Game Theory:

Today's Socratic Method?

(First Draft)

Laurent Dobuzinskis

dobuzins@sfu.ca

Political Science, Simon Fraser University

Paper presented to the Annual meeting of the
Western Political Science Association
Vancouver, BC   April 14, 2017

Abstract

Originally developed as a mathematical instrument for predicting the outcome of conflicts among rational actors and for assisting decision-makers in making decisions in situations where the rules of the game are clear, game theory has branched out in many directions. Its epistemological implications have become varied and complex. While its predictive power has arguably been disappointing in some contexts, it has become a rich source of philosophical and practical questions about the nature of the "games" that are played in politics or the economy; about the meaning of rationality; or about the dialectical interplay of cooperation and competition, self-interest and altruism, or the centrality of reciprocity in social relations. Political theorist or policy analysts starting from very different epistemological or ideological positions may be surprised by the relevance of game theory to their concerns and the unique opportunity it offers to reframe and rethink some of their most cherished certainties. In other words, it is best thought of as a Socratic method rather than a set of ready-mades "solutions" to conventional problems.

Keywords: Nash equilibrium; strategic actors; prisoner's dilemma; cooperative games; evolutionary game theory; methodological individualism; cooperation; reciprocity.

Game Theory:

Today's Socratic Method?

(First Draft)

Introduction

What is the relevance of game theory to political scientists today? A tempting answer—and I argue that this temptation must be resisted—is to say that game theory is useful for those who work within the positivist paradigm but that it does not have much to contribute to other types of research projects. I want to argue that not only is this belief unfounded, but that even if game theory is still closer to the positivist end of the spectrum that runs from the most postmodern or "constructivist" paradigms to positivist methods, provided that this dichotomy is still meaningful considering how the philosophy of social ciences has undermined its validity, game theory has some unique characteristics that speak to the methodological concerns of most researchers in the social sciences. This is why game theory is arguably the most important advance in the methodology of the social sciences in the 20th century and remains a much-discussed topic (Leonard 2010; Amadae 2016; Erickson 2015). The history of game theory strongly suggests that it has branched in many directions and has become a richer and more intriguing program than is usually assumed. If nothing else, it can be an invitation to probe deeper into questions that go the heart of political science research.

Game theory was not originally framed by pioneers such as John von Neumann, Oskar Morgenstern, or John Nash, merely as a way of raising questions. They argued in favour of its predictive power. In spite of very important theoretical advances, however, the predictive power of game theoretic models is more in doubt today than it was when economists and other social scientists became interested in such models. But the paradoxes and epistemological conundrums for which game theory is now

known continue—or ought to continue—to challenge students of politics or political economy. Game theorists themselves have thought deeply about these problems, as I explain below.

In the first section, I briefly retrace the history of game theory to set the context of the arguments I advance in the rest of the paper. In section II, I suggest that the questions that inform the design of game theoretic models ought to be raised even when undertaking projects that do not fall within the purview of game theory strictly defined. Then in section III, I argue that game theory has contributed decisively to a re-thinking and still on-going deepening of our understanding of the relationship between ontology and epistemology in political science, and more specifically of the epistemological significance of unavoidable concepts such as rationality, (self-)interest and reciprocity.

I. How Game theory has Become what it is Today

The contemporary literature on game theory offers varied and sometimes contradictory interpretations. Even more reasons to go back to its roots. There are excellent works on the history of game theory (Leonard 1995; 2006; 2010; Dimand and Dimand 2002; Giocoli 2003; 2009; Erickson 2015) which show how it has been shaped by a series of on-going and sometimes tortuous debates about cross-cutting methodological and discipline-specific issues. But there is general agreement that in it originated with the publication in 1944 of John von Neumann and Oskar Morgenstern's *The Theory of Games and Economic Behavior* (*TGEB*). One of the stated purposes of this seminal work was "to find the mathematically complete principles which define 'rational behavior' for the participants in a social economy, and to derive from them the general characteristics of that behavior" (2004 [1944]: 31). Another milestone in the history of game theory was of course the publication in 1951 of John Nash's article outlining what became known as the "Nash equilibrium" in non-cooperative games. But a few more distant antecedents must also be mentioned.

Strategic thinking is an essential aspect of life in a social setting. People make choices, knowing full well that others do the same; some of these merely involve deciding among practical issues (e.g., choosing a spouse or forging alliances with neighbouring groups), other choices involve moral challenges. I simply take this as an experiential "given." Therefore, it is not surprising that one finds examples of game theoretic reasoning in the Bible or the Talmudic tradition (Aumann 2002; Brams, 2011: chapter 2). And of course, all successful military leaders, going back to Pericles or Hannibal, have devised winning strategies (although the term "strategy" in its modern connotation did not emerge until the beginning of the 19th century). Niccolo Machiavelli famously strategized about war and politics in his *The Prince*. Jean-Jacques Rousseau's allegory of the dilemma faced by hunters who have to chose between cooperating in hunting a deer or going off individually after a rabbit is the origin of the much-discussed eponymous game.[1] Yet in a sense Rousseau was the first critic of the application of game theory to politics: in the *Social Contract*, he goes to great pain to explain that the "General Will" can emerge only when voters vote sincerely listening exclusively to their conscience. Strategic voting can only result in the qualitatively inferior—at least in Rousseau's eyes—sum of individual interests (or "particular wills"). Michael Chwe (2013) argues that Jane Austen also intuited many game theoretic principles and subtly showed that strategic thinking is often used by less powerful individuals in a manner that takes advantage of the "cluelessness" of those who hold power over them.

Antoine-Augustin Cournot was the first economist in the early 19th century to systematically use mathematics to analyze economic problems. In his analysis of monopolies and duopolies, Cournot didn't make use of the concepts deployed by game theorists but anticipated the Nash equilibrium. Nash's genius is to have thought of his famous equilibrium concept as a solution to *all* non-cooperative games.

---

[1] As intellectual fashions go, "stag hunt" seems to have dethroned Prisoner's Dilemma as the most intriguing game (e.g., Skyrms 2004).

The turn of the last century was a crucial time in the history of mathematics, logic and, as it turned out, game theory. Two streams of apparently unrelated inquiries were captivating mathematicians and philosophers in Mitteleuropa: set theory and the game of chess (Leonard 1995: 733; Ferreiros 2012: 12-13). The mathematician Ernst Zermelo's work placed him at the confluence of these streams; he began to think of chess as a set-theoretic problem. He deliberately bracketed out the psychological aspects of the game.[2] In a paper published in 1912, Zermelo proved by following what is today called "backward induction" that it is always the case that one of three "solutions" exist: either white has a winning strategy; or black has a winning strategy; or "each of the two player has a strategy guaranteeing at least a draw (Maschler et al. 2013: 3). None of this was unfamiliar to von Neumann, who knew Zermelo, but he worked on a much more ambitious problem: finding a solution to *all* zero-sum games. In a paper he presented to the Göttingen Mathematical Society in 1926, von Neumann laid down a proof for the so-called minimax theorem. All two-person zero-sum games have either a pure (minimum of the maximum gains/maximum of the minimum gains, i.e., the "saddle point") or mixed (i.e., randomized) winning strategy. Interestingly, as Leonard (1995, 734) notes, "with the minimax theorem, the prevailing probabilistic view of the world in physics was being reflected in von Neumann's theory of human interactions."

John von Neumann pursued other unrelated problems for a decade, then returned to the mathematical problems in the late 1930s. The focus of *TGEB* on economics is largely attributable to Morgenstern, however. The long introductory chapter bears the imprint of Morgenstern's idiosyncratic

[2] In doing so, Zermelo set himself in opposition to what was probably the dominant hypothesis among philosophers and mathematicians interested in the theory of games, namely, that games are not reducible to pure mathematics because the psychology of the players in a large measure explains the outcome. This position was advanced by the chess-master and dilettante philosopher Emanuel Lasker with respect to chess, and the French mathematician Emile Borel, with respects to games of card. It is interesting to note that today the behavioural approach to game theory is rediscovering the importance of psychology.

views on economic theory. Morgenstern was highly critical of practically all schools of thought in economics: "he criticized Hicks, Hayek, Keynes, Samuelson, the business cycle theorists and all his colleagues at the Princeton Economics Department" (Giocoli 2001, 170). Before moving to the US, Morgenstern's research program united two seemingly irreconcilable schools of thought . The first was Austrian economics. Morgenstern attended von Mises' seminar in Vienna and worked as an assistant to Hans Mayer. From both he acquired the typically Austrian scepticism toward Walrasian general equilibrium.[3] Austrian economists (then and now) are opposed to using mathematics to study what they consider irreducibly complex economic problems. In his book *The Limits of Economics* (1937), Morgenstern expressed his critical opinions about the relevance, or lack thereof, of statistics in economic analysis. But—and this is the second facet of his research program—Morgenstern was very much influenced by the epistemological views of the mathematician Karl Menger whom he met in the 1930s. Karl Menger strongly argued in favour of not only the axiomatization of mathematics but also of a similarly formalist approach to the study of social phenomena (Giocoli 2001: 170). Morgenstern's mathematical skills were limited but when the opportunity arose, he teamed with von Neumann in the hope of setting economics on a sound mathematical footing (i.e., set theory), while unmooring it from the neoclassical paradigm and its unreflective use of mathematical tools. The outcome of their common efforts (*TGEB*), covers a wide range of topics in a somewhat disjointed manner: zero-sum games, the concept of expected utility, and cooperative games. It is worth noting that while in its early stages, game

---

[3] Mayer is rarely mentioned by Austrian economists—and in fact Hayek dismissed him as a rather insignificant scholar (an it is true that he did not publish much); this was due largely to the disgraceful way in which Mayer rallied to the Nazis when they took power in 1938 and helped them purged the Jews and liberals from his department. And when the Soviets took over in 1945, he successfully managed to keep his position. But, as Leonard (2013: 77-78) argues on the basis of a close reading of his writings on economic theory, Mayer's views of market processes were not very different from those of Hayek and Mises.

theory was often associated with military applications largely coterminous with lethal zero-sum games, about half of *TGEB* is about cooperation and coalition formation.

The second momentous step in the history of game theory took place in the early 1950s: first in his PhD dissertation (which he defended in 1950) and in an article published in 1951, John Nash proposed a solution to all non zero-sum non-cooperative games. The Nash equilibrium consists of the best response of each player to each other's strategies; in some games, there will be more than one such equilibrium (e.g., "chicken"). As it turned out that the Nash equilibrium is far more relevant to the social sciences than von Neumann's works, but it took quite some time for scholars to perceive the significance of this concept.

In the 1940s and 1950s, game theorists were mostly to be found in the Pentagon or the RAND Corporation, but this is not necessarily the proof that game theory is inherently a child of the Cold War paranoia. It was more a consequence of the fact that mathematicians interested in game theory had few other options to pursue their interests, considering that the rather tepid response from social scientists to the new approach at the time (see below).

The epistemological tenor, let alone the ideological orientation of classical game theory is ambiguous. Although some of its premises were anticipated in more literary forms, it belongs to applied mathematics. Set theory rather than a somewhat circumstantial association with war planning is truly the fundamental origin of game theory. It is the uses to which the theory has been put rather than the theory itself that have an ideological connotation. Just as it makes little sense to speak of the ideological nature of genetics, even if eugenics was a reactionary project, game theory is hard to pin down on an ideological spectrum, even it has occasionally been harnessed by militaristic or elitist projects. It remains that an epistemological critique of game theory is legitimate, albeit challenging. Set theory is vaguely Platonist and the idea that essential aspect of the human experience, such a choosing between

alternative ways of acting, can be reduced to mathematical problems for which there are universally valid "solutions" rests on rationalist premises. Rationalism has been criticized for its tendency to flatten personal, cultural and social differences, and the value conflicts that stand in the way of an objective assessment of strategic options. This is one of the flaws that critics invoke to reject "positivism." Rationalism has been rebuked by both conservative liberal, such as Michael Oakeshott or F.A. Hayek, and by progressive communitarians. Indeed, Hayek's critique of "scientism" and more specifically his scepticism about the use of mathematics in economic analysis account for the lack of enthusiasm toward game theory displayed by the followers of the Austrian school of (Cachanosky 2010). For their part, communitarian and feminist critics have zeroed in on methodological individualism. As long as game theoretic models treat "players" as individuals without ties to their communities and without identifiable gender, critics argue they will obfuscate more than they will explain. As Peregrine Schwartz-Shea (2002: 301) remarks,

> For feminist researchers and others interested in understanding the ways in which societies
>
> shape and are shaped by sex, gender, race, and similar factors, the continuing influence of game
>
> theory is troublesome. Because the game-theoretic individual has no (theorized) sex, much less
>
> gender or race, the substantive issues concerning sex/gender are, at best, invisible in such
>
> analyses.

In a similar vein, S.M. Amadae (2016, 3-61) faults game theory for being rooted in a "neoliberal subjectivity" that betrays a reductionist understanding of human agency characterized by "the commodification of all values," a rationalization of "cheating and free riding," and a "quasi-Darwinian worldview." It rejects the "logic of appropriateness" (Amadae 2016: xvii) embodied in older forms of classical liberalism.

What these critiques rightly imply is that the abstract language of mathematical proofs serves a purpose that is distinct from that pursued by those who want to understand the many facets of life in social contexts shaped by history, traditions and contingent events. But it would be equally wrong to deny that strategic calculations never entered the minds of those who shaped the institutions and practices that account for differences among groups, cultures, and historical epochs. What counts as a good strategy in one context may make little sense in another, but the logic of strategic thinking is rarely absent from these different contexts. Imaginatively "playing" with the suggestion that agents are not always doing what they claim or appear to be doing is a way for social scientists to question unexamined beliefs. Michel Foucault's notion of micropolitics comes to mind here, except that there is no a priori reason to limit the outcomes of these multitudinous games to shifts in the distribution of power attributes; agents can make strategic choices driven by profane motives or religious beliefs, by a desire for esteem, care for others and quite a few other social or personal motivations.

But I concede that when targeting classical game theory, such criticisms hit their mark, although an important qualifier must be added. Game theory occupies a unique place in the rational choice paradigm in that it posits that being rational implies basing one's choices not simply on a solipsistic conception of one's self-interest but also on what *others* have in mind. "Players" interact; they anticipate each other's moves. Notwithstanding von Neumann's intention of setting aside all psychological considerations in the construction of a "theory of games," subjectivity creeped back in (Dufwenberg 2010). Consequently, game theory has moved beyond its classical framework to a new phase in its development. Modeling players who must reason in terms of mutual adjustments to each other's moves has alerted game theorists to the profound uncertainty inherent in most social interactions. This question has become central to the "epistemic program" in game theory which "asks: what do different notions of rationality [,] and different assumptions about what players believe about what others believe about the

rationality of players [,] imply regarding play in a game?" (Dekel and Sinischalchi 2015, 620; see also Perea 2014). In a nutshell, these new directions in game theory start from the more challenging hypothesis that the players know little about each other and have to navigate the difficulties that this poses. While this new phase in the unfolding of the game theory program cannot properly be described as "interpretivist" in the ordinary sense of the term, it opens opportunities for dialogues between game theorists and critical theorists.

Another crucial move away from the simplistic construct of *homo economicus* took place when evolutionary biologists (e.g., Maynard-Smith 1982) began to apply game theoretic models to the study of population dynamics. Evolutionary game theory is by now a field into itself which is not strictly limited to biology. Its defining characteristics is that it dispenses with the notion of rational players altogether, substituting instead competition among strategies. Here too information is less than perfect. (hence the crucial importance of "signaling" in evolutionary game theory.) Not all these theoretical and empirical explorations converge, but a unifying conclusion might be that institutional rules, cultural norms and conventions, as well as psychological dispositions are just as important as sheer self-interest for understanding social dynamics. It is not surprising, therefore, that many political philosophers have taken a strong interest in evolutionary game theory (Skyrms 2004; 2014; Binmore 2005). Experimental game theory—which has arguably become the proverbial "cutting edge" in game theoretic research—has confirmed that players draw from a wide range of norms and beliefs when deciding how to respond to other players' moves.

Applications of game theory in economics and political science have built upon all these advances in various ways—too many in fact to analyze them in detail here. Suffices to say that the convergence of experimental game theory and social psychology has increased the reach and enriched the meaning of the concepts and methods first articulated by Nash or von Neumann. Their contribution

to the social enterprise as a whole is more promising than ever. The diversity of uses to which game theory is now put belies any simplistic characterization of the epistemic implications of the theory. I cannot pretend to analyze here all the reasons that justify this assertion but I can begin this journey by showing how the construction of game theoretic models can inform a wide range of research programs in ways that those who hold somewhat conventional or superficial views about game theory may not suspect.

II. Game Theory as a Research Design Tool

The necessity of carefully laying down a research design before undertaking actual research is now widely acknowledged in political science (King, 1995; Toshkov, 2016). In what follows I consider the extent to which the steps that researchers intent on formulating a game theoretic model must follow could also prove to be relevant to other projects in which the concept of game is used in a more metaphorical manner, or perhaps not at all. This is obviously a matter of assembling the "nuts and bolts" of any workable model but, inevitably, it raises more difficult epistemological questions.

Identifying the "Players"

One of the issues that must be addressed first when considering situations where various actors are involved in either conflictual or cooperative relations is: Who are the "players"? A host of subsidiary questions follow from that initial assessment. What are the payer's motivations? Do they know the rules of the "game" they are involved in? What sort of information do they have on the other players? (Game theorists make an important distinction in this regard between "incomplete" information, i.e., what are the antecedent moves decided upon by the other player(s), and "imperfect" information, i.e., what the other player knows or believes about the nature of the game being played.[4]) And so on. I leave aside for

---

[4] Failures to coordinate on the same game can have disastrous consequence in international relations. The risk is greater when the players come very different cultural and ideological horizons, as appears to be the case

the moment the question of the extent to which the players are "rational" to which I return in the next section. In many problems, the question of who the players are may seem to be rather trivial. For example, if one is interested in the behaviour of legislators in parliamentary committees or of diplomats in say trade negotiations, the structure of the game may seem to be predetermined. But this question brings to light—and this precisely my point about the contribution of game theory to research, namely, that it compels researchers to take a position on issues that are sometimes swept under the proverbial carpet—a very important matter in the field of International Relations: what is gained and what is lost by treating states as (rational) players? Is it really the case that Russia, China, the United Kingdom or the United States pursue their own unified strategies on the international science. As Bruce Bueno de Mesquita (2006) has argued, a very fruitful perspective on diplomacy first articulated in 1933 by the political scientists James Russell and Quincy Wright, namely, that in democracies foreign relations are largely constrained by domestic political considerations, laid dormant for may decades until the recent (re)discovery of the democratic peace thesis. Central to that thesis is the observation that democracies there are many actors who are more or less directly and visibly involved in the formulation of foreign policy priorities; it is therefore only a crude approximation to treat democratic "states" as and single-minded actors on the international scene. This may be truer of authoritarian leaders but then again *who* the leader is matters a great deal (for example, figuring out what motivates Kin-Jon-un is essential in dealing with North Korea).

So far, I have addressed issues that game theorists and scholars who use game theory in their empirical research are well familiar with. I now turn to an issue that may open promising opportunities

---

with the relationship between North Korea and pretty much the rest of the world. Rui De Figuereido et al. (2006) have argued that the American war of independence can best be explained by the difference in beliefs that set apart the British government (committed to parliamentary sovereignty) and the Amercan revolutionaries (convinced of their rights under traditional English liberties).

for dialogue between that category of scholars and those who lean more toward critical theories. By definition, game theory is concerned with *interactions* among agents who in choosing the best way to react to each other's moves must anticipate how others are likely to behave. This is indeed of the most valuable and profound contribution of game theory to the social sciences; strategic thinking presupposes a social environment in which the players, even if they are supposed to be motivated by nothing else than narrow self-interest (I return to this question in the next section), must figure out how their own interests conflicts or meshes that of others. But who are these "others"? Scholars who use game theory do not often venture beyond conventional conceptualization of the institutional or cultural contexts within which the problems they are interested in are framed. But critical theorists are always asking: whose voices are silenced in such contexts. So the question "who are the players?" is far from anodyne. It may not always be practically possible to include in a formal model players who, precisely, have no voice and play no apparently significant role in the problem at hand. But in many other contexts, game theoretic models could be improved by looking further into the implications of this subject.[5] Moreover, if one considers not simply the issue of how to construct a research design suited for the application of game theory to a particular problem but the issue of how to improve research design principles in general, the question acquires an added significance. Even researchers whose methodological or ideological inclinations differ more or less significantly from those of critical theorists would benefit from addressing the sort of questions that the latter raise about why marginalized groups are, in fact, relegated to the margin and whether doing so in one's own research leads to bad research designs.

---

[5] Although not addressing specifically the question of how to include marginalized voices in game theoretic models, theorists such as Landa and Meirowitz (2009) or Johnson (1993; 2008) have made efforts to address some of the concerns of critical theorists about how politics is shaped by diverging interpretations of what social problems are all about.

A related question is why do some individuals or group prefer *not* to play "the game." Could it be that they are playing parallel games? Such games are not entirely distinctive from each other, neither in theory nor in practice (even hermits play some role—in that case a symbolic one—in the world they seek to escape from) but there can be issues in bridging information generated across these games. This notion can in fact be found in the game theory literature[6] but it can also be deployed in a more metaphorical manner, depending on the context and style of the research project. Thus, one way to think of radical critics of conventional politics and morality is to say that they engage in such practices, considerably complicating conventional democratic politics. Ruth Lane (2007) goes even further and argues that individuals must always seek to play two parallel games: one public and one more personal consisting in defending one's private sphere. She weaves game theory with the thoughts of Henry David Thoreau, Michel Foucault and Ludwig Wittgenstein to explore these sort of experiments in what she calls "self-government," by which she means the ways in which individuals construct "a personal value and goal structure answering the specific needs, resources, and desires of the individual person" and adds that social relations should be based on "respect for the self-government of those with whom one disagrees [and] on a respect for, and an affection for, those with whom one agrees" (ibid, 11).If Lane's perspective sounds a little too "micro" (not to mention idiosyncratic), the notion of parallel games can also be useful to draw attention to the strategies and rhetoric of various social movements, such as the anti-globalization movement and its motto "another world is possible."

There is also the possibility suggested by evolutionary game theory that the "players" are strategies in the abstract and not persons or groups. This hypothesis can be useful in designing research

---

[6] Parallel games are a refinement of ordinary repeated games; in such models, the players are simultaneously playing two or more games and their strategies must reflect what happens in all of these games; an example would be political actors simultaneously engaged in domestic and international policy debates (e.g., Alt and Eichengreen 1989)

on institutions. Is it useful to conceptualize norms and rules as proxy actors in political interactions? This is indeed a premise adopted by some proponents of a holistic evolutionary approach to social dynamics. Cultural evolution is a minefield of sort: researchers must avoid the Charybdis of treating culture as an autonomous "force,"[7] and the Scylla of believing that moral norms, societal values and institutional rules are entirely malleable. But leaving aside the vexing question of where do norms come from, it is clear that to make sense of most political problems one must ask:

What are the Rules?

In any game, real or metaphorical, the players are following a set of rules. This can be said of politics in general, outside of a Hobbesian state of nature. One of the lessons of game theory is that a research design should make clear what the rules of the game are and what the players know about them; political science applications will usually elaborate in some detail about the institutional context in which these rules have evolved, how they are enacted, and so on but it is worth noting that the question is relevant to practically all projects, even though it may not be given sufficient attention. Large-n studies of electoral politics may seem to be focused more on structural variables (e.g., income, education, ethnicity, etc.) than on cultural norms or institutional rules, and that is of course perfectly legitimate. Applying the ceteris paribus principle, one may want to treat the legal and/or normative contexts as "givens" and examine how certain measurable key factors influence the outcome. But proponents of "mixed methodologies" have a point when they insist that large-n studies should be complemented or enriched by more focused case studies in which the specificity of certain institutions and/or national cultures and/or historical oddities can be examined more attentively.

---

[7] Rather oddly, F.A. Hayek ended up suggesting in his last book, *The Fatal Conceit* (1988) that cultural evolution is the ultimate determinant of socio-economic progress, which seems to undermine his life-long commitment to liberal individualism.

When the institutional frameworks within which the players operate are unstable (as in a time of crisis) or informal and fuzzy, the question of what they know about the rules and what they believe each one of them knows about these rules becomes vital. These are the problems which are at the root of the new field of epistemic game theory (Perea 2012). Most of that literature is fairly technical and employs mathematical concepts that even quantitatively sophisticated political scientists may be unfamiliar with. But it could be well worth the effort to at least get acquainted with the basis concepts and findings one can glean from an immersion in this literature. In particular, it illustrates the power of Bayesianism, a new direction in inferential reasoning that no political scientist can afford to ignore today.

In a nutshell, Bayesian statistical analysis is a method for updating one's prior beliefs about a given event or series of events. The underlying idea is that uncertainty can be reduced by seeking and carefully weighing the significance of new information—"wait before you jump," as the proverb goes. By extension, this provide an additional window on processes well known to political scientists such the importance of transparency, deliberation, the advantages of networking, etc.

Motivations and Rewards

All game theoretic models associate strategies with their "payoffs." A rational player chooses the strategy that, given the fact that the other players are known to do the same, gives him or her the highest payoffs. Even in formal models, there is some disagreements about how payoffs should be defined (e.g., cardinal or ordinal data). Are we talking about subjective evaluations or observable outcomes, and so on. But the question of what motivates political agents and what are the rewards they seek goes well beyond these technical issues. Indeed, the question I seem to be hinting at here is: what kind of data are used in a research design. And it could easily be retorted that every researcher knows that this question must be posed and proper answers provided. If so, game theory has nothing to teach us.

But the question I am hinting at is a little subtler! What I have in mind is the issue of whether and to what extent political agents *share* a common understanding of what the relevant facts are. An external observer gathers data. But for the players themselves, the data are the ingredients they use to make decisions. Do political scientists pay sufficient attention to the ways in which political agents treat the "data" they measure about them or about the contexts in which they operate? Could strategic players not be manipulating their audiences (including political scientists) about what they consider to be significant? And so on. Again, epistemic game theory might be worth exploring in this regard. There remains the question of what motivates the players to value the payoffs as they do. But I postpone this discussion for now because I examine the notion of self-interest in the next section.

<div align="center">Why Cooperate?</div>

Finally, there is the issue of whether the games under consideration must be thought of as zero-sum, non-zero sum but con-cooperative, or cooperative/bargaining game. Politics being the art of avoiding the Hobbesian state of war, the last two categories are more interesting for political scientists because cooperation is either possible, albeit not certain in non-zero-sum non-cooperative games, or built into the specification of the game in the case of cooperative games. Game theory opens a whole range of fascinating vistas on the ambiguity of socio-economic and political interactions by bringing home the message that there is fundamental uncertainty built-in all situations where more than one equilibrium is possible (e.g., stag hunt); that paradoxical situations arise when the rules of the game frustrate the maximization of one's own interest (e.g., prisoner's dilemma, tragedy of the commons, etc.); about the fragile emergence of cooperation in repeated games; or about the fragility of coalitions. The literature is far too vast to be summarized here (although I revisit some aspects of it in the next section). But are there lessons that can be drawn for research design outside of the boundaries defining game theoretic models? I would suggest that a crucial question that ought to be find its way in the evaluation of a

research project ought to be: what room is there in the model or the theory for uncertainty, unpredictability or unexpected consequences? This is particularly relevant to policy analysis (concerned mostly about the impact of policies) and policy studies (concerned mostly about how policies are enacted). The public policy literature, either explicitly, or implicitly, tends to advocate ways of improving upon past practices in the formulation and implementation of public policies. What I am suggesting is that these well-intentioned recommendations ought to be made guardedly because if there is one thing that game theory teaches us it is that social equilibria are transitory. Social dynamics is not always literally chaotic (although financial markets tend to display such properties); learning does occur (Skyrms 2014) and there are robust ways of bargaining for positive change (Binmore 2005). But again, cooperation is fragile, coalitions collapse, and "black swans" appear out of nowhere (Taleb 2010). So when drawing conclusions from a research project, one should consider whether they are robust enough in the face of always likely realignments, perturbations, crises, etc. This, of course, cannot be fully known in advance but policy advice or projections should never be offered without all sorts of provisos.

III. An Invitation to Epistemological Reflection: What is *Rational* Choice?

Few researchers can dispense with preparing a research design. I hope to have shown that even when working on projects seemingly unrelated to game theory, such research designs can advantageously be informed by questions and methodological puzzles typically associated with the formulation of game theoretic models. I now move to the discussion of more general philosophical issues that are not always of immediate concern to political scientists but nevertheless are implicitly built into many of the theories they use. Epistemological problems are never too far below the surface of works undertaken in all fields of political science and crop up, so to speak, in occasional controversies. Moreover, the contemporary trend toward "mixed methods" and interdisciplinary projects inevitably raises perplexing epistemological quandaries. At the core of these controversies we find dilemmas that are very familiar to

game theorists and indeed account for many of the recent developments to which I alluded in section one. I focus here on one such problem: the meaning of "rational choice" and, more broadly, the place of rationality in social ontology.

<div align="center">Rationality: What's in a Name?</div>

What rationality means or entails is a debate that cuts across many fields and disciplines, and it is not feasible to completely isolate the contribution of game theorists from those originating in rational choice more broadly defined, behavioural economics, social psychology and so on; but hereafter I try to zero in on their contributions as much as possible. Since rationality is obviously central to political discourse and political science research, from political theory to political economy to political behaviour, these questions cannot be brushed aside by political scientists. And as the debate keeps on progressing, opinions about formal theory that might have been justified a while back turn into outdated clichés.

The literature on this matter is so vast that I cannot pretend to give more than a sketchy overview but a few salient points are worth exploring. Ken Binmore (2015, 2) waxes ironic about the fact that the *Oxford Handbook of Rationality* "contains 22 essays with at least 22 opinions on how rationality can or should be construed." It is noteworthy that Binmore, who is one of the most prolific and cited game theorists, is not very impressed with this confusing array of new departures and argues that simply abandoning the well known, albeit much criticized, neoclassical definition would be ill-advised. Nevertheless, he suggests that whenever it is appropriate, looking for ways of improving upon the standard definition is justifiable and even unavoidable. What he recommends is to move cautiously, aiming for "new models of rationality" in order to sort out the wheat from the chaff, as it were. This is sensible but before letting Binmore have the last word on this issue, I want to briefly explore the array of new and admittedly somewhat confusing conceptions of rationality because my aim in this paper is to reflect on intriguing questions rather than to close the debate prematurely.

The neoclassical paradigm of rationality rests on the axioms of completeness (i.e., a rational agent can form preferences over all states of the world), transitivity (i.e., if a is preferred to b and b is preferred to c, then a is preferred to c) and independence of irrelevant alternatives (the idea that relative preferences should not be determined by the choice set[8]). Preferences are assumed to be given, usually by invoking the notion of "revealed preferences": an agent is supposed to prefer what he/she chooses (if I vote for party A, I "reveal" my preference for that party over all the others and there is no need to move further back into the psycho-social mechanisms that have led me to come up with this decision). When events are contingent rather than fixed and certain, rational agents are supposed to calculate their "expected utilities" (a concept fist proposed by John von Neumann) by assigning a probability to their value; this is useful, for instance, in comparing lotteries, it being assumed that rational agents always prefer the lottery with the highest expected utility. In game theory, the calculation of a Nash equilibrium is predicated on the acceptance of these paradigmatic concepts. But how helpful is the Nash equilibrium in political analysis? Are political agents really acting like consumers in conventional microeconomic models? Or for that matter, does marketing—with its emphasis on emotions and social cues—better approximate the actual behaviour of consumers than microeconomics? And could it not be rational in some circumstances to depart from self-interest, perhaps out of what Adam Smith called "sympathy" or some comparable other-oriented reason? In other words, should we bury *homo oeconomicus*? Interestingly, these are not questions that one may wish to raise merely as rhetorical devices for dismissing rational choice as a naïve distortion of the complexity of human affairs, or perhaps more

---

[8] If in an election where there are five candidates representing five different parties (*A, B, C, D, F*), I rank *B*'s candidate as my first choice, I shouldn't suddenly decide to vote for *C*'s candidate upon hearing that *F*'s candidate has dropped out of the race!

sinisterly as a "neoliberal" attack on the political economy of a decent society,[9] but questions that can be explored further in the light of contemporary developments in game theory, and related research projects in the social sciences (behavioural economics, experimental political science) and cognitive science.

It has become clear that "[w]hile game theory is a powerful analytic engine, hundreds of experiments show that its predictions are systematically violated" (Camerer 2004, 374).We have known for quite some time, thanks in particular to the pioneering work of Amos Tversky and Daniel Kahneman (Tversky and Kahneman 1974; 1986; Maital 2004), that in reality people (i.e., not hypothetical "representative agents") do not reason as posited by neoclassical economic theory, meaning in particular that they often do not even seem to realize what the expected utility of a contingent event is. Indeed, experimental evidence shows that the axioms listed previously are repeatedly contradicted. This can be explained by the influence of the context within which decisions are made; this the so-called "framing effect." For example, subjects will respond differently to a question about their willingness to undergo a hypothetical surgical procedure, depending on whether the risk is framed in terms of the chances of surviving it or in terms of dying from it, even though the announced percentages associated with either death or survival are strictly complementary (e.g., respectively 2% and 98%). Subjects are likely to refuse to undergo a procedure when the risk is framed in terms of a low risk of death, and to accept the risk when it is framed in terms of chances of survival. Another illustration of the same idea is given by the Allais paradox: when asked to choose between two lotteries with different values (i.e., expected utilities), one with a higher value, and which includes two or more contingent prizes, and the other of a

---

[9] This, in my view, describes S.M. Amadae's (2016) intention and I would argue that it is unjustified. Amadae's contrast between classical liberalism, which adheres to the no-harm principle, and what the various current she lumps under the label neoclassical is not without merit, even if in my opinion it is overblown. But to say that game theory is entirely consistent with these neoliberal reductionist conceptions of civil society in which narrowly self-interested actors mercilessly compete with each other completely ignores all the current developments in game theory.

lesser value but with only one prize which is non-probalistic (there is certainty that the prize will be awarded), most subjects go for the assured benefit even though it is lower than the expected utility of the probabilistic lottery.[10] Contemporary behavioral game theory starts from the premise that such findings cannot be ignored in order to try to design models with increased predictive power; the jury is still out[11] (but as afar as I am concerned here, it is gain not the actual predictive power of game theory that is its most valuable contribution but the range of questions it raises).

In the same vein, I want to mention Thomas Schelling whose game theoretic approach was developed without relying on the neoclassical principle of rationality. Schelling's players take their clues from their environment (as in the example of the pre-cellular phone era story of two individuals wishing to meet each other in New York City without knowing anything about each other's whereabouts and converging on the idea of finding each other at twelve noon in front of the clock in the middle of Grand Central station). Schelling's social ontology is itself somewhat "playful" in that it implies that epistemology and ontology are co-produced, so to speak rather than hierarchically ordered. This is only hinted at by Schelling who was not a professional philosopher but one cannot read his *Micromotives and Macrobehavior* (1978) without getting the point that individuals construct their own world through their interactions and these emergent structures in turn contextualizes their decisions.[12]

---

[10] For example, lottery L could have an expected value of (0.02 x $1,000 + 0.98 x nil =) $20 and the guaranteed prize of "lottery" L' could be $15, most people will prefer L' (guaranteed prize) over playing L.

[11] See Gale M. Lucas et al. (2015) for a sceptical view.

[12] Incidentally, the debate about the relationship between ontology and epistemology seems to have flared up again in philosophy, economics and political science; see Brian Epstein (2015) who strongly argues in favour of starting with ontological commitments—in his case a rejectrion of methodological individualism—and Robert Sugden's (2016) rebuttal; in their own ways and more or less insistentely, Christian List and Kai Spiekermann (2013), Colin Hay (2006), and Stephen Bates and Laura Jenkind (2007) also raise doubts about the wisdom of categorically prioritizing a (holistic) ontology.

What lessons can political scientists draw from these theoretical explorations, regardless of their intention to employ game theoretic models in their works? One obviously is that it is usually not realistically rational to be "rational"! But I would also suggest that before they dismiss game theory altogether, they would be well-advised to re-think rationality in the light of contemporary advances in behavioral game theory, social psychology and cognitive science. Much of these new directions of research are guided by our new understanding of the centrality of fairness in human interactions. Are expectations of fairness so rooted in the human psyche? The weight of evidence seems to lean in that direction (Bowles and Gintis 2002). If so, social or political instability is likely to result from decisions or actions that frustrate such expectations. Adam Smith already had argued in his *Theory of Moral Sentiments* that there is nothing to which members of society react more passionately than a denial of justice This is not to say that political leaders, when ready to use repressive means, can never succeed in imposing unfair policies on their citizens or some groups among them. But from an analytical point of view, students of politics or policy-making should always be attentive to how people perceive or construct norms of fairness for these are arguably what voters use to assess policies, party platforms, and their degree of allegiance to a regime. I am not re-inventing the wheel here. One of reasons why John Rawls' political philosophy has been so successful is precisely that he insisted that a just society should at the vey minimum be one in which institutional rules are fair. But I would like to suggest that if behavioral game theory and the other research programs I have allude to above are as path-breaking as we are asked to believe, then it is a mistake to prematurely conclude that Rawlsian or Dworkinian conceptions of fair procedures offer the sort of solutions we need. I would rather suggest that we both empirically minded political scientists and political theorists ought to keep on asking probing questions about what constitutes fair arrangements. Is there empirical evidence that show that Rawls' "difference principle" matters (see Michelbach et al. 2003) for a start)? How is fairness interpreted in different

cultural and political settings? How do repressive regime manage to frustrate an innate sense of fairness? Why do democracies act unfairly toward some minorities, at least some if not all the time? How can we learn to "play fair" (and does Binmore [2005] provide a workable answer to that question)? And so on.

The second lesson is that reciprocity appears to be the cement that unite all social norms. Again, I am not suggesting something very new here: form Plato to Marcel Mauss to the contemporary literature on the political economy of altruism (Fehr and Schmidt 1999; Godbout 2000; Gérard-Varet et al. 2000; Kolm and Mercier Ythier 2006), many thinkers have come to a similar conclusion, with or without the benefit of a familiarity with experimental methods! But if so, are political scientists today paying sufficient attention to it? It would seem that economists have taken the lead on this issue. What remains to be (re-)discovered about reciprocal exchanges in political life? Is the "politics of recognition" or the new emphasis on "reconciliation" a manifestation of a re-discovery of this fundamental intuition? Are there perverse consequences flowing from the design of reciprocal arrangements (e.g., some would say that "workfare" is a regress from traditional welfare support programs, etc.)? How does reciprocity factor into the new enthusiasm for a Basic Income? And so on.

Finally, we may want to ask the question: can rationality be divorced from self-interest altogether? It not irrational to follow emotional impulses or to fail to pay due care to the consequences of one's decision, because as I have shown, these biases can be fitted within revised models of rationality.[13] The search for refinements of, or alternatives to, the Nash equilibrium is motivated by such considerations. But more importantly, I would suggest narrow, egotistic self-interest is less robust than enlightened, reasonably altruistic, reciprocal self-interest, as Smith's concept of behaviour guided by an "impart spectator" already implied. But can truly disinterested, purely altruistic acts still be considered

---

[13] Very innovative work is being done in this respect at the intersection between the blogosphere and the academic community by the "rationality community"; see the blog LessWrong.

rational? Not at first sight, and I have serious doubts about this hypothesis, but interestingly Jon Elster (2009) thinks otherwise: he suggests that the final strep, which completely does away with interest, from enlightened to pure altruism does not *ipso facto destroy* rationality. Is this the case? The question is worth posing but I believe that it stretches the idea of rationality far too thinly. I tend to think that pure altruism is a-rational rather than resting at the extended margin of rationality

Conclusion

It is only since the early 1980s that game theory has gained a firm foothold in political science but it already had a long and complicated history before then and since then it has continued to branch out in many directions, both within political science and in several other disciplines. Consequently, many of the things that are said about it, to extoll its advantages or denigrate its value, are probably true of only some aspects of it or of some applications but not of other. There are some core tenets, of course, such the necessity to always model more than one individual player, to posit some sort of competitive interaction, and to logically identify plausible equilibrium points. But, as I have shown, game theory is more like a full tool kit than a single tool. This is not to say that it does not have limitations, nor that it is immune to criticisms but critics who fixate on one narrow definition of its epistemology or its uses can be taken with a proverbial grain of salt. Open minded and more or less pragmatic researchers may be justified in rejecting some variants or models without necessarily having to give up on the whole approach. It is in part to such sceptical or prudent practitioners that I have addressed my suggestions about to best interpret the potentials of game theory. But the message I have attempted to convey, namely, that the benefits one can derive from engaging with the game theory literature stem more from the depth of the questions it suggests rather from the technical answers it offers, should speak to the concerns of a great many researchers regardless of whether they do or do not make any use of game theoretic models.

By recasting game theory as a modern Socratic method, as a source of queries, and a constantly evolving paradigm, I hope to have sparked the interest of even those who are of the opinion that game theory has noting to offer them. I have shown that at the very least, it can be as a heuristic guide for working out generic research designs that can be applied to a variety of projects in many different subfields of the discipline. (After all, applying for a research grant is in itself a game in which the reward goes to the most agile strategist!) Beyond that, I hope to have also convinced some readers that game theory opens fascinating perspectives on the significance of rationality, on the potential for reconstructing political theory on the basis of a commitment to fair reciprocity, and on the merits of a conception of interest that takes on board the interests of others.

References

Alt, James E., Barry Eichengreen. (1989). "Parallel and Overlapping Games: Theory and an Application to the European Gas Trade". *Economics and Politics 1*, 119-144.

Amadae, S. M. (2016). *Prisoners of Reason: Game Theory and Neoliberal Political Economy.* New York: Cambridge University Press.

Auman, R. (2002). "Game theory in the Talmud." Research Bulletin Series on Jewish Law.

Bates, Stephen R., Laura Jenkins. (2007). "Teaching and Learning Ontology and Epistemology in Political Science." *Politics* 27: 55-63.

Binmore, K. (2005). *Natural Justice.* New York: Oxford University Press.

Binmore, K. (2015). Rationality. In S. Z. H. Peyton Young, *Game Theory with Economic Applications.* Oxford: North-Holland.

Bowles, Samuel and Herbert Gintis (2002). Homo Reciprocans. *Nature* 415: 125-128.

Brams, S. J. (2011). Game Theory and the Humanities: Bridging Two Worlds. Cambridge, MA: MIT
  Press.

Cachanosky, N. (2010). "Sontaneous Orders and Game Theory: A Comparative Conceptual Analysis."
  Revista de Instituciones, Ideas y Mercados 52: 52-88.

Camerer, C. F. (2004). "Behavioral Game Theory: Predicting Human Behavior in Strategic Situations."
  In C.F. Camerer, G.Lowenstein and M. Rabin, eds. *Advances in Behavioral Economics.*
  Princeton: Princeton University Press.

Chwe, M. S.-Y. (2013). *Jane Austen, Game Theorist.* Princeton: Princeton University Press.

De Figueiredo, R.J.P., J. Rackove, and B.R. Weingast. (2006). "Rationality, Inaccurate Mental Models,
  and Self-confirming Equilibrium: A New Understanding of the American Revolution."*Journal
  of Theoretical Politics* 18: 384-415.

Elster, J. (2009). *Le désintéressement.* Paris: Seuil.

Epstein, B. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences.* Oxford: Oxford
  University Press.

Erickson, P. (2015). *The World Game Theorists Made.* Chicago: University of Chicago Press.

Fehr, Ernst, Klaus Schmidt. (1999). "A Theory of Fairness, Competition and Cooperation". *Quarterly
  Journal of Economics* 114: 817-868.

Gérard-Varet, S.-C. Kolm, J. Mercier Ythier. (2000). *The Economics of Reciprocity, Giving and
  Altruism.* London: Macmillan.

Giocoli, N. (2003). *Modeling Rational Agents: From Interwar Economics to Early Modern Game
  Theory.* Cheltenham, UK: Edward Elgar.

Godbout, J. (2000). *The World of the Gift.* Montreal: McGill-Queens University Press.

Hay, C. (2006). "Political Ontology." In R. Goodin and C. Tilly, eds. *Oxford Handbook of Contextual Political Analysis.* Oxford: Oxford University Press.

King, G. (1995). "The Importance of Research Design in Political Science." *American Political Science Review* 89: 475-481.

Kolm, S.-C., J. Mercier Ythier. (2006). *Handbook of the Economics of Giving, Altruism and Reciprocity.* Amsterdam: Elsevier.

Lane, R. (2007). *The Game of Justice: A Theory of Individual Self-Government.* Albany, NY: SUNY Press.

Leonard, R. (2010). *Von Neumann, Morgenstern, and the Creation of Game Theory: From Chess to Social Science, 1900—1960.* Cambridge: Cambridge University Press.

List, Christian, Kai Spiekermann. (2013). "Methodological Individualism and Holism in Political Science." *American Political Science Review* 107: 629-643.

Lucas, Game M., Mathew D. McCubbins and Mark Turner. (2015). "Against Game Theory." In R. A. Scott, ed. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource.* http://scholarship.law.duke.edu/faculty_scholarship/3479.

Maital, S. (2004). "Daniel Kahneman: On Defining Rationality." *Journal of Socio-Economics* 33: 1-14.

Maschler, *Michael*, Eilon Solan and Schmuel Zamir, eds. (2013). *Game Theory*. Cambridge: Cambridge University Press.

Maynard Smith, John (1982). *Evolution and the Theory of Games*. Cambridge; Cambridge University Press.

Michelbach, Philip A., John T. Scott, Brian H. Bornstein (2003). "Doing Rawls Justice: An Experimental Study of Income Distribution Norms." *American Political Science Review* 97: 523-539.

Perea, Andres (2012). *Epistemic Game Theory: Reasoning and Choice*. New York: Cambridge University Press.

Schwartz-Shea, P. (2002). "Theorizing Gender for Experimental, Game Theory: Experiments with 'Sex Status' and 'Merit Status' in an Asymmetric Game." *Sex Roles*, 301-319.

Skyrms, B. (2004). *The Stag Hunt and the Eolution of Social Structure.* Cambridge: Cambridge University Press.

Skyrms, B. (2014). *Social Dynamics.* Oxford: Oxford University Press.

Sugden, R. (2016). "Ontology, Methodological Individualism, and the Foundations of the Social Sciences". *Journal of Economic Literature 54*, 1377-1389.

Taleb, N. N. (2010). *The Black Swan: The Impact of the Highly Improbable.* New York: Random House.

Toshkov, D. (2016). *Reasearch Design in Political Science.* London: Palgrave.

Tversky, A. and Daniel Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 1124-1131.

Tversky, Amos and Daniel Kahneman (1986). "Rational Choice and the Framing of Decisions." *Journal of Business* 59: S251-S278.

von Neumann, John and Oskar Morgenstern ([1944] 2004). *The Theory of Games and Economic Behavior.* Princeton: Princeton University Press.