4.

We saw above that Rousseau develops a positional conception of structure in Part II of the *SD*. Rousseau's explanatory variables are things like proximity, geography, and inequalities of strength or power. As these variables change throughout his narrative, he tracks 'how the human soul and passions, by imperceptible adulterations *so to speak* change in Nature' (*SD* 191, italics added).

Much like Rousseau, Kant has a systemic or third-image story to tell.[81] Yet Kant's conception of structure differs significantly from Rousseau's. Whereas Rousseau relies upon a positional conception of structure, Kant develops a transformational model.[82] The positional model views structure as the environment in which action takes place, while the transformational model views structure as a means to action and gives 'central ontological position to social rules, both constitutive and regulative, intended and unintended'.[83] Ultimately, Kant sees anarchy as itself a means of transforming the international system in his essays on perpetual peace and on world history.[84] As Ewan Harrison puts it, 'Kant envisages an open ended process in which anarchy, itself the unintended outcome of state action, is susceptible to deliberate change through the gradual extension of voluntary agency'.[85] For Kant, then, structure leads to its own transformation as the agents in the system unwittingly revolutionise it.

Is Waltz's model of structure closer to Rousseau's positional model or to Kant's transformational model? According to Harrison, 'Waltz relies on a positional model of structure, focusing on a unit's position within a given arrangement of actors'.[86] In *TIP*, Waltz makes it very clear that 'how units stand in relation to another, the way they are arranged or positioned, is not a property of the units. The arrangement of units is a property of the system.'[87] As a result, Waltz insists that 'a structure is defined by the arrangement of its parts'.[88] Indeed, by focusing on variables at the systemic level, Waltz concludes that 'one arrives at a purely positional picture of society'.[89] Unlike Kant, then, who adopts a transformational model, Waltz shares with Rousseau a positional model of structure.

So, even if Waltz is a Kantian at the level of metatheory, Waltz stands much closer to Rousseau when it comes to key explanatory variable of *TIP*. In fact, on specific questions of causality in international relations, Waltz seems to have preferred Rousseau to Kant, despite recognising certain similarities between them. In *MSW*, for instance, Waltz states plainly that 'the explanations of Rousseau and Kant are similar, but Rousseau's is the more consistent and complete'.[90]

Waltz, Rousseau, and the first image reversed

Despite Waltz's Rousseauian conception of structure, one may still wonder what the connection is between Waltz's conception of structure and his thinking about motivations. Because Waltz moved increasingly away from political theory and towards International Relations after 1959, one may also wonder to what extent Waltz remained interested in political philosophers such as Rousseau. According to his biographer, Waltz considered himself a political theorist to the end of his life.[91]

[81] Kenneth Waltz, 'Kant, liberalism, and war', *American Political Science Review*, 56:2 (1962), pp. 331–340; Wade L. Huntley, 'Kant's third image: Systemic sources of the liberal peace', *International Studies Quarterly*, 40:1 (1996), pp. 45–76.

[82] Ewan Harrison, 'Waltz, Kant and systemic approaches to International Relations', *Review of International Studies*, 28:1 (2002), pp. 143–162 (p. 149).

[83] David Dessler, 'What's at stake in the agent–structure debate?', *International Organization*, 43:3 (1989), pp. 441–473 (p. 444).

[84] E.g. Immanuel Kant, 'Idea for a universal history with a cosmopolitan purpose', in Hans Reiss (ed.), *Kant: Political Writings*, trans. Hugh Barr Nisbet (Cambridge: Cambridge University Press, 1991), pp. 41–53 (p. 41).

[85] Harrison, 'Waltz, Kant and systemic approaches to International Relations', p. 149.

[86] Harrison, 'Waltz, Kant and systemic approaches to International Relations', p. 148.

[87] Waltz, *TIP*, p. 80.

[88] Ibid.

[89] Ibid.

[90] Waltz, *Man, the State, and War*, p. 172.

[91] Viotti, *Kenneth Waltz*, p. 35.

Consistent with this, Waltz devotes two pages, or roughly 1 per cent, of *TIP* to refuting Stanley Hoffmann's interpretation of Rousseau. Hoffmann had identified Rousseau's solution to the problem of war and peace as a second-image solution.[92] Waltz's response to Hoffmann indicates that he discovered the 'sameness effect' first and foremost in Rousseau's writings.[93] Because the sameness effect describes how the structure of the system affects the qualities or attributes of the units in the system, it enables us to bridge the gap between Waltz's Rousseauian conception of structure and his conception of human nature.

Whatever one thinks of Waltz's approach to reading political philosophy, he was at least sufficiently careful to recognise when other interpreters had gotten Rousseau wrong.[94] To Waltz's mind, Stanley Hoffmann had attributed 'to Rousseau conclusions about international politics that would subvert the whole of his political philosophy'.[95] 'Preeminently among political theorists', Waltz replied to Hoffmann, 'Rousseau emphasized … the context of action must always be considered, whether dealing with men or with states, for the context will itself affect *attributes and purposes and behavior* as well as alter outcomes.'[96] Although Hoffmann's reading of Rousseau is more nuanced than Waltz sometimes allows, Waltz's critique is well grounded.[97] But what matters here is that Waltz saw Rousseau as the pre-eminent representative of the structural approach to International Relations and identified the sameness effect as his key insight.

Despite describing structure as a 'permissive condition' of war in *MSW* and theory as a 'useful fiction' in *TIP*, with the sameness effect Waltz oversteps these boundaries and makes structure as an 'efficient cause' in its own right. One naturally wonders how structure, if it is merely a mental construct of the theorist, can 'shape and shove' the units in the system.[98] Waltz argues that structure shapes and shoves units 'through socialization of the actors and competition among them'.[99] Since units that do not pursue self-preservation are weeded out under anarchy, those that remain must generally prioritise self-preservation. Rousseau called this the 'Law of Sparta' (*SD*, 138). For Waltz, it is an effect of the ordering principle of the system. 'In anarchy', Waltz says explicitly, 'security is the highest end'.[100] Indeed, this applies to any 'units in a condition of anarchy – be they *people, corporations, states, or whatever*'.[101] Like Rousseau, Waltz thinks that the units in the system are plastic: the situation can determine the motives and qualities of the agents.

Because of such claims, Randall Schweller has concluded that neorealists think anarchy makes states security seekers.[102] Similarly, Joseph Grieco has argued that anarchy makes states 'defensive

[92] Hoffmann, 'Rousseau on war and peace', p. 329.

[93] The 'sameness effect' describes how anarchy compels units to adapt in ways that cause them to resemble one another. Eben Coetzee, 'No rest for the weary: Competition and emulation in international politics. A Waltzian perspective', *Politikon*, 46:1 (2019), pp. 1–19; Michael C. Desch, 'War and strong states, peace and weak states?', *International Organization*, 50:2 (1996), pp. 237–268; Emily O. Goldman and Richard B. Andres, 'Systemic effects of military innovation and diffusion', *Security Studies*, 8:4 (1999), pp. 79–125; Christopher Layne, 'The unipolar illusion: Why new great powers will rise', *International Security*, 17:4 (1993), pp. 5–51; Morgan MacInnes, Ben Garfinkel, and Allan Dafoe, 'Anarchy as architect: Competitive pressure, technology, and the internal structure of states', *International Studies Quarterly*, available at: {https://doi.org/10.1093/isq/sqae111}; Cameron G. Thies, 'State socialization and structural realism', *Security Studies*, 19:4 (2010), pp. 689–717. Cf. Waltz, *TIP*, p. 128.

[94] On Waltz's hermeneutics, see MacKay, 'Kenneth Waltz's approach to reading classic political theory and why it matters'.

[95] Waltz, *TIP*, p. 48.

[96] Waltz, *TIP*, p. 47, italics added.

[97] Rousseau concludes that 'it is not impossible that a well-governed republic wage an unjust war', and that, 'since the well-being of one is another's harm, in keeping with the law of nature each gives preference to himself'. Cf. Jean-Jacques Rousseau, 'Discourse on political economy', in Victor Gourevitch (ed.), *The Social Contract and Other Later Political Writings* (Cambridge: Cambridge University Press, 2019), pp. 3–38 (p. 8); Rousseau, 'Principles of the right of war', p. 155.

[98] Colin Wight, *Agents, Structures and International Relations* (Cambridge: Cambridge University Press, 2006) p. 97.

[99] Waltz, *TIP*, p. 74.

[100] Waltz, *TIP*, p. 126; similarly, Waltz, *Man, the State, and War*, pp. 203, 215.

[101] Waltz, *TIP*, p. 111, italics mine.

[102] Randall L. Schweller, 'Neorealism's status-quo bias: What security dilemma?', *Security Studies*, 5:3 (1996), pp. 90–121.

positionalists'.[103] Yet the claim of Waltz and Rousseau is more radical. *Any* unit in anarchy, whether it be a person, corporation, or state, so long as it lacks certainty about others' future intentions, will be compelled to compete for security. A level of similarity thus arises among units in anarchy.

This emergent likeness of units in anarchy has long made it possible to speak of a 'second image reversed'.[104] Yet Waltz's story about the functional equivalence of units in anarchy implies the conceptual possibility of a 'first image reversed'. Although Waltz thinks there are important differences between people and states,[105] neorealists like Waltz think that structure pushes units in anarchy to emulate the best practices of others.[106] As Rousseau perceived, the strong innovate, and the weak find it 'safer to imitate them than to try to dislodge them' (*SD*, 168). Based on Waltz's statements about survival being the chief aim in anarchy,[107] we can surmise that individuals placed in a state of nature would be socialised and selected by the structure of the system to pursue self-preservation first and foremost. As Waltz himself observed, 'chiliastic rulers occasionally come to power. In power, most of them quickly change their ways.'[108]

Following Waltz's lead, scholars have begun working on the 'first image reversed' empirically.[109] Philosophically, however, what is interesting about the first image reversed is that it points to the malleability of the first image. Given the Rousseauian roots of Waltz's realism, it is not surprising that those following Waltz should arrive here. For Rousseau, too, a distinguishing feature of human beings is the malleability of their nature, dependent upon circumstances (*SD*, 144). That Waltz thinks certain motivations remain constant despite this malleability is owed to the fact that anarchy is a constant and that anarchy compels units to pursue survival first and foremost. Since survival is a morally benign end, the conflicts that result from the security-seeking in anarchy are for the most part tragedies, much as Rousseau envisioned in the *SD*.

Rousseauian realism beyond Waltz

In the previous section, we saw that Waltz remained deeply influenced by Rousseau in his mature writings. In this section, I show what scholars gain by reading neorealist narratives about conflict's origins as Rousseauian tragedies. Although neorealists are a heterogeneous bunch, the lens of Rousseauian tragedy brings three disparate elements of the school into sharper focus and greater coherence.

First, Rousseau's account makes clear that security spirals caused by the pursuit of self-preservation under anarchy are what Catherine Lu calls 'tragedies of innocent suffering'.[110] Ultimately, the dynamics of security competition described by prominent neorealists such as John Mearsheimer depict a situation in which good people cause bad things to happen. Mearsheimer's model assumes a certain level of first-image optimism, though it combines this with third-image pessimism.

Second, Rousseau's account shows why the security dilemma represents a moral dilemma. This is clearest in Herz's original formulation of it, where he motivates the two horns of the dilemma using the same motivations that Rousseau assumes in the *SD*: pity and self-preservation.

[103] Joseph M. Grieco, *Cooperation among Nations: Europe, America, and Non-tariff Barriers to Trade* (Ithaca, NY: Cornell University Press, 1990), p. 36.

[104] Peter Gourevitch, 'The second image reversed: The international sources of domestic politics', *International Organization*, 32:4 (1978), pp. 881–912; Seva Gunitsky, *Aftershocks: Great Powers and Domestic Reforms in the Twentieth Century* (Princeton, NJ: Princeton University Press, 2017); Bruce Porter, *War and the Rise of the State: The Military Foundations of Modern Politics* (New York: Free Press, 1994).

[105] Tong, 'State of nature versus states as firms', pp. 620–1.

[106] Parent and Rosato, 'Balancing in neorealism'.

[107] Waltz, *TIP*, p. 126; Waltz, *Man, the State, and War*, p. 215.

[108] Waltz, *TIP*, p. 128.

[109] Joshua D. Kertzer and Dustin Tingley, 'Political psychology in International Relations: Beyond the paradigms', *Annual Reviews of Political Science*, 21 (2018), pp. 319–339 (pp. 329–30).

[110] Catherine Lu, 'Tragedies and International Relations', in Toni Erskine and Richard Ned Lebow (eds), *Tragedy and International Relations* (New York: Palgrave MacMillan, 2012), pp. 158–171 (p. 159).

Although pity plays a less prominent role in subsequent theorising about the security dilemma, neorealists continue to encourage greater strategic empathy. Rousseauian pity may, then, play an underappreciated role in neorealist arguments.

Third, Rousseau's account clarifies the logic of the security dilemma itself. If we had perfect information about the future or about the motives and intentions of others, the security dilemma could probably be overcome. But because of the type of creatures that we are – creatures that can barely know our own motives in the present, much less what others may intend in the future – it is probable that we will never transcend the security dilemma. Although this might seem to implicate human nature, Rousseau's account absolves human nature by showing that anxiety about the future depends upon non-innate conditions such as anarchy and temporal self-consciousness.

Fear and security-seeking in neorealism

The most prominent neorealist since Waltz is probably John Mearsheimer. Like Waltz, Mearsheimer ends up mirroring many of Rousseau's pessimistic conclusions about interstate relations in his *Tragedy of Great Power Politics*. For instance, Mearsheimer writes that 'the structure of the international system forces states *which seek only to be secure* nonetheless to act aggressively toward each other'.[111] He then surmises that 'states recognize that the more powerful they are relative to their rivals, the better their chances of survival. Indeed, the best guarantee of survival is to be a hegemon, because no other state can seriously threaten such a mighty power.'[112] Rousseau similarly argues that each state 'feels weak so long as any are stronger than it. Its security and self-preservation require it to make itself more powerful than all its neighbors; it can increase, sustain, and exert its forces only at their expense.'[113]

Like Rousseau, Mearsheimer sees this situation as tragic. He concludes, 'this situation, which no one consciously designed or intended, is genuinely tragic. Great powers that have no reason to fight each other – *that are merely concerned with their own survival* – nevertheless have little choice but to pursue power and to seek to dominate the other states in the system.'[114] Put simply, the anarchic environment compels the units to compete for power and sometimes to act aggressively, contradicting the 'rather harmless goal' of the actors.[115]

One might object to a Rousseauian interpretation of Mearsheimer by pointing to passages where he identifies himself as a Hobbesian.[116] The problem with this objection is that Mearsheimer's account of conflict's origins differs fundamentally from Hobbes's. Whereas Hobbes highlights the role of passions over reason in the origination of war,[117] Mearsheimer thinks that dispassionate calculation drives aggression.[118] Likewise, Rousseau thinks that the emergence of the state of war 'demands a cool head and reason'.[119] Not only that, but Mearsheimer also rejects arguments that evolution under anarchy has predisposed human nature towards aggression because he denies that 'humans are naturally aggressive'.[120] In short, Mearsheimer follows Rousseau in thinking that agents in anarchy *behave like* power-maximising egoists but *are not by nature* such egoists.

Ultimately, the tragic heuristic used by neorealists such as Mearsheimer implies a combination of third image pessimism with first image optimism because there is nothing particularly tragic

[111] Mearsheimer, *The Tragedy of Great Power Politics*, p. 3, italics added.

[112] Ibid.

[113] Rousseau, 'Principles of the right of war', p. 162.

[114] Mearsheimer, *The Tragedy of Great Power Politics*, p. 3, italics added.

[115] Mearsheimer, *The Tragedy of Great Power Politics*, p. 32.

[116] Mearsheimer, *The Tragedy of Great Power Politics*, p. 424; John J. Mearsheimer, *Great Delusion: Liberal Dreams and International Realities* (New Haven, CT: Yale University Press, 2018) pp. 135–6.

[117] Jerónimo Rilla, 'Are Hobbesian states as passionate as Hobbesian individuals?', *Review of Politics*, 85:3 (2023), pp. 285–303 (p. 287).

[118] Mearsheimer, *The Tragedy of Great Power Politics*, pp. 37–40.

[119] Rousseau, 'Principles of the right of war', p. 156.

[120] Mearsheimer, *Great Delusion*, pp. 43–4 n. 64; citing Dominic D. P. Johnson and Bradley A. Thayer, 'The evolution of offensive realism', *Politics and the Life Sciences*, 35:1 (2016), pp. 1–26 (p. 5).

about malevolent agents causing harm.[121] For power politics to be tragic, malevolent people must be the exception to the rule. Consequently, Mearsheimer makes an exception for Hitler, who is not a tragic figure and is thus an anomaly for his theory.[122] Although one can point to a few anomalies, the neorealist view remains that 'most states are rational most of the time'.[123] For neorealists, human beings are generally benign agents responding rationally to the selective pressures of their environment. The conflicts that result, then, are tragedies of innocent suffering.[124]

Pity and strategic empathy in neorealism

Ken Booth and Nicholas Wheeler have pointed out that, strictly speaking, there is no dilemma in conventional definitions of the security dilemma.[125] To find the dilemma in the security dilemma, one must turn to the theorist who first coined the term. In John Herz's original formulation of it, the security dilemma describes 'a fundamental social constellation, a mutual suspicion and a mutual dilemma: the dilemma of "kill or perish", of attacking first or running the risk of being destroyed'.[126] That the decision between killing or being killed presents a dilemma for Herz is owed to the fact that he thinks 'most human beings', in addition to the motive of survival, 'seem to possess a basic feeling of pity or compassion that is provoked by the observance of the suffering of another human being'.[127] While survival occasionally requires killing (motivating the 'kill' horn of the dilemma), pity gives human beings a general aversion to killing (motivating the 'be killed' horn of the dilemma).

In its original formulation, then, the security dilemma constitutes a *moral dilemma*. Moral dilemmas refer to situations in which an agent is morally motivated to do each of two actions; the agent can do each of the actions; but the agent cannot do both of the actions.[128] The result is that the agent is condemned to moral failure because no matter what they do, they will do something wrong.[129] For Herz, this inevitable moral failure gives rise to feelings of guilt or bad conscience.[130] Yet Herz denies that this is caused by 'such a thing as an innate 'power instinct'.[131] Instead, he concludes that 'it is the mere instinct of self-preservation which, in the vicious circle described above, leads to competition for ever more power'.[132]

Although subsequent theorists of the security dilemma have not always emphasised the role of pity to the extent that Herz did, there are echoes of it in contemporary discussions of 'strategic empathy'.[133] Much like Rousseauian pity, strategic empathy refers both to an affective capacity and to a cognitive process. Stephen Walt sees strategic empathy as one of neorealism's key lessons.[134] In his view, it is important for overcoming what Lee Ross called 'the fundamental attribution error'. This error describes the human tendency to see the behaviour of others as reflections of the latter's personality or character rather than as a response to the situations they are in, while tending to see our own behaviour as a response to the circumstances we are facing rather than as being

[121] *Aristotle's Poetics*, pp. 75–7; *SD*, 156, 203–4.
[122] Mearsheimer, *The Tragedy of Great Power Politics*, p. 182.
[123] Mearsheimer and Rosato, *How States Think*, p. 101.
[124] Lu, 'Tragedies and International Relations', p. 159.
[125] Booth and Wheeler, *Security Dilemma*, p. 8.
[126] Herz, *Political Realism and Political Idealism*, p. 3.
[127] Herz, *Political Realism and Political Idealism*, p. 6.
[128] Terrance McConnell, 'Moral dilemmas', in Edward N. Zalta and Uri Nodelman, *The Stanford Encyclopedia of Philosophy* (Stanford University, Metaphysics Research Lab, 2024), available at: {https://plato.stanford.edu/archives/spr2024/entries/moral-dilemmas/}.
[129] Ibid.
[130] Herz, *Political Realism and Political Idealism*, pp. 7–8.
[131] Herz, *Political Realism and Political Idealism*, p. 4.
[132] Ibid.
[133] For a review of the literature, see Claire Yorke, 'Is empathy a strategic imperative? A review essay', *Journal of Stategic Studies*, 46:5 (2023), pp. 1082–1102.
[134] Walt, 'The geopolitics of empathy'.

solely a manifestation of 'who we are'.[135] Similarly, our capacity to feel what others are feeling plays a central role in what Booth and Wheeler call the 'security dilemma sensibility'.[136] The security dilemma sensibility refers to an actor's ability to understand the role that fear might play in the attitudes and behaviour of others and the role that one's own actions might play in provoking their fears.

Despite Booth's and Wheeler's call for greater sensibility to the security dilemma, IR theorists tend to focus more on material variables than on emotional ones. Only recently have IR theorists returned to the question of how developing compassion facilitates de-escalation by leading actors to reinterpret their interests and moderate their positions on key issues of contestation.[137] Stephen Walt even suggests that the West's inability to empathise with the Russian perspective on Ukraine may have made it more difficult to avoid escalation.[138] Whatever one thinks of this, the lens of Rousseauian tragedy provides a helpful framework for understanding the psychology underlying neorealist writings on empathy.

Time-consciousness and the security dilemma

By reading neorealism as Rousseauian, the importance of time-consciousness for the security dilemma also becomes clearer. Because Hobbes's natural man is conscious of time, Hobbes concluded that men are born natural enemies.[139] Rousseau's lesson about the dangers of time-consciousness was not lost on Herz, nor has it been completely forgotten since. In fact, it lives on in what contemporary neorealists such as Dale Copeland and Sebastian Rosato call 'the problem of the future'.

Following Rousseau rather than Hobbes, Herz suggests that the acquisition of time-consciousness transforms humans from peaceful brutes into predatory animals. As Herz sees it, there is 'a great divide between death-conscious – dare we say, civilized? – man, and pre-civilized man plus all other living beings. With the development of consciousness of the natural limitations of one's life, this life itself assumes a different character.'[140] Once man can conceive of past, present, and future, 'no longer, as in the case of death- and time-unconscious beings, does the present, lived moment partake of the character of "eternity"; it is merely a link between a remembered past and a future which is already lived through in advance'.[141] Only the advent of language, and with it time-consciousness, brings man 'anxiety as to his neighbors' intentions'.[142] As Rousseau first pointed out, only a time-conscious creature 'can conceive of its own mortality, fear its own death, and thus give rise to the dynamics Hobbes explores'.[143]

Because almost everyone agrees with neorealists that the international system is anarchic, that states seek security, and that great powers possess offensive capabilities, the key to explaining how otherwise defensively motivated people begin to compete for security is uncertainty about the future. Like Herz, Waltz thought that, for units under anarchy, 'the condition of insecurity – at the least, the uncertainty of each about the other's future intentions and actions – works against their

[135]Lee Ross, 'The intuitive psychologist and his shortcomings: Distortions in the attribution process', *Advances in Experimental Social Psychology*, 10 (1977), pp. 173–220.

[136]Booth and Wheeler, *Security Dilemma*, p. 7.

[137]Baker, 'The empathic foundations of security dilemma de-escalation', p. 1252.

[138]Stephen M. Walt, 'The West is sleepwalking into war in Ukraine', *Foreign Policy* (23 February 2022), available at: {http://foreignpolicy.com/2022/02/23/united-states-europe-war-russia-ukraine-sleepwalking/}.

[139]Arthur M. Melzer, *The Natural Goodness of Man: On the System of Rousseau's Thought* (Chicago: University of Chicago Press, 1990), p. 20; Bahar Rumelili, 'Integrating anxiety into International Relations theory: Hobbes, existentialism, and ontological security', *International Theory*, 12:2 (2020), pp. 257–272 (pp. 261–5).

[140]Herz, *Political Realism and Political Idealism*, p. 2.

[141]Herz, *Political Realism and Political Idealism*, p. 2; similarly *SD*, pp. 146, 167.

[142]Herz, *Political Realism and Political Idealism*, p. 3.

[143]Williams, *The Realist Tradition and the Limits of International Relations*, p. 60.

cooperation'.[144] Although Mearsheimer gives uncertainty about intentions a more central role than Waltz did, he contradicts himself by claiming that states cannot know others' intentions and that anarchy forces states to have hostile intentions.[145] Only Sebastian Rosato articulates the intentions assumption in depth and consistently.

For Rosato, states face acute uncertainty about others' current intentions because of problems of access and reliability.[146] Problems of access stem from the fact that usually only a handful of decision-makers know what a state intends to do in terms of foreign policy in the mid to long term, while problems of reliability occur because evidence is frequently equivocal, and because states sometimes have incentives to deceive or may intend to pursue interests in a variety of ways. Future intentions are even less certain than current ones because there is no way to access them first hand, as they do not yet exist, because second-hand information about them faces the same problems of reliability that confront current intentions, and because they are subject to change as conditions unfold.[147]

The result has been described by some neorealists as the 'problem of the future'.[148] Due to the consequences that could follow from misjudging others' intentions, agents in anarchy need to be nearly certain of 'benign intentions in order to forego (*sic*) security competition'.[149] But, because the future does not yet exist, future intentions cannot be known directly and, even if correctly inferred, are subject to change. As a result, agents in an anarchic environment are unlikely to ever have near certainty about others' future intentions.[150] Time-consciousness thus undergirds the tragic account of security competition not only in Rousseau's writings but also in contemporary neorealism.

Conclusion

This article began with the question of whether neorealism implies a conception of human nature. According to the conventional wisdom, neorealism makes no assumptions about human nature. We are now in a position to see why that is incorrect. Not only does neorealism imply a view of human nature, as several scholars have argued,[151] but in fact its conception corresponds most closely to Rousseau's image of human nature in the *Second Discourse*. In seeing neorealist anthropology as Rousseauian, this account diverges from previous readings of neorealism as Hobbesian.[152]

This argument matters because the political philosophies of Hobbes and Rousseau point to opposite policy prescriptions. If Hobbes is right, then policymakers should approach conflicts along the lines of the deterrence model. According to the deterrence model, the frequent presence of greedy units in the system means that, as Jervis once put it, 'power must be met by power'.[153] From this perspective, concessions to greedy states are likely only to empower or embolden them. If Rousseau is right, however, policymakers should approach conflicts along the lines of the spiral model. According to the spiral model, units in anarchy are more fearful than greedy, so concessions and restraint are more likely reassure than to embolden. From this perspective, there is a worry

[144]Kenneth N. Waltz, 'Conflict in world politics', in Steven L. Spiegel and Kenneth N. Waltz (eds), *Conflict in World Politics* (Cambridge: Winthrop Publishers, 1971), pp. 454–474 (p. 460); Waltz, *TIP*, p. 105.

[145]Mearsheimer, *The Tragedy of Great Power Politics*, pp. 31, 34.

[146]Rosato, *Intentions in Great Power Politics*, pp. 20–31.

[147]Rosato, *Intentions in Great Power Politics*, pp. 30–2.

[148]Copeland, 'Rationalist theories of international politics and the problem of the future'; Rosato, *Intentions in Great Power Politics*, pp. 71–2.

[149]Rosato, *Intentions in Great Power Politics*, p. 5.

[150]Rosato, *Intentions in Great Power Politics*, pp. 21–32.

[151]Crawford, 'Human nature and world politics'; Elshtain, 'Woman, the state, and war'; Freyberg-Inan, *What Moves Man*; Freyberg-Inan, 'Rational paranoia and enlightened machismo'.

[152]E.g. Donnelly, 'The discourse of anarchy in IR', p. 417; Freyberg-Inan, *What Moves Man*, p. 61; Polansky, 'Drawing out the Leviathan'.

[153]Robert Jervis, *Perception and Misperception in International Politics* (Princeton, NJ: Princeton University Press, 1976), p. 78.

that the policies recommended by the deterrence model will produce 'a self-reinforcing cycle that can eventually lead to preemptive or accidental war'.[154] This is because, regardless of one's motives, arming is likely to appear threatening to others.[155] The fact that the philosophical anthropologies of Hobbes and Rousseau point to opposed policies, combined with the fact that neorealism retains clout among students of IR and in some foreign policy circles, indicates the significant stakes of describing neorealism as Rousseauian rather than Hobbesian.[156]

Ultimately, the Rousseauian reading of neorealist anthropology brings to light a central moral concern of neorealists, what I call the tragedy of sheep in wolves' clothing. Although we all know the old proverb 'beware of wolves in sheep's clothing', Rousseauian realism turns this admonition on its head. If Rousseau is right, human beings are generally more like sheep than wolves.[157] So, if we correctly understood our human nature, we might realise that our enemies are not innately evil or inhuman. On the contrary, they may simply feel compelled by the pressures of their situation to compete for security, and sometimes even to attack others, just like we do. Unfortunately, by treating our enemies as inhuman or beyond the moral pale, we may just make a bad situation worse.[158] Thus, the Rousseauian warning becomes 'don't forget the sheep beneath the wolves' clothing'. In other words, remember that your rivals are human too, however ferocious they may appear.

Video Abstract. To view the online video abstract, please visit: https://doi.org/10.1017/S0260210525000026.

Acknowledgements. For their insightful comments and apt critiques, I am grateful to Adriana Albanus, Kevin Bustamante, Susan Collins, Michael Desch, Hannes Furlan, Mary Keys, Joseph Parent, Emma Planinc, Sebastian Rosato, three anonymous reviewers, the editors of RIS, and the participants of the Notre Dame political theory research workshop: Evelyn Behling, Richard Cassleman, Luke Foster, Seaver Holter, Shasta Kaul, Samuel Piccolo, Linus Recht, Emma Schmidt, Benjamin Sehnert, and Max Smith. I am also grateful to the Institute for Humane Studies for their support (grant no. IHS016880).

Competing interests. The author declares none.

Joseph Clarkson is a postdoctoral teaching scholar at the University of Notre Dame specialising in political theory and International Relations. His dissertation examined the sources of political unity in the Weimar-era state theories of Carl Schmitt and Hermann Heller. More broadly, his research focuses on accounts of state formation and sovereignty as well as the causes of war in the history of political thought. Before pursuing his PhD in Political Science at the University of Notre Dame, Joseph studied Philosophy, Classics, and German through the University Scholars Program at Baylor University.

[154] Andrew Kydd, 'Arms races and arms control: Modeling the hawk perspective', *American Journal of Political Science*, 44:2 (2000), pp. 228–244 (p. 228).

[155] Kenneth N. Waltz, 'Reason, will and weapons', *Political Science Quarterly*, 74:3 (1959), pp. 412–419 (p. 413).

[156] I am grateful to an anonymous reviewer for suggesting this point.

[157] The language of 'sheep' and 'wolves' is borrowed Joseph Parent and Joshua Baron's response to Andrew Kydd. Kydd had argued, 'while a group of sheep can get along fine with each other if they were fully convinced that all the animals in the flock were sheep, some may be wolves in sheep's clothing'. Andrew Kydd, 'Sheep in sheep's clothing: Why security seekers do not fight each other', *Security Studies*, 7:1 (1997), pp. 114–155 (p. 116). The neorealists Parent and Baron suggested that Kydd had been too hasty in dismissing the possibility of savage competition among purely benign actors, complaining that Kydd 'out-of-hand dismisses the problem of savage competition between sheep should they be comfortable that no wolf was in their midst'. Parent and Baron, 'Elder abuse', p. 207 n. 39. Rather than 'sheep' and 'wolves', Rousseau uses the image of Glaucus to make this point. Rousseau writes that 'human soul adulterated in the lap of society' is 'like the statue of Glaucus which time, sea, and storms had so disfigured that it resembled less a God than a ferocious beast' (*SD*, 126). Glaucus first appears in philosophy in Plato's *Republic* 10.611d. Grace Roosevelt suggests that Rousseau's use of the Platonic image is 'aimed at the "beastly" qualities that Hobbes posited as natural in human beings'. Roosevelt, *Reading Rousseau in the Nuclear Age*, p. 28.

[158] Empirical evidence supports this hypothesis: Baker, 'The empathic foundations of security dilemma de-escalation', pp. 1260–3; Abigail S. Post, 'Words matter: The effect of moral language on international bargaining', *International Security*, 48:1 (2023), pp. 125–165.

Cite this article: Joseph Clarkson, 'The Rousseauian roots of neorealism', *Review of International Studies* (2025), pp. 1–19. https://doi.org/10.1017/S0260210525000026