B.J.Pol.S. **41,** 191–210 *Copyright © Cambridge University Press, 2010*
doi:10.1017/S0007123410000372
First published online 10 November 2010

Review Article: 'Major Combat Operations Have Ended'? Arguing about Rational Choice

ANDREW HINDMOOR*

Arguing about rational choice theory remains a popular pastime. Following the publication of Green and Shapiro's *Pathologies of Rational Choice Theory,* a backlash against the use of rational choice theory within political science gained momentum. This article shows how, since 1994, sceptics have refined and extended the critique of rational choice and how practitioners have defended their approach, and a more general argument has emerged. In the 1990s, attitudes towards rational choice theory constituted a fundamental fault-line within the discipline, but changes to the way in which rational choice is practised and defended, together with some broader changes in the social sciences, have created more areas of common ground and taken some of the urgency out of this debate.

In an article published in 1993 to celebrate the twenty-fifth anniversary of the founding of the journal *Public Choice*, Dennis Mueller predicted that 'rational choice and political science will be indistinguishable in another generation': that rational choice 'will be a field within economics, and will encompass all of political science'.[1] This was certainly a provocative claim but not one which could be dismissed out of hand. In 1980, around 20 per cent of the articles published in the *American Political Science Review* used rational choice (RC) theory. By 1990, more than 40 per cent did so.[2] Yet at this moment, a backlash against the use of RC theory within political science was gaining momentum. Green and Shapiro's argument that successful applications of RC theory were few and far between and that much work simply restated 'existing knowledge in rational choice terminology' proved an early focal point.[3] Also of significance was the gradual accumulation of experimental evidence showing that, contrary to the assumptions routinely made within RC, people are not perfectly rational or exclusively self-interested. In the United States, the 'perestroika' movement provided critics with a wider platform from which to press their concerns, and by 1996 Shepsle was describing RC theory as experiencing a 'mid-life crisis'.[4]

Shortly before the publication of *Pathologies*, Grafstein likened the debate about RC theory to that on abortion rights. What he meant was that RC is a subject about which any

* School of Political Science and International Studies, University of Queensland (email: andrew.hindmoor@uq.edu.au). The author thanks Stephen Bell, Keith Dowding, Albert Weale and two anonymous *Journal* referees for their generous comments on an earlier draft of this article.

[1] Dennis Mueller, 'The Future of Public Choice', *Public Choice*, 77 (1993), 145–50.

[2] Donald Green and Ian Shapiro, *Pathologies of Rational Choice Theory* (New Haven, Conn.: Yale University Press, 1994), p. 3.

[3] Green and Shapiro, *Pathologies of Rational Choice Theory*, p. 6.

[4] For an openly partisan account of the aims and influence of the perestroika movement, see: Kurt Jacobsen, 'Perestroika in American Political Science', *Post-Autistic Economics Review*, 32 (2005), Article 6; Kenneth Shepsle, 'Statistical Political Philosophy and Positive Political Theory', in Jeffrey Friedman, ed., *The Rational Choice Controversy* (New Haven, Conn.: Yale University Press, 2006), pp. 213–22, at p. 213.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

'self-respecting' social scientist ought to have an opinion.[5] But as a means of characterizing the debate in the years which were to follow, the analogy works in a further respect. For discussion about RC prompted passionate and sometimes intemperate exchanges exposing fundamental and apparently irreconcilable differences between participants. These included ontological differences concerning the degree to which individuals are compromised 'on high' by aggregate or structural factors;[6] epistemological issues relating to the merits of an instrumentalist theory of knowledge; methodological issues about formalism and deductive models; and political issues relating to the efficacy of markets and the merits of government intervention.

To focus for a moment upon the last of these issues, many have objected to RC because it was 'inherently' right-wing and had provided the New Right in Britain and United States with intellectual justification and a steady stream of policy ideas.[7] As Christiano suggests, there is a certain irony in any attributing policy influence to RC.[8] If politicians and other actors really are zealously self-interested, then RC arguments about which policies they *ought* to pursue should have no motivational force. If politicians are going to succumb to rent-seeking pressures because it is in their electoral interests to do so, then the publication of a further paper in *Public Choice* estimating the costs of rent-seeking should not make any difference to their behaviour. Yet there can be little doubt that theorists like Buchanan and Tullock found in RC a convenient tool with which to develop a theory of state failure.[9]

Arguing about RC theory remains a 'cause célèbre' within social and political theory.[10] Over the last fifteen or so years, the critique of RC theory has been sharpened and extended, whilst practitioners have found new ways of defending the use of their method. A part of what I want to do in this article is review these contributions from my own position of engaged agnosticism. But I want to do something more: to argue that some of the urgency has gone out of this debate and that attitudes toward RC theory no longer constitute such a fundamental fault-line within the discipline.

It is not that political scientists have simply become indifferent to RC theory. The rapturous reception accorded to Mackie's attempted demolition of Riker's work in *Democracy Defended* and, more recently, to Taylor's *Rationality and the Ideology of Disconnection* and Marglin's *The Dismal Science: How Thinking like an Economist Undermines Community* shows that full-blooded assaults still hold an appeal and that the debate about RC retains a political edge.[11] Neither is this one-way traffic. RC theorists remain more than ready to accuse their critics of

[5] Robert Grafstein, 'Rational Choice Inside and Out', *Journal of Politics*, 54 (1992), 259–68, p. 259.

[6] Phillip Pettit, *The Common Mind: An Essay on Psychology, Society, and Politics* (Oxford: Oxford University Press, 1996), p. 111.

[7] Desmond King, *The New Right* (Basingstoke, Hants.: Macmillan, 1987), p. 104. Also, see Peter Self, *Government by the Market* (Basingstoke, Hants.: Macmillan, 1993), and Hugh Stretton and Lionel Orchard, *Public Goods, Public Enterprise, Public Choice*. (Basingstoke, Hants.: Macmillan, 1994).

[8] Thomas Christiano, 'Is Normative Rational Choice Theory Self-Defeating?' *Ethics*, 115 (2004), 122–41.

[9] See Andrew Hindmoor, 'Public Choice Theory', in Colin Hay, Michael Lister and David Marsh, eds, *The State: Theories and Issues* (Basingstoke, Hants.: Palgrave Macmillan, 2005), pp. 79–97. Keith Dowding and I have previously argued that it is the optional use of the Pareto criterion which gives much RC scholarship its distinctive political flavour (Keith Dowding and Andrew Hindmoor, 'The Usual Suspects: Rational Choice Theory, Socialism and Political Theory', *New Political Economy*, 2 (1998), 451–66).

[10] Frank Lovett, 'Rational Choice Theory and Explanation', *Rationality and Society*, 18 (2006), 237–72, p. 237.

[11] Gerry Mackie, *Democracy Defended* (Cambridge: Cambridge University Press, 2003); Michael Taylor, *Rationality and the Ideology of Disconnection* (Cambridge: Cambridge University Press, 2006);

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

failing to understand even the most basic issues in the philosophy of science.[12] So what has changed? I want to suggest that changes to the way in which RC theory is both practised and defended, which in part reflect the weight of earlier criticisms, imply a more modest and, to its critics, less objectionable role for RC theory. In addition, I suggest that, over roughly the same period, broader changes within the discipline have created more areas of common ground between RC theorists and their critics. I do not want to exaggerate the extent of the rapprochement. Important differences remain and RC theory remains controversial. In suggesting that 'major combat operations have ended', I am being intentionally ironic. George W. Bush's equivalent assurance on the flight deck of the USS Lincoln in May 2003 proved somewhat wide of the mark, but nevertheless I want to suggest that, whilst the RC debate remains significant, it is no longer as divisive.

I start in a moment by outlining Green and Shapiro's argument. I then suggest that their critique may have encouraged RC theorists to engage in more careful empirical testing, thus ironically reducing the significance of their argument, but this may be a side issue. The fundamental objection to RC theory centres upon the assumptions it uses, rather than the empirical results it achieves. The 'hard core' of the RC method is said to consist of two assumptions: decision-theoretic rationality and self-interest.[13] Mueller, for example, defines RC as the application of the methods of economics to the study of politics and writes that the basic behavioural postulate is that man is an egoistic, rational utility-maximizer.[14] Critics argue that these assumptions are woefully inadequate and that it is because they are inadequate that RC's empirical record is so poor. There is nothing new in this argument. Barry Hindess rehearsed many of the best lines in *Choice, Rationality and Social Theory*.[15] In this opening section I do, however, suggest that broader developments within political science and economics have given such arguments greater momentum. The story of how the RC debate has developed is, in significant part, the story of how RC theorists have responded to this argument about assumptions. In order to track this debate, I break the rest of the review into three further sections.

In the first section, I discuss two relatively straightforward arguments. The first, partial universalism, suggests that RC explanations will prove effective when people *are* rational and self-interested but accepts that these assumptions will not always hold. The second is that RC theorists need not and frequently do not assume self-interest and decision-theoretic rationality. Judged in terms of the explanatory leverage they credit to RC, these are two very different arguments. I put them together in this section because they tackle head-on the argument about assumptions.

(F'note continued)
Stephen Marglin, *The Dismal Science: How Thinking Like an Economist Undermines Community* (Cambridge, Mass.: Harvard University Press, 2008).

[12] Daniel Diermeier, 'Rational Choice and the Role of Theory in Political Science', in Jeffrey Friedman, ed., *The Rational Choice Controversy* (New Haven, Conn.: Yale University Press, 2006), pp. 59–60. For a more recent exchange, see Keith Dowding, 'A Pandemonium of Confusions: Kay and Marsh on Tiebout', *New Political Economy*, 13 (2008), 335–48.

[13] Pettit, *The Common* Mind, p. 265. Also, see Uskali Mäki, 'Symposium on Explanations and Social Ontology: Explanatory Ecumenism and Economic Imperialism', *Economics and Philosophy*, 18 (2002), 235–57; and Viktor Vanberg, 'On the Economics of Moral Preferences', *American Journal of Economics and Sociology*, 67 (2008), 605–28, p. 605.

[14] Dennis Mueller, *Public Choice III* (Cambridge: Cambridge University Press, 2003), p. 1. Also, see Mäki, 'Symposium on Explanations and Social Ontology'.

[15] Barry Hindess, *Choice, Rationality and Social Theory* (London: Unwin Hyman, 1988).

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

In the following section, I identify and assess four arguments which seek to side-step objections to the assumptions of self-interest and decision-theoretic rationality, by offering alternative accounts of what it is that RC seeks to explain and how it explains it. The 'received interpretation' here is that RC is a form of intentional explanation which explains people's actions in terms of their conscious mental states and political outcomes and events in terms of those actions and interactions.[16] Lovett describes such an understanding of what RC is explaining as 'ubiquitous'.[17] The four alternative arguments I consider here are: (i) that RC theory operates with a revealed preference view of utility and does not attempt to provide psychologically realistic explanations of why people act in certain ways; (ii) that RC aims to explain the resilience of patterns of behaviour rather than the reasons for that behaviour; (iii) that RC is a heuristic aid to thought; and (iv) that RC explanations explain via unification.

In a recent review article in this journal, Gerring describes a 'widespread turn' towards mechanism-centred explanation across the social sciences.[18] RC theory has not proven immune to this trend. Chong argues that RC theory ought to be understood as a method for 'exploring new causal mechanisms', whilst Macdonald points to the mechanistic base to much RC theory.[19] This is not necessarily inconsistent with the claim that RC is a form of intentional explanation. Intentions might be viewed as one possible explanatory mechanism. Pettit says that RC explanations 'appeal to aggregative mechanisms … as well as to motivational ones'.[20] Mechanisms cause things to happen in particular ways. But in an ontologically complex world, mechanisms can generate counteracting effects.[21] A key question, therefore, concerns the way in which the effects of the operation of individual mechanisms can be analysed. One answer to this question involves the application of the 'method of isolation'. I discuss this method and the relevance it has to arguments about the use of unrealistic assumptions in the final section.

THE CASE AGAINST RATIONAL CHOICE

Green and Shapiro's *Pathologies* had a tremendous impact because it claimed to be judging RC theory by the standards its practitioners had set for themselves. Traditionally, one way in which RC theorists have been able to side-step arguments about inaccurate assumptions is by invoking Friedman's instrumentalist argument that theories should be judged in terms of their predictive success rather than the realism of their assumptions.[22]

[16] Debra Satz and John Ferejohn, 'Rational Choice and Social Theory', *Journal of Philosophy*, 91 (1994), 71–87, p. 71.

[17] Lovett, 'Rational Choice Theory and Explanation', p. 268.

[18] John Gerring, 'The Mechanismic Worldview: Thinking Inside the Box', *British Journal of Political Science*, 38 (2007), 161–79, p. 161.

[19] Dennis Chong, 'Rational Choice Theory's Mysterious Rivals', in Friedman, ed., *The Rational Choice Controversy*, pp. 37–58; Paul MacDonald, 'Useful Fiction or Miracle Maker: The Competing Epistemological Foundations of Rational Choice Theory', *American Political Science Review*, 97 (2003), 551–65. Also, see Lovett, 'Rational Choice Theory and Explanation', p. 253.

[20] Pettit, *The Common Mind*, p. 269. Elster, by contrast, draws a clear distinction between intentional and causal explanations (see Jon Elster, 'Introduction', in Elster, ed., *Rational Choice* (Oxford: Blackwells, 1986), and Jon Elster, 'The Nature and Scope of Rational-Choice Explanation', in Ernest LePore and Brian McLaughlin, eds, *Actions and Events: Perspectives on Donald Davidson* (Oxford: Blackwells, 1985).

[21] Tony Lawson, *Economics and Reality* (London: Routledge, 1999).

[22] Milton Friedman, 'The Methodology of Positive Economics', in *Essays on Positive Economics* (Chicago: University of Chicago Press, 1953). For a critical discussion of what Friedman actually said (as

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

Green and Shapiro seemed to demonstrate that this predictive record was nevertheless quite poor. Their argument can be summarized as follows: (i) The value of any theory ought to be measured in terms of its empirical achievements; (ii) in the case of RC theory, these achievements are extremely limited; (iii) applications of RC theory exhibit a number of methodological pathologies, including the search for confirming rather than falsifying evidence, post hoc theory development and arbitrary domain restriction; and (iv) the prevalence of these pathologies can be accounted for in terms of the theory's method-driven rather than puzzle-driven approach.

One of the initial lines of defence here was attack. In a series of essays, subsequently published as *The Rational Choice Controversy*, a number of practitioners argued that Green and Shapiro had fallen foul of their own methodological standards and unfairly searched for areas of work, such as the paradox of voting, where RC theory was already known to be empirically weak, and ignored areas of work, like legislative politics and agenda-setting, where it was strong.[23] Shepsle also argued that, whatever its limitations, there was no credible rival to RC – his first law of wing-walking being that you should not let go until you have something else to hold on to.[24] To which Green and Shapiro rather bitingly noted that this law 'might be easier to take seriously if one could develop a degree of confidence that the aircraft in question were in fact airborne'.[25]

If the initial tone of the debate was quite hostile, it may nevertheless be that, in the longer term, *Pathologies* encouraged RC theorists to invest more time in empirical testing. Sanchez-Cuenca argues that the 'last ten years or so' have 'witnessed an impressive integration of formal modelling and empirical testing': with Boix's *Democracy and Redistribution*, Epstein and O'Halloran's *Delegating Powers*, Huber and Shipan's *Deliberate Discretion* and Tsebelis's *Veto Players* being singled out for praise.[26] In a recent paper with Hampsher-Monk, I have calculated that, in accordance with Sanchez-Cuenca's contention, the proportion of articles published within the *American Political Science Review* and the *American Journal of Political Science* employing RC theory and including some form of empirical testing has increased. Between 1984 and 1993, only 52 per cent of the RC articles in these journals included empirical tests. Between 1994, when *Pathologies* was published, and 2006 this proportion increased to 70 per cent.[27] Critics might still argue that too much

(F'note continued)
opposed to what he is usually understood as saying), see Uskali Mäki, '"The Methodology of Positive Economics" Does Not Give Us the Methodology of Positive Economics', *Journal of Economic Methodology*, 10 (2003), 495–505.

[23] Shepsle, 'Positive Political Philosophy'; Chong, 'Mysterious Rivals'; Morris Fiorina, 'Rational Choice, Empirical Contributions and the Scientific Enterprise' (pp. 85–94), Norman Schofield, 'Rational Choice and Political Economy' (pp. 189–212) and Susanne Lohmann, 'The Poverty of Green and Shapiro' (pp. 127–54) all in Friedman, ed., *The Rational Choice Controversy*.

[24] Shepsle, 'Statistical Political Philosophy', p. 217.

[25] Donald Green and Ian Shapiro, 'Pathologies Revisited: Reflections on our Critics', in Friedman, ed, *The Rational Choice Controversy*, pp. 235–76, at p. 256.

[26] Ignacio Sánchez-Cuenca, 'A Preference for Selfish Preferences: The Problem of Motivations in Rational Choice Political Science', *Philosophy of the Social Sciences*, 38 (2008), 361–78, p. 363; Charles Boix, *Democracy and Redistribution* (Cambridge: Cambridge University Press, 2003); David Epstein and Sharyn O'Halloran, *Delegating Powers* (Cambridge: Cambridge University Press, 1999); John Huber and Charles Shipan, *Deliberate Discretion?* (Cambridge: Cambridge University Press, 2002); and George Tsebelis, *Veto Players: How Political Institutions Work* (Princeton, N.J.: Princeton University Press, 2002).

[27] Iain Hampsher-Monk and Andrew Hindmoor, 'Rational Choice and Interpretive Evidence: Caught Between a Rock and a Hard Place?' *Political Studies*, 58 (2009), 47–65.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

RC theory remains untested, or that, when tested, it only generates 'banal' results.[28] But such reservations should not be allowed to detract from the extent of the changes which have occurred.

The renewed emphasis upon empirical testing may be welcome but it is unlikely to lead to any mass conversions to the RC faith. The key issue for critics has always been about the assumptions RC makes. The argument here is straightforward. RC experimental evidence has conclusively shown that people have incomplete and sometimes intransitive preferences; that they are vulnerable to framing, anchoring and priming effects; that they miscalculate risk; that they excessively discount the future; and that they frequently display a willingness to co-operate with others that is difficult to reconcile with the assumption of self-interest.[29] In short, the criticism here is that far from being perfectly rational and zealously self-interested, people are boundedly rational and have bounded will-power and self-interest.[30]

I do not have the space available here to offer a thorough review of this work. But, as a token example, consider research on the ultimatum game. Here, two players must decide how to divide a sum of money which has been given to them. The first makes a proposal about how to divide the money and the second can either accept or reject this offer. If the second player rejects it, then neither player receives anything. If the second player accepts the proposed division, the money is split accordingly. The game is only played once and is usually played anonymously. RC theory predicts that the first player will propose giving themselves nearly all of the money, leaving the second player with a choice between getting almost nothing and getting nothing. But when the ultimatum game was first run with cash-motivated subjects, Guth observed that the modal proposal was to split the money equally.[31] This result has since been replicated hundreds of times, with one review article reporting that the average offer made to the second person is between 30 and 40 per cent: a finding a number of critics refer to in their commentaries upon RC.[32]

The status of this experimental evidence is contested. Economists have argued that when choices are relatively simple, the stakes are relatively high and situations are repeated, and participants in experiments routinely act in ways that are consistent with

[28] Green and Shapiro, *Pathologies of Rational Choice*, p. 6.

[29] Summaries of and guides to this research can be found in: Wolfgang Pesendorfer, 'Behavioral Economics Comes of Age: A Review Essay on Advances in Behavioral Economics', *Journal of Economic Literature*, 44 (2006), 712–21; Elinor Ostrom, 'A Behavioral Approach to the Rational Choice Theory of Collective Action', Presidential Address, American Political Science Association, *American Political Science Review*, 92 (1990), 1–22; Norman Frohlich and Joe Oppenheimer, 'Skating on Thin Ice: Cracks in the Public Choice Foundation', *Journal of Theoretical Politics*, 18 (2006), 235–66; Bryan Caplan, *The Myth of the Rational Voter: Why Democracies Choose Bad Policies* (Princeton, N.J.: Princeton University Press, 2007); Dan Ariely, *Predictably Irrational: The Hidden Forces that Shape our Decisions* (New York: Harper Collins, 2008); and Daniel Khahneman and Amos Tversky, *Choices, Values and Frames* (Cambridge: Cambridge University Press, 2000).

[30] Sendhil Mullainathan and Richard Thaler, 'Behavioral Economics', in Neil Smelser and Paul Bates, eds, *International Encyclopedia of the Social and Behavioral Sciences* (New York: Elsevier, 2001).

[31] Werner Guth, Rolf Schmittberger and Bernd Schwarze, 'An Experimental Analysis of Ultimatum Bargaining', *Journal of Economic Behavior and Organization*, 3 (1982), 367–88.

[32] Colin Camerer and Richard Thaler, 'Anomalies, Ultimatums, Dictators and Manners', *Journal of Economic Perspectives*, 9 (1995), 209–19; Jon Elster, 'Rational Choice History: A Case of Excessive Ambition', *American Political Science Review*, 94 (2000), 685–95, p. 692; Raymond Boudon, 'Beyond Rational Choice Theory', *Annual Review of Sociology*, 29 (2003), 1–21, p. 7. Also see the discussion in Vanberg, 'On the Economics of Moral Preferences', pp. 606–7.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

the standard assumptions of RC theory.[33] Others have expressed more general concerns about drawing strong conclusions from evidence that is often limited. Santos is typical in suggesting that experiments which can tell us something about how people behave in one carefully circumscribed and artificially created situation cannot necessarily tell us a great deal about how people generally reason and make decisions.[34]

Yet the accumulated experimental evidence about people's behaviour is difficult to dismiss out of hand or to regard as being caused by individual idiosyncrasies which will cancel each other out when the behaviour of enough people is examined. Furthermore, the rise and rise of historical institutionalism provides political scientists with a methodological alternative to RC which makes very different assumptions about agency. Historical institutionalism tends to a view that agents are embedded within, and their agency constrained by, institutions; that agents' preferences are formed endogenously within institutions (rather than being exogenously fixed); and that embedded agents lack the cognitive resources always to identify the utility-maximizing strategy. This is why, within the new institutionalism, clear distinctions have traditionally been drawn between RC and historical institutionalism.[35]

The story here is, however, not a straightforward one. In recent years, there is some evidence of what Thelen calls 'border crossing' and a narrowing of the differences between these approaches.[36] Three examples will suffice. Scharpf accepts that institutions can shape norms and behaviour and that, as a result, 'we will not gain much mileage from assumptions of self-interested action per se'.[37] His actor-centred institutionalism nevertheless seeks to show how game theory can be used to explain the results of strategic interaction and so reconcile approaches 'conventionally treated as being mutually exclusive'.[38] In a study of the development of vocational training regimes in Germany, Britain, the United States and Japan, Thelen argues that the question 'Why do institutions take the form they do?' is best answered using historical rather than RC institutionalism.[39] Yet from this starting point, she nevertheless shows how the evolution of training regimes can be understood as the product of, amongst other things, self-interested bargaining between unions, employers and the state. Her general argument, that institutional reproduction and institutional change have to be closely studied together and are in some important ways quite closely linked, is presented as a challenge to existing RC *and* historical institutionalist literature. Finally, Rothstein argues that the problem of explaining how some countries escape from 'social traps' in which efficient co-operation for common purposes can come about only if people trust that most other people will also choose to co-operate 'cannot be handled solely by either the rationalist or the culturalist approach'.[40] Within his overall

[33] Ken Binmore, 'Why Experiment in Economics?' *Economic Journal*, 109 (1999), F16–F24.

[34] Ana C. Santos, 'Behavioral Experiments: How and What Can We Learn about Human Behaviour?' *Journal of Economic Methodology*, 16 (2009), 71–88. Also, see Nikos Siakantaris, 'Experimental Economics under the Microscope', *Cambridge Journal of Economics*, 24 (2000), 267–81.

[35] Most clearly in Guy Peters, *Institutional Theory in Political Science: The New Institutionalism* (London: Pinter, 1999).

[36] Kathleen Thelen, 'Historical Institutionalism in Comparative Politics', *Annual Review of Political Science*, 2 (1999), 369–404.

[37] Fritz Scharpf, *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research* (Boulder, Colo.: Westview, 1997), p. 41.

[38] Scharpf, *Games Real Actors Play*, p. 36.

[39] Kathleen Thelen, *How Institutions Evolve: The Political Economy of Skills in Germany, Britain, the United States and Japan* (Cambridge: Cambridge University Press, 2004), p. 26.

[40] Bo Rothstein, *Social Traps and the Problem of Trust* (Cambridge: Cambridge University Press, 2005), p. 29.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

argument, an important role is played by the notion of a 'collective memory', which can be manipulated by strategic actors in order to enhance levels of overall levels of trust but 'cannot be reduced to simple utility maximization'.[41]

If the boundaries between historical institutionalism and RC theory have become blurred, those between RC and a more recently minted constructivist institutionalism remain frozen. Constructivist institutionalism, as it has been expounded by, principally, Hay, Blyth and Schmidt, holds that *both* RC and historical institutionalism are 'overly structuralist' and do not grant 'purposeful actors a proper role'.[42] Indeed, it is the belief that historical institutionalism has ceded too much ground to RC, which appears to have provided a part of the motive for the development of a constructivist alternative.[43] Central to this position is a conviction that it is the 'ideas that actors hold about the context in which they find themselves rather than the context itself which informs the way in which actors behave'.[44] This focus upon ideas distinguishes constructivist institutionalism not only from historical institutionalism – where institutions are given explanatory primacy – but also from RC – where interests derived from an actor's structural location are the prime mover.

Developments within economics as well as political science have conspired to emphasize the limitations of RC assumptions and the availability of alternative approaches. We might start here by noting that there is a measure of nonchalance in Mueller's earlier definition of rational choice as the application of the methods of economics to politics, in so far as he is actually equating economics with the methods of neo-classical economics. Such has been the grip of the neo-classical approach upon economics that this difference might once have seemed slight. This is, however, no longer the case. Davis argues that 'there is now considerable evidence that economics is undergoing a significant change in which a collection of new research programs all at odds in important respects with standard neoclassical economics is increasingly dominating the economics research frontier'.[45]

As examples of this new research, Davis cites game theory, evolutionary economics, behavioural economics, experimental economics and neuro-economics. With the notable exception of game theory, which has formed a staple of RC analysis, these research

[41] Rothstein, *Social Traps*, p. 160.

[42] See J. Olson, 'Change and Continuity: An Institutional Approach to Institutions and Democratic Government', *European Political Science Review*, 1 (2009), 3–32, p. 3; Colin Hay 'Constructivist Institutionalism', in Rod Rhodes, Sarah Binder and Bert Rockman, eds, *The Oxford Handbook of Political Institutions* (Oxford: Oxford University Press, 2007), pp. 56–74; Colin Hay, 'Constructive Institutionalism … Or, Why Ideas into Interests Don't Go', in Daniel Beland and Robert Cox, eds, *Ideas and Politics in Social Science Research* (Oxford: Oxford University Press, forthcoming); Mark Blyth, 'Any More Bright Ideas? The Ideational Turn in Comparative Political Economy', *Comparative Politics*, 29 (1997), 229–50; Mark Blyth, *Great Transformations: Economic Ideas and Institutional Change in the Twentieth Century* (Cambridge: Cambridge University Press, 2002); and Vivien Schmidt, 'Discursive Institutionalism: The Explanatory Power of Ideas and Discourse', *Annual Review of Political Science*, 11 (2008), 303–26.

[43] Vivien Schmidt, 'From Historical Institutionalism to Discursive Institutionalism: Explaining Change in Comparative Political Economy' (paper given at the annual meeting of the American Political Science Association, Boston, 2008).

[44] Colin Hay and Ben Rosamond, 'Globalisation, European Integration and the Discursive Construction of Economic Imperatives', *Journal of European Public Policy*, 9 (2002), 147–67, p. 147. The whole constructivist institutionalist paradigm is subject to a powerful critique by Stephen Bell, 'Do We Really Need Constructivist Institutionalism to Explain Institutional Change? In Defence of an Agent-Centred Historical Institutionalism', available from the author at Stephen.Bell@uq.edu.au.

[45] John Davis, 'The Turn in Economics and the Turn in Economic Methodology', *Journal of Economic Methodology*, 14 (2007), 275–90, p. 275.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

programmes, all originating in disciplines outside economics, constitute a challenge not only to neo-classical economics but to its political science offspring.

This is perhaps most obviously the case with behaviourism. Within economics, behaviourism has acquired a considerable cachet.[46] Within political science, behaviouralist arguments about bounded rationality, bounded willpower and bounded self-interest once provided the foundations for research on voting and party identification and incremental decision making within organizations.[47] Once sidelined by RC, evidence for a resurgence in behaviourist-inspired research can be seen in Baumgartner and Jones's work on agenda setting and punctuated equilibrium, Sunstein and Thaler's work on 'nudging' and the architecture of choice, and Jonathan Bendor's work on turnout, party competition and bureaucratic decision making.[48] Behaviourism, like historical institutionalism, now provides a credible rival to RC theory.

Experimental economics does not constitute a rival approach to RC in the same way as behaviourism. Indeed, experiments can be used to test RC predictions. But experiments of the sort devised by Donald Green in the United States and by Peter John in Britain need not make any particular assumptions about how people reason and so provide an inductive alternative to a deductive RC method.[49] Neuro-economics poses a different kind of challenge by bringing into question something which has previously been seen as a key source of RC theory's advantage: that 'rational action is its own explanation'.[50] Neuro-economics is a reductionist programme which promises to explain rational action by identifying the neural mechanisms by virtue of which people assess rewards, calculate future consequences, compare options, negotiate deals, perceive threats and modify

[46] Pesendorfer, 'Behavioral Economics Comes of Age'.

[47] The classic references for which remain: Angus Campbell, Philip Converse, Warren Miller and Donald Stokes, *The American Voter* (Chicago: University of Chicago Press, 1960); and Charles Lindblom, 'The Science of Muddling Through', *Public Administrative Review*, 19 (1959), 74–88. Indeed, political science can claim partial ownership here as Herbert Simon's undergraduate and doctoral degree were both in political science (Jonathan Bendor, 'Herbert Simon: Political Scientist', *Annual Review of Political Science*, 6 (2003), 433–71).

[48] Frank Baumgartner and Bryan Jones, *The Politics of Attention: How Government Prioritizes Problems* (Chicago: University of Chicago Press, 2005); and *Agendas and Instability in American Politics* (Chicago: University of Chicago Press, 1993); Bryan Jones, *Politics and the Architecture of Choice* (Chicago: University of Chicago Press, 2001); Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions About Health, Wealth and Happiness* (New Haven, Conn.: Yale University Press, 2008); and Jonathan Bendor, Daniel Diermeier and Michael Ting, 'A Behavioral Model of Turnout', *American Political Science Review*, 97 (2003), 261–80; Jonathan Bendor, Dilip Mookherjee and Debraj Rav, 'Satisficing and Selection in Electoral Competition', *Quarterly Journal of Political Science*, 1 (2006), 171–200; Jonathan Bendor and Sunil Kumar, 'The Perfect is the Enemy of the Best: Adaptive versus Optimal Organizational Reliability', *Journal of Theoretical Politics*, 17 (2005), 5–39.

[49] Donald Green, Eline de Rooij and Alan Gerber, 'Field Experiments on Political Behavior and Collective Action', *Annual Review of Political Science*, 12 (2009), 385–95; and Peter John and Tessa Brannan, 'How Different Are Telephoning and Canvassing? A 'Get out the Vote' Field Experiment in the British 2005 General Election', *British Journal of Political Science*, 38 (2008), 565–74, and references therein.

[50] Martin Hollis, *Models of Man: Philosophical Thoughts on Social Action* (Cambridge: Cambridge University Press, 1977), p. 26. In 'Beyond Rational Choice', Raymond Boudon makes a great deal of the argument that rational choice is its own explanation, approvingly citing Coleman to the effect that the 'rational actions of individuals have a unique attractiveness as the basis for social theory. If an institution or a social process can be accounted for in terms of the rational actions of individuals, then and only then can we say that it has been "explained" … and need ask no more questions' (James Coleman, *Individual Interests and Collective Action: Selected Essays* (Cambridge: Cambridge University Press, 1986), p. 1).

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

their preferences.[51] Neuro-economics has not, as yet, generated significant applications within political science. Still, it is not hard to see how existing work within economics might be easily extended to political decision making.

DEFENDING RATIONAL CHOICE ASSUMPTIONS

In the introduction to *Pathologies,* Green and Shapiro state that they do not object to the use of RC theory *per se*; only to the attempt to use it to explain any and every political phenomenon. Confronted with experimental evidence about how people behave, one alternative argument, therefore, available to RC theorists is to maintain that explanations which assume decision-theoretic rationality and self-interest will prove effective when people *are* rational and self-interested, even if there are many cases in which the assumptions will not prove valid and in which RC explanations consequently will be inappropriate. A 'partial universalism' of this sort carries the qualified blessing of Green and Shapiro. They suggest that RC models are likely to prove effective in those situations where: (i) the stakes are high and actors are self-conscious optimizers; (ii) preferences are well-ordered and relatively fixed; (iii) actors are presented with a clear range of options and little opportunity for strategic opportunity; (iv) strategic uncertainty is minimal; and (v) actors have the opportunity to learn from feedback.[52] Hay complains that this kind of partial universalism implies a 'profound ontological unevenness and inconsistency – some parts of the social universe are simple, some complex'.[53] But by Hay's usual standards, this is fairly unconvincing. Why cannot certain parts of the world be simpler than others?

Partial universalism concedes considerable ground to RC's critics in so far as it would seem to put a large part of politics off-limits. If, for example, Brennan is right about the logic of choice and the prevalence of expressive rationality in the electoral arena, then it may be that RC theory cannot tell us anything useful about voting.[54] George Orwell once remarked, the quickest way of ending a war is to lose it and partial universalism looks like a partial surrender. But, in so far as even the most intense of critics is likely to concede that political actors are often self-interested and capable of making informed choices – if only because in some situations there is a norm encouraging such behaviour, it also creates the grounds for an academic division of labour and the possibility of a potentially peaceful coexistence.

There is, however, another option available to practitioners. Confronted with evidence that people are not really infallible, self-interested decision-makers, RC theorists can argue that their position has been unfairly caricatured and that they had never said that they were. In her 1997 Presidential Address to the American Political Science Association, Elinor Ostrom reviewed the evidence relating to people's decision-making capacities and motives. She concluded not that RC theory ought to be abandoned but instead that it ought to be revised so as to allow for the possibility, within its basic assumptions, of error, learning and the existence of ingrained reciprocity norms.[55] Over the last decade or so, a

[51] Carl Carver and Anna Alexandrova, 'No Revolution Necessary: Neural Mechanisms for Economics', *Economics and Philosophy*, 24 (2008), 381–406, p. 382.

[52] Green and Shapiro, 'Pathologies Revisited', p. 267.

[53] Colin Hay, 'Theory, Stylized Heuristic or Self-fulfilling Prophecy? The State of Rational Choice Theory in Public Administration', *Public Administration*, 82 (2004), 39–62, p. 56.

[54] Geoffrey Brennan and Loren Lomasky, *Democracy and Decision: The Pure Theory of Electoral Preference* (Cambridge: Cambridge University Press, 1993).

[55] Ostrom, 'A Behavioural Approach to the Rational Choice Theory of Collective Action'.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

growing number of practitioners seem to have heeded Ostrom's advice. As a consequence, Quackenbush, Lovett and Ferejohn specifically deny that RC theorists are wedded to the assumption of self-interest.[56] On such a reading, RC is perhaps best understood as a 'research tradition' rather than a 'research programme': one constituted by a shared set of epistemic and methodological norms of investigation rather than a 'hard core' of assumptions.[57]

Although she might well baulk at the label, Ostrom's own work on *Governing the Commons* might serve as an exemplar of such a revised and improved RC approach, combining, as it does, an account of how calculative but boundedly rational and self-interested actors struggle with and sometimes manage to overcome collective action problems.[58] Citations might also be awarded here to Dennis Chong's *Collective Action and the Civil Rights Movement* and *Rational Lives: Norms and Values in Politics and Society*, Brennan and Hamlin's *Democratic Devices and Desires*, or Brennan's other work on economizing on virtue or esteem, Regina Smyth's *Candidate Strategies and Electoral Competition* and Jonas Tallberg's *Leadership and Negotiation in the European Union*.[59] Finally, and with perhaps the greatest accompanying fanfare, *Analytical Narratives* promised to develop a 'middle ground between ideographic and nomothetic reasoning' characterized by the combination of 'thin' game-theoretic forms of reasoning with 'thick' descriptions 'capturing the influence of history, the importance of uncertainty, and the capacity of people to manipulate and strategize, as well as the limits of their ability to do so'.[60]

In the course of a case study of central banking, Bell recognizes the existence of what he calls RC2 models, which update and more richly specify the context and behaviour of rational actors and allow for the possibility of endogenous formation of preferences.[61] And he accepts that these models represent a significant improvement upon what he calls RC1 models, in which preferences are assumed to be fixed, motives selfish and reasoning flawless. But Bell is far from satisfied. He calls upon RC theorists to abandon their method and 'move formally in the direction implied' in their recent work by embracing

[56] Stephen Quackenbush, 'The Rationality of Rational Choice Theory', *International Interactions*, 30 (2004), 87–107, pp. 93–4; Lovett, 'Rational Choice Theory and Explanation', p. 254; and John Ferejohn, 'Rational Choice Theory and Social Explanation', *Economics and Philosophy*, 18 (2002), 211–34, p. 218.

[57] On the notion of a research tradition, see Larry Laudan, *Beyond Positivism and Relativism* (Boulder, Colo.: Westview, 1996). For an application to the RC debate, see Kaisa Herne and Maija Setala, 'A Response to the Critique of Rational Choice Theory: Lakatos' and Laudan's Conceptions Applied', *Inquiry*, 47 (2004), 67–85.

[58] Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge: Cambridge University Press, 1990).

[59] Dennis Chong, *Collective Action and the Civil Rights Movement* (Chicago: University of Chicago Press, 1991); and *Rational Lives: Norms and Values in Politics and Society* (Chicago: University of Chicago Press, 2000); Geoffrey Brennan and Alan Hamlin, *Democratic Devices and Desires* (Cambridge: Cambridge University Press, 2000); and 'Economizing on Virtue', *Constitutional Political Economy*, 6 (1995), 35–56; Geoffrey Brennan and Phillip Pettit, *The Economy of Esteem* (Oxford: Oxford University Press, 2004); Regina Smyth, *Candidate Strategies and Electoral Competition in the Russian Federation: Democracy Without Foundation* (Cambridge: Cambridge University Press, 2006); Jonas Tallberg, *Leadership and Negotiation in the European Union* (Cambridge: Cambridge University Press, 2006).

[60] Robert Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal and Barry Weingast, *Analytical Narratives* (Princeton, N.J.: Princeton University Press, 1998). The 'middle ground' quotation can be found on p. 12 and the account of what constitutes a 'thick' description forms p. 14. The book itself was subject to a particularly savage review by Elster, 'Rational Choice History'.

[61] Stephen Bell, 'The Limits of Rational Choice: New Institutionalism in the Test Bed of Central Banking Politics in Australia', *Political Studies*, 50 (2002), 477–96.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

historical institutionalism.[62] This is not a proposal for more 'border crossing'. It is a call for RC to abandon its sovereign status. RC theorists are likely to remain unimpressed. One argument available to them here is to suggest that there is a trade-off between explanatory breadth and depth and that RC can be used to generate more widely applicable explanations of types of events, which can then be modified to fit the details of specific token cases.[63] But putting aside the details of this dispute, there is surely a more general point to be made. In terms of the assumptions they make about individuals' decision-making capacities and motives, there is, as Bell recognizes, a great deal of difference between RC1 and RC2.

WHAT IS RATIONAL CHOICE EXPLAINING?

In this part of the article, I want to shift the focus away from what it is that RC assumes to what it is that it explains. I have so far followed the 'received interpretation' that RC is a form of intentional explanation which accounts for people's actions in terms of their mental states.[64] Von Wright drew a distinction that is still useful here between:

> interpretation or understanding on the one hand and explanation on the other. The results of interpretation are [an] answer to a question 'What *is* this?' Only when we ask *why* there was a demonstration, or which were the 'causes' of the revolution, are we in a narrower and stricter sense trying to explain what there is, the facts.[65]

Von Wright goes on to suggest that, in the social sciences, explanations are frequently 'quasi-causal'.[66] The explanation of a particular event entails identifying its 'contributory causes' with these causes being connected to the event not via a covering law but by a set of singular statements which constitute the premises of practical inferences.[67] Before this quasi-causal explanation can take place, however, it is first necessary to interpret an action which will mean attributing intentions to agents. RC explanations remain distinctive and controversial in so far as they assume not only that agents' actions can be explained in terms of their beliefs and desires but that their actions were the best possible actions they could have taken given their beliefs and that their beliefs were the best possible beliefs given the information available.[68] Critics will, of course, baulk at the optimality constraint here, but the point is that this conventional account of what it is that RC theorists are doing is one that other political scientists will at least recognize. There are, however, alternative accounts available of what it is that RC explains: alternatives that, in turn, can be used as a defence against the charge that RC employs inadequate assumptions.

One particularly hard-line response is that RC theory does *not* attempt to provide psychologically realistic explanations of why people act in certain ways and that criticisms about the accuracy of its assumptions are, therefore, misplaced. RC theory operates with

[62] Bell, 'The Limits of Rational Choice', p. 488.
[63] Keith Dowding, 'The Compatibility of Behaviouralism, Rational Choice and New Institutionalism', *Journal of Theoretical Politics*, 6 (1994), 105–17.
[64] Satz and Ferejohn, 'Rational Choice and Social Theory', p. 71.
[65] Georg Henrik von Wright, *Explanation and Understanding* (London: Routledge & Kegan Paul, 1971), p. 134, emphasis in original.
[66] Von Wright, *Explanation and Understanding*, p. 153.
[67] Von Wright, *Explanation and Understanding*, p. 142.
[68] Elster, *Rational Choice*; Andrew Hindmoor, *Rational Choice* (Basingstoke, Hants.: Palgrave Macmillan, 2006), pp. 190–2.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

a revealed preference view of utility which absolves it from the need to make any claims about mental states. Choices reveal preferences, and rather than thinking about people trying (and often failing) to maximize their utility, the utility function is simply a formal representation of those choices. Binmore suggests:

> one of these days, psychologists will doubtless come up with a workable theory of what goes on in our brains when we decide something. *In the interim, economists get by with no theory at all of why people choose one thing rather than another*. The modern theory of utility makes no attempt to explain choice behaviour. It assumes that we already know what people choose in some situations and uses this data to deduce what they will choose in others.[69]

Nevertheless, the revealed preference view cannot escape making some assumptions about agents, their preferences and their rationality. In order to use information about how people have behaved in one situation to predict how they will behave in another, it must be assumed that agents' preferences are continuous, complete, reflexive and transitive. This, it might be argued, amounts to a substantive set of claims about what goes on in people's brains, which have been empirically tested and frequently found wanting.[70]

Phillip Pettit offers another account of what it is that RC explains. He accepts that most of the time most people are, in non-market settings, content to let their actions be dictated by habits or cultural norms rather than by calculations of self-interest.[71] People are not, by and large, 'rational fools'.[72] They do favours for friends, vote when called upon to vote and give honest directions to strangers. But, for most people, there are natural limits to such behaviour. If, in a particular situation, acting out of habit or norm is likely to significantly threaten self-interest, then 'alarm bells will ring' and people will come to consider their own interests more carefully and, as a result, act in a more self-interested manner.[73] On this account, self-interest normally plays little or no actual role in motivating action. But it is nevertheless 'virtually present in deliberation, for there are alarms which are ready to ring at any point where the agent's interests get to be possibly compromised and those alarms will call up self-regard and give it a more or less controlling deliberative presence'.[74]

Once an alarm bell has rung and an agent has been motivated to act in a self-interested manner, there is obviously no difficulty in seeing how self-interest might enter into the explanation of their actions and why, therefore, RC theory might prove valuable. But how can the virtual presence of self-interest explain anything in those cases where self-interest

[69] Ken Binmore, *Playing for Real: A Text on Game Theory* (Oxford: Oxford University Press, 2007), p. 111, emphasis added.

[70] In particular, behaviourists have argued that preferences can sometimes be shown to be intransitive (so that whilst $A > B$ and $B > C$, $C > A$) and non-continuous. See Khahneman and Tversky, *Choices, Values and Frames*, and the other references in fn. 30. Keith Dowding argues that in order to interpret behaviour, we need to assume that the behaviour is meaningful and that 'we need the formal axioms [of revealed preference theory] to be able to interpret the behaviour' (Keith Dowding, 'Revealed Preference Theory', unpublished manuscript).

[71] Phillip Pettit, *The Common Mind*, pp. 272–8; 'Functional Explanation and Virtual Selection', *British Journal for the Philosophy of Science*, 47 (1996), 291–302; and 'Rational Choice, Functional Selection and Empty Black Boxes', *Journal of Economic Methodology*, 7 (2000), 33–57.

[72] Amartya Sen, 'Rational Fools: A Critique of the Behavioural Foundations of Economic Theory', *Philosophy and Public Affairs*, 6 (1977), 317–44.

[73] Pettit, 'Rational Choice, Functional Selection', p. 42.

[74] Pettit, 'Rational Choice, Functional Selection', p. 42.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

did *not* actually lead an agent to behave in the way they did? Pettit argues that the virtual presence of self-interest may explain the resilience of a particular course of action. Imagine a person who acts in a particular way simply because there is an established cultural norm encouraging them to do so but that they would continue to act in this way even *if* an alarm bell was rung and they started to act purely on the basis of their self-interest. The fact that they would have continued to behave in the same way, even if something that did not happen had happened, cannot help us to explain their actions; yet it can help us to explain the resilience of that action. The fact that RC theory does not present us with an accurate account of how people actually reason does not mean that it is unable to explain the resilience of particular patterns of behaviour.

A third argument here is that RC should be judged not in terms of the realism of its assumptions or the extent of its empirical successes but as a heuristic 'aid to thought'.[75] One version of this argument can be traced back at least as far as Popper's defence of the 'zero method'. Popper argued that it made pragmatic sense to assume perfect rationality and self-interest in the social sciences in order to generate benchmark predictions which could be adjusted in the light of empirical evidence.[76] On such a reading, RC theory serves as a cognitive short-cut: a way of coping with complexity whilst generating falsifiable predictions.[77] In reincarnating Popper's argument, Schotter suggests that RC is 'strong but wrong', and that there is an unavoidable trade-off between a theory's capacity to make predictions and the accuracy of those predictions.[78] The source of RC theory's comparative advantage clearly lies in its capacity to generate clear-cut predictions – that people will defect in a one-shot prisoner's dilemma; that people will propose unequal divisions in a dictator game; and so on. Experimental evidence shows that these predictions are frequently inaccurate. But these predictions, although false, may nevertheless be 'very instructive' in so far as they can be used to inform subsequent research.[79] RC 'does not provide us with the truth about human behaviour but rather with a very compelling platform with which we can seek the truth'.[80]

My reading of Schotter's argument is that RC theory serves as the methodological equivalent of a set of learner plates: a necessary but unrewarding stage, through which most people pass. Hindricks offers a slightly more positive assessment in the course of an essay defending the use of unrealistic assumptions within economics.[81] He argues that models ought to be judged not in terms of the realism of their assumptions but their 'explanatory scope'. RC models resting on implausible assumptions are valuable if they can be used to deduce original and interesting conclusions, which subsequently encourage

[75] Robert Lane, 'What Rational Choice Explains', in Friedman, ed., *The Rational Choice Controversy*, pp. 107–26, at p. 124.

[76] Karl Popper, *The Poverty of Historicism* (London: Routledge & Kegan Paul, 1960).

[77] Indeed, Røgeberg and Nordberg argue that the use of 'absurd assumptions' within economics can be justified because they 'alleviate cognitive limitations in humans by integrating a variety of facts in untrue mock-explanations'. In other words, the fact that we are all boundedly rational explains the comparative advantage of models which assume that this is not the case (see Ole Røgeberg and Morten Nordberg, 'A Defence of Absurd Theories in Economics', *Journal of Economic Methodology*, 12 (2005), 543–62, p. 544.

[78] Andrew Schotter, 'Strong and Wrong**:** The Use of Rational Choice Theory in Experimental Economics', *Journal of Theoretical Politics*, 18 (2006), 498–511.

[79] Schotter, 'Strong and Wrong', p. 504.

[80] Schotter, 'Strong and Wrong', p. 500.

[81] Frank Hindriks, 'False Models as Explanatory Engines', *Philosophy of the Social Sciences*, 38 (2008), 334–60.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

other researchers to investigate whether the same conclusions hold when these underlying assumptions are modified and presumably made more realistic. Consider Downs's popularization of the median voter theorem.[82] If there are two parties, political space is one-dimensional, parties can move to and occupy any point in this space, parties are vote-maximizers, voting is positional rather than directional, voters' preferences are fixed and there is perfect information, then Downs demonstrates that parties will converge upon the position of the median voter. These assumptions might be dismissed as being implausible. Yet the 'explanatory scope' of Downs's work has been such that generations of theorists have used his basic model to delineate the circumstances in which party competition takes a centripetal form.[83]

A slightly different way of phrasing the same argument is to say that RC theory performs a valuable service for political science by exposing the limitations of existing research and identifying new empirical puzzles. Consider Olson's *Logic of Collective Action* and its prediction that unless the number of individuals in a group is quite small, or unless there is coercion or some other device to make individuals act in their common interest, rational self-interested individuals will not act to achieve their common or group interests.[84] As a prediction, this is wrong. Collective action problems are, after all, often overcome. But it might be argued that the real value of Olson's work is to be found in the way in which it exposed the limitations of pluralist theory and encouraged political scientists of all persuasions to consider the circumstances in which people join groups. Similarly, whilst the paradox of voting has proved to be an empirical disaster zone for RC theory, it has stimulated research into the decision to vote. In this vein, Chong concludes that 'like any good theory, rational choice theory raises interesting issues for study and charts a program and methodology of research'.[85]

Finally, and whatever the obvious inadequacies of its assumptions, a final argument here is that RC theory is heuristically useful because it can tell us something about how the world would be *if* in the future people were perfectly rational and entirely self-interested. Just as engineers stress-test buildings and regulators (hopefully) stress-test banks, institutional designers can use the assumptions of RC theory to see whether institutions are capable of withstanding 'knavish' behaviour or can 'economise on virtue'.[86] Alternatively, Hay suggests that RC theory can serve as a useful fable, warning people of the dangers of a world in which people are 'rational fools' and life is nasty and brutish.[87]

The claim that RC theory stands as a 'useful aid to thought' and, more specifically, the claim that it serves the broader interests of political science by generating baseline predictions, setting research agendas and facilitating counterfactual analysis is not uncontroversial. Sceptics might argue that the predictions RC makes are so inaccurate as to be of no value; that designing institutions on the assumption that people are knaves simply encourages knavish behaviour; and that, with the notable exception of Tsebelis's work on *Veto Players*,

[82] Anthony Downs, *An Economic Theory of Democracy* (New York: Harper & Row, 1957).

[83] This research is reviewed in Andrew Hindmoor, *Rational Choice* (Basingstoke, Hants.: Palgrave Macmillan, 2006), pp. 30–45.

[84] Mancur Olson, *The Logic of Collective Action* (Cambridge, Mass.: Harvard University Press, 1965), p. 2.

[85] Chong, 'Mysterious Rivals', p. 39.

[86] Robert Goodin, 'Institutional Design and Rational Choice', in Robert Goodin, ed., *The Theory of Institutional Design* (Cambridge: Cambridge University Press, 1996); Brennan and Hamlin, *Economising on Virtue*.

[87] Hay, 'Theory, Stylized Heuristic or Self-fulfilling Prophecy'.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

which has generated an active research programme within comparative politics, RC theory has become increasingly method-driven and uninteresting.[88] Yet the heuristic defence of RC theory nevertheless serves to narrow the gap with the rest of political science. There is a great difference between arguing that RC will be a field within economics and will encompass all of political science and suggesting that RC is the handmaiden of progress within political science.

One final argument in this part of the paper takes us back to *Pathologies of Rational Choice*. Green and Shapiro, it will be recalled, equate theoretical value with empirical potency: 'A theory of politics has no payoff if its hypotheses do not survive empirical scrutiny'.[89] In doing so, they embraced an epistemological standard shared by many – but not all – RC theorists. Cox, for example, suggests that RC theory can be used to provide 'more satisfying' explanations, even when it does not add to our existing empirical knowledge.[90] How so? One argument here, which Cox appears to edge towards, is that explanation can involve showing how something fits into a more general pattern. Explanation, that is, can come from unification. Just as Newton showed how tides, projectiles and planetary motion could be explained using the same set of underlying principles, RC theory shows how it is possible to account for a diverse range of phenomena using a small number of core assumptions.[91] On such a reading, the most valuable works in the RC corpus are not those which exhibit an 'impressive integration of formal modelling and empirical testing' but those like, for example, Russell Hardin's explanation of ethnic conflict in *One for All*, which apply the RC method to new areas of work.[92]

The unificationist argument, or, more precisely, the argument that explanation through unification can be achieved independently of causal explanation, remains controversial. In Kinkaid's discussion, he draws a distinction between a unification which consists in the provision of a common classificatory scheme or descriptive vocabulary, where no satisfactory scheme previously existed, and *physical* unification, where phenomena previously regarded as having quite different causes or explanations are shown to be the result of a common set of mechanisms or causal relationships.[93] The distinction is relevant because critics are likely to argue that whilst RC theory provides a common descriptive vocabulary and set of methods, it falls far short of physical unification, because – precisely as *Pathologies* showed – it has been unable to use a small number of assumptions to reveal common causal relationships. Furthermore, a critic might argue that, in the absence of this physical foundation, unification simply licenses a method-driven *Flight from Reality* in which, rather than explaining some event or process, effort is exclusively

[88] Tsebelis, *Veto Players*.

[89] Green and Shapiro, *Pathologies of Rational Choice*, p. 32.

[90] Gary Cox, 'The Empirical Content of Rational Choice Theory: A Reply to Green and Shapiro', *Journal of Theoretical Politics*, 11 (1999), 147–69, p. 162.

[91] The Newtonian example is taken from Harold Kincaid, *Philosophical Foundations of the Social Sciences* (Cambridge: Cambridge University Press, 1996), p. 55. The strongest defence of unification as a goal of explanation is still that provided by Philip Kitcher, 'Explanatory Unification and the Causal Structure of the World', in Philip Kitcher and Wesley Salmon, eds, *Scientific Explanation* (Minneapolis: University of Minnesota Press, 1989).

[92] Russell Hardin, *One for All: The Logic of Group Conflict* (Princeton, N.J.: Princeton University Press, 1995).

[93] Harold Kincaid, *Philosophical Foundations of the Social Sciences* (Cambridge: Cambridge University Press, 1996), p. 97.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

devoted to showing how that event or process might be reconciled with the use of the RC framework.[94]

THE METHOD OF ISOLATION

I have already suggested that the 'turn' towards mechanism-centred explanation creates a point of common ground between RC and the rest of political science. Mechanisms cause things to happen in particular ways. But particular mechanisms only cause things to happen in predictable ways when left to operate in isolation. In the social sciences, prediction is extremely difficult because most outcomes are the results of the operation of a number of mechanisms. So how can we know what difference particular mechanisms make? One answer to this question involves the application of the 'method of isolation'. Models constructed using the method of isolation take account of just some of the factors which are at work in the real world: the potential influence of other factors being 'sealed-off'.[95] The question I address in the final section is this: can the use of the method of isolation justify the use of unrealistic assumptions within RC? If it does, then RC theorists can acknowledge that the 'hard core' assumptions they employ are inaccurate but nevertheless claim that the RC models which employ them can explain events and outcomes.

Within the *Journal of Economic Methodology*, the method of isolation has been argued to justify the use of unrealistic assumptions.[96] Consider Schelling's famous 'chequerboard' model of racial discrimination, which was lauded as a classic example of mechanistic explanation.[97] Here, two sets of counters (dimes and pennies in Schelling's initial presentation) are arranged randomly on a chequerboard, each cell of which adjoins eight other cells. Schelling postulates a condition to determine whether a coin is to remain in its initial place. For example, a rule can be introduced stipulating that a coin is to be moved when more than two-thirds of the coins on the adjoining cells are of a different type. Schelling shows that, over time, and almost regardless of the initial distribution, such a rule results in an outcome in which the two sets of coins occupy different parts of the board. The point is to show how even mildly segregationist preferences can, over time, lead to a quite rigid racial division within cities.

As a model of residential moving and racial segregation, the chequerboard model makes a number of unrealistic assumptions, including costless moving and the absence of any income differences between the two sets of coins. For this reason, Schelling's model cannot be thought to constitute a complete explanation of the dynamics of racial segregation. Yet the use of these unrealistic assumptions might be defended on the grounds that they

[94] Ian Shapiro, *The Flight from Reality in the Human Sciences* (Princeton, N.J.: Princeton University Press, 2005).

[95] Uskali Mäki, 'On the Method of Isolation in Economics', *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 26 (1992), 316–51, p. 321.

[96] Caterina Marchionni, 'Contrastive Explanation and Unrealistic Models: The Case of the New Economic Geography' *Journal of Economic Methodology*, 13 (2006), 425–46; Menno Rol, 'Idealization, Abstraction, and the Policy Relevance of Economic Theories', *Journal of Economic Methodology*, 15 (2008), 69–97; Frank Hindriks, 'Tractability Assumptions and the Musgrave–Maki Typology', *Journal of Economic Methodology*, 13 (2006), 401–23; and Hindricks, 'False Models as Explanatory Engines'.

[97] Tyler Cowen, 'Do Economists Use Social Mechanisms to Explain?', in Peter Hedstrom and Richard Swedberg, eds, *Social Mechanisms: An Analytical Approach to Social Theory* (Cambridge: Cambridge University Press, 1998); Thomas Schelling, 'Models of Segregation', *American Economic Review*, 59 (1969), 488–93; 'Dynamic Models of Segregation', *Journal of Mathematical Sociology*, 1 (1971), 143–86; and *Micromotives and Macrobehavior* (New York: W.W. Norton, 1978).

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

allow Schelling to 'seal-off' his model from the effects of other mechanisms in order to show how, when left undisturbed, the sorting mechanism operates.[98] In this way, Schelling's model can be said to provide a 'credible' account of racial discrimination even though it employs unrealistic assumptions.[99]

There is, however, a significant problem here. The method of isolation can justify the use of assumptions which are unrealistic in the sense of being partial. Schelling's model 'isolates away a host of elements', including income differentials in order to focus upon the operation of the sorting mechanism.[100] But it might be argued that the core assumptions of RC theory are unrealistic in the very different sense of being simply false. People are not infallible and self-interested decision makers. The problem with RC theory comes, in other words, not with what is left out of the models but what is put in.

So, at this point, it would seem that the method of isolation needs after all to be run in conjunction with an alternative account of what it is that RC is seeking to explain: taking us back to the arguments of the previous section. One such alternative here can be found in Ferejohn and Satz's argument that RC theory constitutes a form of structural explanation in which it is variations in the constraints on actors' choice sets and not variations in their motives and decision-making capacities which carry the explanatory burden.[101] Consider the example of a RC theorist who wants to explain historic variations in fertility rates. Ferejohn argues that the theorist will want to start by assuming that agents are rational and that they seek to maximize their wealth, but these assumptions do not themselves explain anything. Their role is instead to provide a fixed backdrop against which it can be more clearly seen how parametric shifts in constraints – presumably including such things as changes in inheritance laws and the introduction of labour-saving technologies – affect behaviour.

In analysing the operation of structural, that is aggregative, mechanisms like free riding, sorting, structure-induced equilibrium, cycling and agenda setting, RC theorists need not argue that people really *are* perfectly rational and self-interested or that variations in decision-making capacities and motives associated with the operation of individual-level mechanisms like weakness of will, framing and anchoring are of no explanatory significance. The RC theorist can simply argue that they are practising a division of labour and 'sealing-off' the influence of these other factors in a manner consistent with the method of isolation.

As I have presented it here, the structural argument implies a certain degree of modesty. RC theory offers incomplete explanations of political processes and outcomes which focus upon the operation of aggregative mechanisms. I argued before that partial universalism creates the possibility of a peaceful coexistence between RC and other branches of political science. The same applies here. RC theorists who commit to the structural defence need not deny the potential value of other explanatory approaches which analyse the

[98] Following Hindricks, it might also be argued that Schelling's model has been shown to have 'explanatory depth'. For details of the various ways in which Schelling's initial assumptions have been relaxed and modified, see N. Emrah Aydinonat, 'Models, Conjectures and Exploration: An Analysis of Schelling's Checkerboard Model of Residential Segregation', *Journal of Economic Methodology*, 14 (2007), 429–54.

[99] Robert Sugden, 'Credible Worlds: The Status of Theoretical Models in Economics', *Journal of Economic Methodology*, 7 (2000), 1–31.

[100] Marchionni, 'Contrastive Explanation and Unrealistic Models', p. 428.

[101] Satz and Ferejohn, 'Rational Choice and Social Theory'; Ferejohn has subsequently argued that: 'this is pretty much the standard way that rational choice theory is applied to explaining social phenomena' (John Ferejohn, 'Rational Choice Theory and Social Explanation', *Economics and Philosophy*, 18 (2002), 211–34, p. 226).

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

operation of either individual-level mechanisms, like weakness of will, or even aggregative mechanisms, like group think, which assume sub-optimal decision making. But in judging the value of the resulting explanations, a key question suggests itself. Is the operation of the aggregative mechanisms identified by RC theorists using the assumption that individuals are rational and self-interested robust to variations in decision-making capacities and motives? If the answer to this question is 'Yes', then these aggregative mechanisms will continue to operate when people are boundedly rational and boundedly selfish, RC explanations can be widely applied and the assumption that people are perfectly rational and self-interested justified on the purely pragmatic grounds that it makes it easier to construct models. If the answer to this question is 'No', then the status of the initial assumptions does matter as aggregative mechanisms will only operate when people *are* perfectly rational and self-interested and RC models will only have a limited explanatory reach. Indeed, at this point, the structural argument simply becomes a variant of partial universalism.

The correct answer to this question about the reach of RC structural explanations will, I suspect, vary depending upon the particular mechanisms under consideration. Consider, once again, free riding. Looking back at Olson's original argument, he formulates the collective action problem on the specific assumption that individuals are self-interested.[102] But Chong argues that the free-riding mechanism still operates in those situations in which participants are motivated by 'extra-rational' considerations.

> The rational choice analysis of strategic behaviour is likely to be important in any examination of collective action, whether we make simple or complex assumptions about motives. Social norms, expressive and symbolic actions, and even habitual behaviour tend to have strategic components that are amenable to rational choice analysis, even if instrumental rationality does not appear to be the whole story.[103]

Equally, whilst Downs's initial argument rested upon the assumption that politicians are vote-maximizers who 'act solely in order to attain the income, prestige and power which come from being in office', the same competitive mechanism pushing parties towards the electoral centre has been argued to operate when politicians are policy seekers or motivated purely by ideological considerations.[104] There are, however, other mechanisms whose operation would seem to be parasitic upon the assumption of self-interest. The levels of free riding within a single play of the prisoner's dilemma, as one variation of the collective action problem, would, for example, seem to be significantly affected by actors' underlying motives.[105] Here, and as in the ultimatum game, the evidence suggests that people consistently behave contrary to the standard expectations of RC theory.

A common complaint about RC theory is that it is unnecessarily formal and that this formalism is used to hide its dependence upon implausible assumptions. Yet, in this context, it might equally well be objected that it is the informality of RC theory which often poses as a significant problem. Too many RC models that point to the operation of aggregative mechanisms fail to clearly specify the range of circumstances in which deductions from the model remain valid.[106]

[102] Olson, *Logic of Collective Action*, p. 16.
[103] Chong, 'Mysterious Rivals', pp. 52–3.
[104] Downs, *An Economic Theory of Democracy*, p. 27.
[105] Hindmoor, *Rational Choice*, pp. 122–5.
[106] Nancy Cartwright, 'The Vanity of Rigor in Economics: Theoretical Models and Galilean Experiments', in *Hunting Causes and Using Them* (Cambridge: Cambridge University Press, 2007), pp. 217–61.

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press

CONCLUSION

At the very start of this article, I cited Shepsle's description of RC as enduring a mid-life crisis. The analogy is an appealing one, because the timing is right. If we date the origins of RC to the publication of Downs's *An Economic Theory of Democracy*, Riker's *The Theory of Political Coalitions* and Olson's *The Logic of Collective Action* in the late 1950s and early 1960s, then RC was approaching its fortieth birthday when Green and Shapiro published *Pathologies*.[107] Stretching Shepsle's analogy, critics will want to argue that after promising so much, the limitations of RC theory were painfully exposed in the early 1990s and that despite incessant protestations to the contrary, there has been no subsequent recovery, but care must be taken here. It is perfectly possible to slide into middle age in the absence of any crisis. Proponents will want to argue that RC has grown older gracefully; that it has acquired greater depth and lost some of the rough edges exposed by earlier critics. In particular, proponents will point to the wider range of assumptions and empirical tests now employed.

Having previously characterized myself as an engaged agnostic, I am reluctant to end this lengthy review with any clear-cut assessment of whether RC has withstood the criticism levelled at it. What does seem clear is that RC – or, perhaps, more specifically, the way in which RC is defended – has changed, and has changed in a way which has narrowed the gap with the rest of political science. Partial universalism, the use of a wider range of assumptions, the method of isolation and the defences of RC as a heuristic device and as explaining resilience take some of the sting out of earlier criticisms. What is also clear is that RC is no longer the *enfant terrible* of the political science world. Over the last decade, historical institutionalism and, more recently, constructivist institutionalism and behaviourism have emerged as credible rivals and attracted considerable attention. The RC debate will go on. It no longer seems plausible to argue that RC will subsume the rest of political science. But neither does it seem likely that RC will simply disappear in the same way as, for example, systems theory. Nonetheless, the debate has changed.

[107] Downs, *An Economic Theory of Democracy*; Olson, *The Logic of Collective Action*; and William Riker, *The Theory of Political Coalitions* (New Haven, Conn.: Yale University Press, 1962).

https://doi.org/10.1017/S0007123410000372 Published online by Cambridge University Press