STUDIA METODOLOGICZNE
NR 39 • 2019, 99–121
DOI: 10.14746/sm.2019.39.4

Piotr Swistak

The Curious Case of Formal Theory in Political Science: How Did It Emerge as a Discipline, Why Its Existence Is a Sign of a Failure, and Why the Science of Politics Is Not Possible Without It?

Abstract. American political science has evolved a subfield which is commonly referred to as *formal theory*. Political scientists identify themselves as specializing in formal theory, departments advertise faculty positions in formal theory and put together formal theory subfields that offer undergraduate and graduate curricula. The roots of the field can be traced to Thomas Hobbes. Hobbes' message, however, seems to have been utterly ignored by the social science. William Riker's second launch of "Hobbesian advice", in 1950's and 60's, proved more successful and put the field of formal theory on the map of political science. Yet, the very existence of the formal theory field can be seen as the failure of both Hobbes and Riker. There seems to be a continuing need for teaching social scientists why they should construct a proper science and how they should do it. This paper is an attempt to meet this need. I believe that the future science of politics will have to follow in the footsteps of Hobbes and Riker. And so will other social sciences. My point in the paper is not new; the way I make it, is.

Keywords: formal theory, axiomatic theory, formal theory field in political science.

1. Introduction

The "formal theory" label used by political scientists has always struck me as odd. Possibly because I could not imagine a field of "formal theory" evolving, for instance, in physics. In fact, I do not know of any other discipline of science in which a form of a theory became a label for a field of inquiry. Clearly, the very use of a qualifier "formal" implies that such theories are special, possibly rare, and that the common understanding and use of the notion of theory is different. For anyone interested in the general methodology of science, this singularity of political science may be interesting to note and to reflect on. Hence, one of my objectives is to bring this case to the attention of those who may find it worthwhile from the perspective of philosophy of science or general methodology.

My other objective is to provide an easy to read text on (formal) theory construction for social scientists. In my experience there is a widespread need for a basic reflection on the form and the benefits of a properly constructed theory in all social sciences with a possible exception of economics.[1] The absence of short and simple texts on the subject reinforces this problem. My goal is to reach those who would want to learn on their own as well as those who would want their students to better understand why this form of theory construction is important to use and necessary to accept as an integral part of any scientific endeavor.

My third, and last, objective is to engage yet a different group of scholars. Many philosophers and mathematicians will find the idea of writing about theory construction as antiquated and trivial. This is understandable. Yet, many in this group work within well formulated theories and rarely, if ever, face a problem of theory construction. Even fewer would ever deal with a case in which they observe an empirical phenomenon and have to formalize it from scratch. This is a very different problem than working with deductions and one that can be sufficiently difficult and fascinating to attract mathematicians like John von Neumann to the social sciences. Still, I know of a good number of mathematicians and other scientists who seem to dismiss social science as a domain where any interesting deductive work has been or can be done. Famously, Stanisław

[1] In my experience it is very common for students graduating with a political science degree not to have seen a formal theory nor to have reflected on the benefits of having one. This is less frequent, though still not uncommon, among graduate students and faculty members in political science. The same can be said about sociologists, psychologists and anthropologists. Economics is the only exception since it routinely exposes students to formal theories, like preference theory and expected utility theory, that are fundamental to modern economic theory. These claims derive from my many interactions with students and faculty from all disciplines of social science associated with dozens of universities worldwide. Barring a close-to-zero probability of being exposed to a very skewed sample, my observation should have a dose of generality to it.

Ulam has once challenged Paul Samuelson (Samuelson 1969, p.9) to name "one proposition in all of the social sciences which is both true and non-trivial." Some thirty years later Samuelson provided an answer to this challenge.[2] We are less likely, perhaps, to witness such a challenge today, since the underlying attitude towards economics has changed, though I think a challenge of naming a non-trivial theory in political science, or sociology, is as likely to come from a mathematician today as it was back when Ulam has posed his challenge to Samuelson. For those who cannot name a "true and non-trivial proposition" in political science I would like to offer a few very simple examples.

In short, regardless of the reader's academic discipline this paper may have something useful to offer to everyone. The value for political scientists may be in realizing the significance of the proper formulation of a problem. I hope they will see that "formal theory" is more than a subfield of political science, or a short-lived fad, it is rather the only proper way to go about theory construction if we ever hope to turn political science into a cumulative body of knowledge. This point should also be useful for other social scientists including economists some of whom get frustrated with dismal predictive power of "formal" theories in economics and turn, unwisely so, against the very scientific method used to construct these theories. Some mathematicians may take my bait and choose to look up my answers to the Ulam challenge. Perhaps they will start constructing, and solving, their own puzzles of human behavior. Finally, anyone with interest in the philosophy of science, or general methodology, may find the very existence of the formal field in political science, and some historical details behind it, to be an interesting and useful anecdote with some important lessons to learn or to teach.

I begin the paper by recalling Hobbes' advice from *Leviathan* that, ironically, was used by Wilder, an influential mathematician in 1950s, to open a chapter on axiomatic theory. The case of a prominent mathematician who brings up Hobbes whose advice was utterly ignored by political science all the while

[2] Samuelson later wrote about Ulam's testing question: "This was a test I always failed. By now, some thirty years later (…) an appropriate answer occurs to me: The Ricardian theory of comparative advantage." (Samuelson 1969, p.9)

Leviathan remained a required reading for all in the discipline[3] makes for an anecdote that was difficult for me to pass. More importantly, though, I find it important for the social scientists to remember that the proper form of a theory has been argued for and pointed to by Hobbes centuries before it started being used in the modern social science. Someday, perhaps, Hobbes' main legacy would be seen as that of a protagonist of a scientific social science. Possibly this was also on Wilder's mind when he tuned to Hobbes for an opening quote.

Hobbes' advice on theory construction fell on deaf ears among social scientists and it took some three hundred years to change that. The change came through some major developments in economics; in political science it was largely driven by the efforts of William Riker and the so called Rochester School. I will later describe the intellectual roots and the significance of the Rochester School for modern political science. While Riker and his students have effectively changed modern political science the very existence of the "formal theory" field proves that the main massage of Hobbes and the Rochester School is still lost on the discipline. The day this message settles in there will be no need to single out formal theory lest make it a special field of political science. While nearly all political scientists have heard of Prisoners' Dilemma and many can write it down and explain, only a handful would fully understand the details of the model including the properties of payoffs.[4] People use products of a proper theory construction but do not understand their basic properties. The very purpose of a proper form of science is completely lost on them. For this reason, in the last part of the paper I will go back to Hobbes' point one more time. To make the point obvious I will use some of the simplest examples I could have

[3] Hobbes' *Leviathan* is one of the most important works in the history of social science and one of fundamental importance for political science. The reason, however, is the importance of Hobbes' explanation of the emergence of state rather than his singular effort to follow Euclid's theory construction. A short comment preceding excerpts from *Leviathan* in the *Princeton Readings in Political Thought* (Cohen 1996) are typical in that regard.

[4] I can name scores in the discipline who thought that using identical payoffs for different players means that we assume that they have identical utilities and requires interpersonal utility comparisons, or those who thought that negative payoffs signify displeasure while positive pleasure, or finally those whom I have asked to tell me what is the difference between ordinal and cardinal payoff only to find out they have never heard of either. Some people in this set are prominent scholars, many use games in their research. This, by no measure, can indicate a success of Hobbes or Riker.

thought about. Even if I fail to convince anyone to change the ways they practice social science I hope to be convincing enough to make them see that the standard practice of social science cannot possibly be a long term equilibrium. It may take decades or centuries but at some point Hobbes will not only get his message through but will effectively change the social science.

2. The message of Hobbes that social science chose to ignore

Perhaps the most striking aspect of formal theories in the social sciences is their ubiquitous absence rather than occasional presence. For most social scientists a theory corresponds, roughly, to a statement "x affects y" like in "the level of education affects the level of income." The relation between x and y would be typically assumed to be of the form $y = bx + a$, and a and b are parameters estimated from the data. The more "complex" theories would assume that y is affected by more variables x, x', x'' etc, albeit in the same linear fashion: We won't see anyone conjecture $y = \int x^3 e^{-2^{(x-3)}} dx$, for instance, since how could we possibly come up with this functional form if we don't have a set of more fundamental principles, a theory, to derive it from? Most statistical tools require that we specify a function, or a class of functions, before we do any data analysis. The functions used in such statistical analyses are commonly understood as theories. Linear functions of the general form have been, by far, the most common type of "theories" generated in the social science. This literally is what most social scientists understand by "theory." It is embarrassingly uncommon for them to think of a theory the way their colleagues in science think about theories, that is, as built of undefined terms (the elementary building blocks), assumptions (axioms) defining their properties, and deductive consequences (theorems) of these assumptions. For this reason, I find it perversely amusing to see Raymond Wilder, a leading mathematician of his days, opening an essay on theory construction with a quote from Thomas Hobbes' "Leviathan" originally published in 1651:

> "The errors of definitions multiply themselves according as the reckoning proceeds; and lead men into absurdities, which at last they see but cannot avoid, without reckoning anew from the beginning."[5]

[5] Hobbes (1651, 1981 edition).

A pure mathematician using a quote from a father of political science and one of the great classics of the social science to tag an explanation of the modern form of scientific theory presents an awkward dilemma for the social science: Has Hobbes' advice, realized with a great success by mathematics and hard science, been blatantly ignored by most of the mainstream social theory for centuries, and all that while "Leviathan" has never stopped to be one of the most widely assigned and read pieces of social science? Well, yes. And it puts social sciences in an embarrassing situation, especially that Hobbes' advice is hardly controversial.

Hobbes' point, to put it in the crudest way, is that science[6] cannot be built on ill defined, or non-defined, concepts. This sounds rather obvious. If we make inferences using concepts that we don't understand—and how can we understand something that is ill defined? —we can only make a bad problem worse.[7] For sure we won't be able to explain these concepts to others, which is to say that if we don't know what we are talking about, it would be unreasonable to expect someone else to understand it. The first problem leads to absurd conclusions, the second prevents an accumulation of knowledge; either and both make the development of science impossible. This is a general observation and one that should apply to any endeavor aspiring to be called "science."

Hobbes might have felt uneasy about preaching the obvious as well. After all, he was merely repeating points that have been made twenty centuries ago by Aristotle and then implemented by Euclid in "Elements." "Elements" were, in fact, Hobbes' blueprint for constructing a scientific argument, so much so that he called it "the only science that it hath pleased God hitherto to bestow on mankind."[8] Hobbes organization of "Leviathan" is meant to

[6] I use the term "science" in its most general meaning, not one that merely restricts it to "physical science."

[7] In Hobbes' words "(…) a man that seeketh precise truth, had need to remember what every name he uses stands for; and to place it accordingly; or else he will find himself entangled in words, as a bird in lime-twigss; the more he struggles, the more belimed."

[8] Indeed, Euclid's "Elements" have been the only axiomatic theory in existence for many centuries. The modern form of axiomatic theory is typically credited to Hilbert and his 1899 axiomatization of Euclidean geometry. For a fascinating account of the history Hobbes accidental inspiration by "Elements" see Macpherson (1985.)

follow the construction of "Elements."[9] In fact, "Of Man," the first book of "Leviathan," is not only an implementation of a properly constructed theory but it also constitutes a proof that science—and indeed any reasonable form of human communication—cannot be practiced in any other way. Hobbes' implementation of Euclid's blueprint was so good that some three and a half centuries later "Leviathan" still stands out in the social sciences—if we clean up the antiquated language and shrink the long sentences—as an example of clarity and rigor. This is as much a praise of Hobbes as it is a criticism of the social sciences.

The fact that a scientific theory has to be constructed in a proper way was obvious to Hobbes and by 1950's it was obvious to anyone working in mathematics and in physical sciences and to a growing number of economists especially those with background in mathematics or in physics[10]. This must have made Raymond Wilder[11] conclude in 1952 that this form of theory construction "is acknowledged now as a fundamental part of the scientific method in every realm of human endeavor" (p.1622.) Wilder has been wrong when he wrote these words and he would have been wrong today. Most social scientists practice "science" that lacks "a fundamental part of the scientific method."

Wilder was right on a different count though. The tide was about to turn. In 1950's the change has already been sweeping economics and soon it was about to hit political science.

[9] It has been pointed to me by one of the reviewers that Hobbes had contacts with Descartes and Galileo, hence Euclid was hardly the only source of his convictions. My objective here is not, however, to provide a comprehensive historical account.

[10] By now to be accepted into a major graduate program in economics, an undergraduate degree in mathematics, or a coursework that comes close to getting one, is almost a prerequisite.

[11] Raymond Louis Wilder (1896-1982) was a leading mathematician of his times. His 1952 "Introduction to the Foundations of Mathematics" includes a chapter on theory construction, "The Axiomatic Method." It was later reprinted as an independent essay in "The World of Mathematics." More than half a century after its first publication "The Axiomatic Method" remains an excellent introductory writing on the subject. (Wilder spent most of his academic career in the Department of Mathematics at the University of Michigan, Ann Arbor; he was a member of the National Academy of Sciences and served as the President of the American Mathematical Society and the Mathematical Association of America.)

3. William Riker and the emergence of the formal theory field in political science

While the early uses of formal theory in political science can be traced to Black (1948), Downs (1957), Shapley and Shubik (1954) and a few others, there is no doubt that a subfield of political science that goes under the name of formal theory, rational choice theory or positive political theory (these names are often seen as interchangeable) is a work of one man. His name is William Riker. For anyone who witnessed the emergence and the growth of this field, "formal theory" feels largely synonymous with William Riker. Riker, in a sense, succeeded at implementing Hobbes' advice for political science.

Critical for Riker's success was the timing of his efforts. By 1950's mathematicians and logicians have refined the notion of a formal theory, which in these disciplines is better known as an axiomatic theory, and it became a standard of theory construction for an increasing number of scientists including a growing number of economists. In 1944 game theory has been launched by von Neumann and Morgenstern's "Theory of Games and Economic Behavior." It quickly became a fad among mathematicians and a small, but influential, group of social scientists. John Nash and Lloyd Shapley were among the earliest followers of the fad. In 1950's they were both graduate students in mathematics at Princeton and they both chose to work on game theory under Albert Tucker. Their doctoral theses were to lay the foundations of non-cooperative and cooperative game theory. (Their solution concepts are known as Nash equilibrium and Shapley value. John Nash and Lloyd Shapley were both awarded Nobel Prizes for these contributions.) About the same time Albert Tucker, came up with a name and an interpretation of a game that was brought to his attention by Merrill Flood and Melvin Dresher.[12] The name he proposed was Prisoner's Dilemma. Prisoner's Dilemma went on to become the single most important and most influential game in game theory. Finally, 1951 saw the publication of Kenneth Arrow's "Social Choice and Individual Values" which contained what is now known as Arrow's Theorem. The result

[12] This story has been recounted in many writings. One of the more recent ones is (Holt 2004).

gave rise to social choice theory and became an important factor in awarding Arrow a Nobel Prize. These developments have proved to be transformative for all social sciences but they were particularly consequential for the science of politics. At a closer look the core of most of them was par excellence political.

Game theory is a general theory of interactions and as such it is relevant for all social sciences. Of the two branches of game theory, cooperative and non-cooperative, cooperative game theory was created as a theory of coalition formation. Coalition formation is, of course, the domain of political science which made cooperative game theory impossible to ignore by political scientists. The relevance of non-cooperative games soon became equally obvious. The Prisoner's Dilemma hit at the very core of political science by providing a general model for a class of problems that require solution by government. The game's equilibrium came to be seen as the argument for the necessity of state and a proof that political science cannot be subsumed by economics. Jon Elster made this point explicitly clear by defining politics as "the study of ways of transcending the Prisoner's Dilemma." (Elster 1976, p.249) In this sense Prisoner's Dilemma became a label for the very essence of political science.

Another transformative idea came from a different line of research. Central to this research was the work of Kenneth Arrow. Arrow (Arrow 1951, 2nd ed. 1963) set out to formally define the core properties of a democratic choice and proved that they cannot all be satisfied. One way to think about Arrow's result is as a formal theory of any voting rule that we would want to use in elections. Before we start arguing if we want to use a simple majority rule, or a plurality rule or approval voting or, perhaps, the American electoral system, we would want to sit down and specify basic properties of any acceptable democratic voting rule. For example, we would want a rule that treats all voters equally (all votes count the same) and one which when used in a two-candidate election where every voter prefers one candidate over the other, would name the preferred candidate the winner. The way I always explain Arrow's Theorem is by stating such—overtly obvious—properties of a democratic choice and ask students to think of them as articles of a constitution. I have yet to find a single individual who would not support it. As it turns out, democratic rule of choice, as defined by the articles our constitution, is a concept with an empty connotation. Arrow has proved that "articles" of this "constitution" are inconsistent. This result, known as Arrow's Theorem, is considered by many to be the most important finding in the modern political science. This is one of the best lessons Hobbes could have hoped for. If Arrow did not construct an axiomatic (formal) theory of choice, we would have never known that what we may want in democracy is not possible to have.

It is important for me to note that my description of the developments of 1940's and 50's is not indicative of how they were seen, or recognized, in political science. All these results came from mathematics and/or economics and very few political scientists knew about them or had sufficient mathematical training to understand this work.[13] Riker who had considerable training in economics was an exception. Not only was he familiar with this research but he clearly understood its importance for political science. He was also aware that the timing to change political science could not have been better and through his own research he began to advance the new paradigm (Riker, 1963, 1982; Riker and Ordeshook, 1973). The scientific form of this paradigm, widely seen as employing formal theory construction, became central to Riker's brand of political science. When in 1960's the University of Rochester was looking for faculty to build analytically rigorous social science departments, Riker was an obvious choice for political science. Backed by a very generous support of the university Riker built a department that made Rochester a famous brand and began a transformation that created the field of political science commonly known as formal theory.

Riker's students, students of Riker's students, and countless followers populate now major departments of political science offering curricula in formal theory at undergraduate and graduate levels. Political science seems to have undergone a permanent and irreversible change. And yet, I am not sure if the scope and the pace of the change would have been seen as satisfactory by Hobbes, or by Wilder.

The bulk of political science still does not use the proper theory construction. The very use of the "formal theory" label suggests that most "theories" are not constructed that way; to put it bluntly, it suggests that most of political

[13] Private communication with William Riker and Peter Orheshook.

theory is not really scientific. The inertia of the discipline and the resistance to change have proved to be considerable. For this reason, papers like this one will be written for a foreseeable future and arguments will be given for and against the use of formal theory construction in social science. The day when symbolic notation, like the one used in the formulation of preference theory, will cover blackboards of classes in political science the way they do now in physics, is still very distant.

For this reason, anyone teaching formal theory in political science finds it imperative to explain to students why formal political science looks so much different than what they see elsewhere in the discipline. I often envy physicists who don't have to justify the use of mathematics or other properly formalized constructs since everyone in physics knows that this is the proper form of science. In the next two sections I offer two arguments which I often use to convince students of politics about the necessity of the proper theory construction.

What I have consistently found to work best as convincing explanations are arguments that use short and simple examples. Each of the next two section is built around one such example.

4. An example that explains the necessity of formal political science

Sometime during the winter break of 2006/07 Nicholas Grossman, who was then a graduate student at the University of Maryland, stopped by my office to tell me about a peculiar finding he had overheard on the radio when driving from Connecticut to Maryland. The data concerned a study involving voters' preferences over three Democratic front runners, John Edwards, Barack Obama and Hilary Clinton, who were to compete for the Democratic nomination in the US presidential election of 2008. The puzzle behind the peculiar results can be shortly described as follows.

Subjects in the study were asked to consider two candidates, Edwards and Clinton, for instance, and pick the one they would vote for. They were then asked to do the same for Edwards and Obama, and Clinton and Obama. In yet another question they were asked to vote for one from the set of Edwards, Obama and Clinton. Interestingly, Edwards was picked by a substantial majority against Clinton and against Obama and yet when subjects were picking one out of three, Edwards ended up with the lowest percentage of votes. This has struck Grossman as an interesting, and a possibly odd, result. Indeed, the data seems counterintuitive enough to warrant a closer examination.

To make the puzzle more explicit let's conjecture specific distributions of choices. Assume, for instance, that Edwards was chosen by 60% over Clinton and by 60% over Obama; for completeness, let's also assume than 60% chose Clinton over Obama. Suppose, moreover, that when asked to pick one out of three, 38% chose Clinton, 32% Obama, and 30% Edwards. What, if anything, might be wrong with this outcome?

If, indeed, there is something wrong with the assumed distributions, the problem must originate in subjects' preferences over the candidates. If, for instance, a subject does not really know any of the three candidates but feels compelled to answer the questions, he may well make his choices at random. But choosing at random means that a person who chose Edwards over Clinton may choose Clinton over Edwards when asked about his preferences again. His choice in one instance is unrelated to his choice in the other. Is it an effect of such random choices that strikes us as "odd" in the data? Do we, indeed, intuitively expect some "consistency of choices" and this consistency is violated by the distribution of votes cast by the subjects?

If the lack of consistency of choices constitutes the essence of the puzzle, then the solution would require that we define the conditions of consistency and then prove that they are violated by the data.

And so, let's first consider the properties of choices that would agree with our intuition of consistency. If, for instance, a decision maker chooses x over y, basic consistency, as I have already noted, would require that he does not choose y over x when we ask him again. Another similarly straightforward aspect of consistency would have a preference relation behave similar to a "greater than" relation on numbers: if a subject prefers x over y and y over z then he should also prefer x over z. Yet another set of intuitions would apply to cases in which a decision maker is indifferent between alternatives. Here our intuition would suggest that indifference is a form of equivalence. Thus, for instance, a "consistent" indifference relation should not be affected by the order in which alternatives are presented: if a person is indifferent between x and y, he should also be indifferent between y and x. In addition, anyone who is indifferent between x and y and y and z should also be indifferent between x and z. One final condition would assume that subjects know the alternatives and, hence, can choose between them. More specifically, we will assume that when a decision maker compares two alternatives he would either prefer one over the other or be indifferent between them.

If we agree to define our intuitions the way I have suggested, we are all set to construct a "formal theory" of individual decision making. In the formulation I will use the following notation: $D = \{x, y, z, \ldots\}$ denotes a finite set of alternatives; $\prec$ denotes a binary preference relation (strict), $\sim$ denotes a binary indifference relation, $\neg$ denotes "it is not true that," $\forall$ denotes "for all," $\exists$ denotes "there exists," $\in$ denotes "belongs to," $\Rightarrow$ denotes "implies," $\vee$ denotes "or," and & denotes "and." A formal system $\langle D, \prec, \sim \rangle$ which satisfies the following five axioms is called *theory of preferences*.

A1. (COMPARABILITY) $\forall x,y \in D: (x \prec y \vee y \prec x \vee x \sim y)$.
A2. (ASYMMETRY OF $\prec$) $\forall x,y \in D: (x \prec y \Rightarrow \neg (y \prec x))$.
A3. (TRANSITIVITY OF $\prec$) $\forall x,y,z \in D: ((x \prec y \;\&\; y \prec z) \Rightarrow x \prec z)$.
A4. (SYMMETRY OF $\sim$) $\forall x,y \in D: (x \sim y \Rightarrow y \sim x)$.
A5. (TRANSITIVITY OF $\sim$) $\forall x,y,z \in D: ((x \sim y \;\&\; y \sim z) \Rightarrow x \sim z)$.

The theory we have just formulated, known as theory of preferences, constitutes the most general concept of rationality in the social sciences. For this reason alone, it may well be the most important formal theory of all. But there is another important reason for its significance. A properly constructed theory of choice made it possible for the science of choice to grow. The expected utility theory, formulated by von Neumann and Morgenstern (1945), is a generalization of the preference theory to the domain of choices that involve lotteries, i.e., probability distributions over alternatives in the domain. Game theory, finally, is a generalization of theories of individual decision making, i.e., theory of preferences and the expected utility theory, to cases in which the outcome is affected not only by the action of a decision maker but also by the actions of others. These connections are important to note because without a proper formulation of a theory of preferences the path that links it with game theory would not have been possible. For any science to progress, it must be cumulative and this cannot be done outside of properly constructed theories.

Now that we have defined a consistent choice behavior, better known as rational choice, we can return to the data and ask if there is any evidence that any of the axioms has been violated. Given the design of the study the only axiom that can be violated is the transitivity of the preference relation. Does, then, 60% choosing Edwards over Clinton, 60% choosing Edwards over Obama, 60% choosing Clinton over Obama, and 38% choosing Clinton, 32% Obama and 30% Edwards when asked to pick one out of three indicate that some of the subjects must have had intransitive preferences?

The answer is given by the following example. Consider a group with the following (transitive) preferences over Edwards (E), Clinton (C) and Obama (O): 8% prefer E over O over C or EOC, in short, 22% ECO, 30% CEO, 8 COE, 30% OEC and 2% OCE. It is easy to check that this distribution of preferences results in the conjectured data. And so, the intuitive feeling that something must be wrong with Edwards winning by considerable majority over Clinton and over Obama and then coming last when voters were to pick one out of three candidates was plain wrong.

This, of course, would not come as a surprise to anyone working with deductions. After all the point of any deductive science, i.e., a formal theory as a political scientist would say, "(…) is to start with something so simple as not to seem worth stating, and to end with something so paradoxical that no one will believe it." (Russell 2009) Implicit in Bertrand Russell's words is an advice that working with proper science we are better off putting our intuitions aside since what we are heading for may be negatively correlated with them. This seems to be the case with the example of the voting study.

Another lesson to be learned from the example is that formal theory, as the proper form of science, is necessary simply because human brain cannot properly process tasks of certain complexity. Just like we do not attempt to multiply two five digit numbers in our head we should not attempt to make scientific inferences outside of a properly constructed formal theory. As Richard Feynman has notably put it: "Science is a way of trying not to fool yourself.

The first principle is that you must not fool yourself, and you are the easiest person to fool." (Coyne, 2009)[14] While we widely recognize and acknowledge our arithmetical deficiencies and wisely use calculators when multiplying five digit numbers, we are not nearly as humble when it comes to our ability of practicing social science. If we were, the formalization of the social science would have started with Hobbes.

5. A trap for a barefoot empiricist

Barefoot empiricism is a common accusation used against many social scientists. How can we, in a simple way, see that sheer data analysis may not make sense without a theoretical reflection? It would be good to see it using the simplest theory possible. If sheer data analysis without any analytical work turns out to be unreasonable in the simplest instance then, obviously, it can only get worse as we move to more complex cases.

Imagine that John, a student of international relations, has been studying networks of military alliances. A pure empiricist, John gave no thought to theoretical properties of such networks and set out to search for interesting empirical relations in the data. The data spanned many decades and for each point in time and every two countries it could tell John whether the pair was in military alliance or not. For exploratory purposes John picked an arbitrary year and started looking at the data. To aid his exploration he would occasionally draw a graph, like these in Figure 1, with countries represented by vertices and alliance relation by edges; absence of an edge between two vertices means that the two countries were not allied.

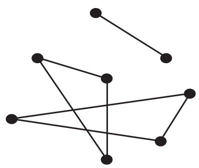

Figure 1

[14] I want to thank Jon Bendor for bringing this quote to my attention.

Imagine now two unrelated scenarios of what could have been the outcome of John's analyses.

Situation 1. After a careful examination of alliances in all groups of countries, and at all points in time, John discovered a remarkable regularity: It was always the case that if *a* was allied with *b* and *b* was allied with *c* then *a* was also allied with *c*. Since this property is called transitivity in the jargon of "relational properties," we can say that John has discovered ubiquitous transitivity of the alliance relation. This regularity has struck him as one of obvious significance and since no one else had noticed it before, John felt that he made an important discovery.

Now, let's reverse the time, assume that Situation 1 has never happened and John did not notice transitivity simply because there was no transitivity in the data. Situation 1 aside, however, the data might have held other interesting regularities for John to observe. The next situation describes one such scenario.

Situation 2. Suppose that John, like many scholars of international relations, was interested in the concept of power. One possible index of country's power is its number of allies. Since more allies makes one more powerful (other factors equal), John started examining patterns in the number of allies. One thought that has crossed John's mind was to look for cases where every country has a different number of allies. In such cases countries would form a strict "power hierarchy" with the country with most allies at the top and the country with the least allies at the bottom. Interestingly, though, John has failed to find them. At all points in time and in all possible subsets of countries there were always at least two countries with the same number of allies. There has never been a set of countries with a strictly hierarchical order that John was looking for—not a single case like that across the recorded history! This observation has struck him as interesting and significant. Since no one else had discovered it before, John felt that he has made an important discovery.

Let's now revisit the two scenarios. Since all that John has done is data analysis, we now need to do some thinking for John. Let's begin with Situation 1 and the case of transitivity. If we consider all theoretically possible graphs of alliances it is obvious that some graphs have a transitive alliance relation, say the graph in Figure 1, but other graphs may not. This means that the transitivity of the alliance relation discovered by John was a genuine empirical finding.

A philosopher would say that a proposition "In all networks of alliances the alliance relation is transitive" is a synthetic proposition—one that may be either true or false depending on what we find in the data. In Situation 1 the data have rendered this proposition true.

Now, let's turn to Situation 2 and, by analogy, ask if the proposition "In all networks of alliances there are at least two countries with the same number of allies" is a synthetic proposition as well. Is it so that in some cases the proposition is true but in some it is false? If this proposition is synthetic then John's discovery, like the discovery of transitivity, would mean a genuine empirical finding. But what if it is not a synthetic proposition? In that case, given the definition of a synthetic proposition this proposition must always be true. But if it is always true, regardless of the data, then it must follow deductively from the structure of the data. Does it?

Clearly, to move ahead with the problem we need to realize the structure of the data—i.e., the underlying formal theory—and try to figure out what follows from its axioms. Fortunately, the theory is trivial, even from a barefoot empiricist's point of view and may well be the simplest theory there is. It consists of a finite set of objects, countries in our interpretation, and a single binary relation, the alliance relation α, on this set. The only assumption we make is the symmetry of this relation: for any *a* and *b*, if *a* is an ally of *b* then *b* is an ally of *a*. Our formal theory is a system $\langle D, \alpha \rangle$ with a single axiom:

A1. (SYMMETRY OF α) $\forall a,b \in D: (a \alpha b \Rightarrow b \alpha a)$.

This, of course, is a well-known theory in mathematics, a subfield of graph theory which is called a theory of undirected graphs. What is important for John, however, is what follows deductively from this simple axiom. And what follows is that his empirical observation was not "empirical" at all. No matter what system of alliances you take there will always be at least two countries with the same number of allies.

To prove this proposition let's suppose, by contradiction, that there is a system with n different countries (n>2) where all countries have different number of allies. The only possible number of allies any country can have is one of the n numbers: 0, 1, …. , n-1. Thus if all countries have different number of allies, then one of them must have 0 allies, another one 1 ally, another one 2 allies, …, and yet another one n-1 allies. But if a country has n-1 allies, it must be allied with all other countries in the system, and this means that the system cannot contain a country with 0 allies. This, however, contradicts the fact that all countries have different number of allies, and thus proves the proposition.

What John has thought to be an empirical discovery, turns out not to be an empirical discovery at all—it is a deductive consequence of the way he has defined the problem. And so the proposition "In all networks of alliances there are at least two countries with the same number of allies" is not a synthetic proposition; it is not an empirical regularity at all. Philosophers call such propositions analytic. This means that if we properly specify the concepts and the assumptions we use—construct a proper theory, that is—then this proposition follows from the axioms.

Whenever we deal with properly defined concepts, there will always be empirical findings whose status we won't know if we don't check if the observed regularity does not follow from the set of assumptions that define these concepts. In other words, unreflective empiricism is only safe when a theory is trivial: either all deductive consequences of the assumptions we make are obvious or there are no deductive consequences at all. Whenever we venture into the simplest of nontrivial theories, like the theory of alliances, we would not know if our empirical discovery is genuine or not unless we first show that it is not a deductive consequence of theory's assumptions. Atheoretical empiricism only makes sense in a world that has a transparently trivial structure.

Perhaps my example is too simple to fool everyone though I know for a fact that it works on great many people. It shows that our brains are not good with deductions. Even if we realize the form of the theory we are working with we still cannot see some very simple consequences of the assumptions we make. But if our logical skills are dismal then the very idea of pursuing inferences without properly specifying all assumptions needed—i.e., without having a formal theory—makes about as much sense as trying to execute an addition without knowing what numbers we are adding. And so, my last example goes to the core of the problem: How good are our logical powers?

6. The dismal shape of our scientific backbone

Since all valid inferences have to use logic, logic is our scientific backbone. I often come across people who think that logic is an abstract construction of philosophers and mathematicians and as such belongs with other esoteric topics that are of interest to no one but academics. They are wrong. First and foremost, logic is the way we (should) make inferences, not just in science, but in everything we do.

To see logic at work it is enough to recall a visit to a supermarket and reflect on one of the shoppers in front of us whose train of thought must have gone somewhat like this:

> "Oh, good, the cashier has scanned all my products. Why does he keep staring at me? Oh, I know. I have to pay now. OK, let's pay. How can I pay? Well, by cash, credit card, or a check. So, where is my money? It must be in a purse which is in my backpack. Let's get it and pay cash. How much do I have to pay? All right, let's see how much cash is in the purse. I have a ten, and a five, and some singles. No, I don't have eighty-four dollars. All right, let's use a credit card then. But, where are my credit cards? They can be in my purse. Are they? Let's check. No, they are not. Well, they must be in my wallet then. But my wallet can only be in my purse or in the back pocket of my pants. Since it is not in my purse, it must be in the pocket. Oh yes, it is. (…)"

In science people make inferences the same way, albeit faster.

If we can watch logic at work in a supermarket, logic must clearly be the backbone of our actions and thoughts. Since it is important to have a strong backbone we ought to take a closer look at ours.

Consider, for instance, the following dictum of Carl Sagan, a prominent astrophysicist and a popularizer of science. (Davidson, 1999, p. xv.)

> If "experts" could always be trusted to make the right moral decision, then public participation would not be necessary—but they cannot be, and so it is.

If we have read over this argument without any alarms going off, my point that our backbone is not in good shape would have been made. This is because Sagan's inference is wrong; his dictum contains a logical error. Note that the saying has the following structure: "If *p* then *q*. But we observe *not p*, thus we conclude *not q*." For Sagan's dictum to hold we would need "if *p* then *q*" implying "if *not p* then *not q*." But, it does not.

The dictum's logic is not all that more challenging than what we saw in the supermarket story and yet, if we don't clearly see what is wrong with it we would have to concede that our backbone has failed us. Well, a slip, you may say, is not a big deal. Don't we all make mistakes? Yes, we do. This is one point of the story. Another point of the story is that we would expect science to be free of such mistakes. If an engineer designs a bridge, we expect him to be careful about the validity of inferences he used in the construction. This is clearly something we would wish of anyone whose actions affect us. Even the most methodologically flamboyant social scientist will place his trust in logic when heading for a surgery. Whether we want to admit it or not, we all assume that people whose actions affect us do not make logical mistakes.

Picking on Carl Sagan is important for another reason. It is the smart people we should watch because it is their arguments that we trust most. The problem, of course, is that their brains, just like everyone else's, are not perfect. The more important problem with logical mistakes is that while we all make them; it is very seldom that we have an opportunity to recognize that we do. Here is my very own blunder. I must have had many of them, but, unlike the case below, there was no one around to point them out to me.[15] The blunder uses Wason's selection task experiment which I have later replicated dozens of times in my classes.

Suppose we have four cards and each has a letter written on one side and a number on the other. The four cards lie on a table and show A, B, 4 and 7. Which cards do you need to turn over to determine if the proposition "if A is on one side then 4 is on the other" is true? Should we open cards A and 4, or should we do something else? I am most gratified to find that a good majority of people are guessing A and 4. So did I. Some people guess A and 7 though, and they are correct. My guess of A and 4 was wrong and I would not be able to talk myself out of the trouble. The virtue of science is that whenever I am wrong anyone can prove me wrong and make me see my mistake and correct my reasoning.

Let's then go through the reasoning here. Why is opening A and 7 the right answer? For simplicity let's use p to denote our proposition "if A is on one

[15] Many years ago professor Andrzej Nowak of the University of Warsaw has shown me this puzzle and was a witness to my blunder. The puzzle is well known among psychologists as Wason selection task (Wason 1968).

side then 4 is on the other." Two things are clear. First, we don't need to open B since no matter what number we find on the other side, it has no bearing on *p* being true or false, *p* remains true regardless. Second, we have to open A to check if 4 is indeed on the other side; if it not then *p* is, clearly, false. Now, how about 4 and 7? Let's consider 4 first. If we open 4 and the letter on the other side is A then *p* is true for this card. But if the letter is not A then *p* is also true—*p* only claims that "if A is on one side then…" No matter what is on the other side of the 4 card *p* remains true. If we want to know if *p* is true or false opening this card is useless. The same, however, is not true about the card with 7: If we find A on the other side it renders *p* false but if the letter is different it renders *p* true. To conclude, if we want to establish if *p* is true or false we have to open cards A and 7 only.

The most compelling aspect of Wason's example is that it uses the simplest of logical structures—a single implication. If we can show that most people get it wrong at this level, I don't think we can get a more persuasive argument for constructing science in a proper form.

7. Conclusion

If our brain's software which we use for reasoning contains glitches, as Wason's selection task experiment would strongly suggest, then we face double peril when practicing science. First, we don't clearly see deductive consequences of the assumptions we tacitly or overtly impose on a problem. And second, and more disturbing perhaps, the framework which we use to make inferences is itself ridden with errors which make us see some true propositions as false or false propositions as true. When practicing science that involves any nontrivial inferences we don't really have a choice—we have to do it in properly constructed theories. In political science they are called formal theories.

The day when political scientists will stop using the qualifier "formal" will be the day when the vision of Hobbes would prevail. As long as there is a need to identify "formal theory" as a specific subfield of political science neither Hobbes, nor Wilder, nor Riker would find the status quo satisfactory.

In conclusion I should note that it was not my objective to come up with an overview of the formal theory field or its comprehensive history. In my judgment, a different vision of the paper was more useful and this is the vision I have implemented. My reasons and objectives are explained in the introduction.

One reason I did not to pursue certain topics is that others have done that before and I found no good reason to either repeat their points or try to improve on their work. In closing it may be useful to mention a few of these sources. Morris Fiorina (1975) wrote an early overview aimed at explaining to a general audience the role and the reasons for using formal theories in political science. It is a very good paper and I would recommend it to anyone interested in the subject. Paper by Amadae and Bueno de Mesquita (1999) provides an interesting and a comprehensive historical account of the, so-called, Rochester School in political science. Lalman, Oppenheimer and Swistak (1993) compiled a comprehensive overview of the field of formal theory in political science. Shepsle's "Analytical Politics" (2010) is the best undergraduate textbook in the discipline. And, finally, anyone interested in the most recent advances in the field should consult the book of Humphreys (2017.) It provides a wealth of specific applications.

References

Amadae, Sonja M., and Bruce Bueno de Mesquita. 1999. "The Rochester School: The origins of positive political theory." *Annual Review of Political Science* 269-295.

Arrow, Kenneth J. 1951, 2nd ed. 1963. *Social Choice and Individual Values.* Yale University Press.

Black, Duncan. 1948. "On the Rationale of Group Decision Making." *Journal of Political Economy* 23-34.

Cohen, Mitchell, and Nicole Fermon, editors. 1996. *Princeton Readings in Political Thought.* Princeton, New Jersey: Princeton University Press

Coyne, Jerry. 2009. "Seeing and Believing." *New Republic*, February 4: 32-41.

Davidson, Keay. 1999. *Carl Sagan: A Life.* New York: John Wiley & Sons.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

Elster, Jon. 1976. "Some Conceptual Problems in Political Theory." In *Power and Political Theory: Some European Perspectives*, by Brian Barry (Ed.). London: Wiley.

Fiorina, Morris. 1975. “Formal Models in Political Science.” *American Journal of Political Science* 133-159.
Hobbes, Thomas. 1651 (reprinted 1981). *Leviathan.* London: Penguin Classics.
Holt, Charles A., and Alvin E. Roth. 2004. “The Nash Equilibrium: A Perspective.” *Proceedings of the National Academy of Sciences* 3999-4002.
Humphreys, Macartan. 2017. *Political Games.* New York: Norton.
Lalman, David, Oppenheimer, Joe and Piotr Swistak. 1993. “Formal Rational Choice Theory: A Cumulative Science of Politics.” In *Political Science: The State of the Discipline II*, by Ada Finifter (Ed.) , 77-104. Washington DC: American Political Science Association.
Macpherson, Crawford. 1985. “Introduction.” In *Leviathan*, by Thomas Hobbes, 9-63. Penguin.
Oaksford, Mike and Mike Chater. 2009. “Précis of Bayesian rationality: The probabilistic approach to human reasoning.” *Behavioral and Brain Sciences* 69-84.
Riker, William and Peter Ordeshook. 1973. *An Introduction to Positive Political Theory.* Englewood Cliffs, NJ: Prentice-Hall.
Riker, William. 1982. *Liberalism Against Populism.* San Francisco: Freeman.
Riker, William. 1963. *The Theory of Political Coalitions.* New York: Yale University Press.
Russell, Bertrand. 2009. *The Philosophy of Logical Atomism.* Routledge.
Samuelson, Paul A. 1969. “Presidential Address The Way of an Economist.” In *International Economic Relations*, by Paul A. Samuelson (Ed.) , 1-11. London: Palgrave Macmillan.
Shapley, Lloyd and Martin Shubik. 1954. “A method for evaluating the distribution of power in a committee system.” *American Political Science Review* 787-792.
Shepsle, Kenneth. 2010. *Analyzing Politics.* New York: Norton.
Von Neumann, John and Oskar Morgenstern. 1945. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.
Wason, Peter. 1968. “Reasoning about a rule.” *Quarterly Journal of Experimental Psychology* 273-281.

Piotr Swistak,
Department of Government and Politics,
and Applied Mathematics, Statistics and Scientific Computation Program,
University of Maryland, College Park, Maryland, 20742, USA
e-mail: pswistak@umd.edu