

[Work in progress, please do not quote without explicit authorization from the author]

Visualizing Data: A Theoretical Rationale\*  
(*First draft: April, 20, 2001*)

**Itai Sened**<sup>†</sup>

---

\* Prepared for presentation at the Research Workshop in Applied Statistics at the Center for Basic Research in the Social Sciences at Harvard University. April, 25, 2001. Thanks are due to John Sprague and Andrew Martin for useful comments on earlier drafts. Remaining errors are to the credit of the author alone.

<sup>†</sup> Itai Sened is an associate professor of Political Science at Washington University in St. Louis.

## ***1. Introduction***

The way in which theory and data are jointly used in scientific research is far from trivial. Presumably, theory advances our knowledge of the logic underlying the phenomena we study and data is used to test our theoretical predictions. Yet, we all know that treating the relationship between empirical evidence and theory in these terms is an oversimplification. We all vaguely remember that we never ‘prove’ any hypothesis but ‘fail to disprove it’. This roundabout language reminds us that the use of empirical evidence in evaluating the merits of theoretical knowledge is not as straightforward as it may seem. Unfortunately, this common knowledge is often set aside when we continue, in our routine conduct of social science, to use language that suggests that what we derive theoretical predictions and bring forth empirical data to test them.

The recent advent of visualizing tools, I argue, brings about a revolutionary new way to combine theory and empirical evidence in scientific research that overcomes the need to refer to the naïve notion of ‘hypothesis testing’ as the central role of empirical evidence in scientific research. More generally, this essay encourages this new methodological approach by arguing that from the standpoint of philosophy of science it may be the only justifiable approach in conducting empirical research. It argues a fundamental and new philosophical rationale for using visualizing techniques to bridge the inherent gap between empirical evidence and theoretical knowledge and provide a sound and fresh alternative to ‘naïve hypothesis testing,’ as the central logic of the philosophy and practice of social sciences in general and political science in particular.

The following section discusses the inherent difficulty in combining theoretical knowledge with empirical evidence in conducting scientific research. The third section explains what is wrong with ‘dogmatic falsificationism,’ the prevalent philosophical

justification behind the current practice of reliance on ‘naïve hypothesis testing’ in the social sciences. Section four reviews Lakatos’ (1978a) philosophy of science as a reasonable way around this difficulty. Section five augments Lakatos’ argument, using a forgotten insight from Wittgenstein’s view of the problem. To my knowledge, Wittgenstein was the first to suggest that the relationship between theoretical knowledge and empirical evidence may be regarded as an attempt at matching pictures, one structural and theoretical in nature, the other empirical and ‘semantically (or substantively) detailed’ in nature. Section six provide three example starting from a simple linear model, through an event count model and finally a Bayesian analysis of voting behavior in Israeli election in the last two decades.

Rather than preach to the skeptic or even try hard to convince them, the purpose of this essay is to provide those who find some common good sense in conducting research in this way with a coherent theoretical argument to justify the use of visual tools in scientific research. This approach is often rejected on the grounds that it does not provide us with as clear criteria as more traditional statistical tools to evaluate, reject or ‘fail to disprove’ scientific hypotheses. The argument of this essay is simply that: science does not prove or disprove anything. Good science provides useful tools to help us understand aspects of our lives that we find worth studying. To do so it must demonstrate that the theoretical knowledge we accumulate corresponds somehow to the phenomenal world we seek to explain. Traditional statistical analysis tends to reduce this illustration to significance tests and characterization of the functional forms of multivariate relations and correlations that may best summarize the underlying structure of the empirical data at hand. Visual tools provide richer pictures of reality. In doing so they encourage the scientist to develop more intricate and elaborate theoretical explanations and, in the end, force the scientist to work

harder at proving the ‘goodness of fit’ between the structure of the theoretical argument (or ‘picture’) and the structure of the empirical data (or our ‘picture of reality’).

We are obligated to develop precise measures of this ‘goodness of fit,’ but just as a picture is often worth a thousand words, an ‘inter-ocular traumatic impact’ is often more reliable and convincing than many technical measures of goodness of fit. More importantly, the argument of this essay is that visualizing data helps us make the long overdue transition from naïve hypothesis testing derived from ‘dogmatic falsificationism’ to more current views in philosophy of science that reject dogmatic falsificationism and shift the emphasis from the search for ‘truth’, hypothesis testing and critical tests, to consistency and proliferation as the central criteria for scientific progress in general and the role of empirical research within the scientific enterprise in particular.

## ***2. The Challenge in Combining Theory and Data in Scientific Research***

At the beginning of the 17<sup>th</sup> century Francis Bacon, an influential judge, statesman, alchemist and philosopher proposed the ‘inductive scientific method’ in his *Novum Organum* (1900 (1620)). This new empiricism searched to literally forge empirical data into scientific knowledge. According to Bacon, the scientific search for the truth must start with assembling ***all*** the relevant instances and different aspects of the phenomenon under study. In a second stage, the data thus collected should be organized in useful categories, or variables as we have come accustomed to label them. The third stage consists of finding correlations between the relevant categories or variables. In the fourth stage of any such scientific enterprise, theory is literally forged out of the correlations that the third stage has brought to light or identified.

As early as the middle of the eighteenth century, David Hume, in his *Enquiry Concerning Human Understanding* (1748) has highlighted the limitation of the ***inductive***

*reasoning* approach to scientific research. The process of collecting all the relevant information on any given phenomenon we may wish to study is, of course, infinite. The choice of categories or *variables* is never obvious and the correlations between variables may or may not be statistically and/or substantively significant. Even if they are statistically significant they may be due to spurious relations, intervening variables or measurement errors. But most damaging to the inductive method is the realization that correlations between variables never amount and can never be augmented to the status of theoretical knowledge. They may provide a better picture of the data and the reality but correlations never prove causality. They are, at best, a useful way to organize, summarize and describe observed patterns in the empirical data. Correlations among meaningful variables are extremely important in as much as they outline what we need to explain as they provide us with systematic and precise characterization of the structure of the empirical evidence but it is misleading to see in them part of the explanation itself. If two variables correlate, then the theoretical analysis will have to come up with an explanation for this correlation, but the correlation itself does not explain anything.

Yet the reasons for a correlation to be found or absent from a data set may be other than the worthiness, or lack thereof, of the relation postulated by the theoretical argument. Just as intervening variables and measurement errors can induce into a data set spurious relations and correlations, they may hide genuine ones. The reduction of the information inherent in a data set to any functional form is bound to be misleading in as much as any reductionism is. Since the scientific enterprise requires that we distinguish between important relationships and ‘white noise,’ we must use different judgment criteria to go about this exercise in reductionism. It is here that the approach of visualizing data is categorically different from traditional statistical practice. The

traditional approach is to emphasize the functional form that best summarizes the pattern in the data and to bring statistical measures of fitness to bear on the validity and reliability of this summary. In visualizing data we allow ourselves more freedom of expression as we emphasize the compound web of relations and correlations in the data, highlighting, at times, a particular pattern, but paying much more attention to what shapes this pattern and the ‘unsorted’ environment in which it is found.

The patterns we sort out of the data, in any exercise in quantitative analysis, and assess the validity and reliability of, may or may not be a genuine one and it may or may not reveal a significant point or a theoretical postulate, but everything else is in the residuals. Visualizing data leaves the reader with so much more of this ‘everything else.’ In this way, unexpected consequences, new insights and alternative explanations are more likely to emerge and enrich the scientific enterprise.

Shortly after Francis Bacon published his *Novum Organum* (1620) Rene Descartes published in France his *Discourse on the Method of rightly conducting one’s reason and reaching the truth in the sciences* (1637). *The Method*, as it is often referred to, is regarded as a masterpiece in philosophical reasoning and the origin of the logic of deductive reasoning. It suggests almost the mirror image of inductive reasoning as ‘the method ...[of] reaching the truth in the sciences.’ The first step in any exercise in deductive reasoning consists of a choice of a set of non-controversial and irrefutable assumptions. In the second stage, an infallible method is chosen that is used, in the third stage, to derive, logically or mathematically, conclusions based on the set of assumptions, chosen at the outset as the foundation for the enquiry. In Descartes’ deductive scientific method empirical evidence comes into the picture only at the last stage. The conclusions obtained deductively in the third stage by deriving them mathematically or logically from

the set of assumptions chosen in the first stage, using the method one adheres to in the second stage, are subjected to a 'reality check' by testing whether they are consistent with the empirical evidence we possess of the studied phenomenon.

Again, as early as the middle of the 18<sup>th</sup> century David Hume (1748) highlighted the shortcomings of deductive reasoning as a tool to reach scientific truth. Assumptions are just that: they are assumptions. Even in as a precise scientific field as topology, mathematicians often choose different assumptions to construct different topological spaces in order to study different abstract ideas. These topological choices often lead to dramatically different conclusions, depending on the topological space chosen for the sake of any particular study. The choice of the method is almost as arbitrary and may lead to desperately different conclusions. In contemporary social sciences game theory is probably the most dominant method of deductive reasoning and the prisoners' dilemma is probably the best known exemplar of the usefulness of this method for the study of human interactions in the market place or in politics (Doron and Sened, 2001). But game theorists have, available to them, two different game theoretical approaches: cooperative game theory and non-cooperative game theory (Doron and Sened, 2000: 19-27). The debate about which one of them is a more appropriate foundation for the study of social sciences is on going but no one questions the logical adequacy and the coherence of any of these two approaches. And yet, if the prisoners' dilemma game is analyzed using non-cooperative game theory, the only scientific conclusion that can be reached is that humans will find it either very hard, in the context of the infinitely repeated game (e.g. Calvert, 1995) or impossible, in the one shot game, to cooperate. If we use cooperative game theory, which is just as mathematically rigorous and sound as non-cooperative game theory, the only scientific conclusion that we can reach is that players will cooperate for the benefit of all (Buchanan, 1986). Thus, the choice of two

different methods, each perfectly sound and perfectly rigorous as a logical framework of analysis, may lead us to reach diametrically opposed conclusions.

But once again, the most damaging criticism of the deductive method of scientific reasoning is the difficulty encountered when it comes to put it to empirical test, at least in the sense of empirical testing intended by Descartes (1637). This difficulty stems from three main sources. First, theoretical structures, of the deductive type discussed here, are inherently logical and devoid of empirical content. To endow theories with ‘empirical relevance’ we use what Hempel (1966: 70-2) called ‘bridge principles’ that attach to perfectly abstract logical or mathematical primitives, an empirical content. For example, in the *spatial theory of electoral competition* (Austen-Smith, 1983) we use Euclidean topological spaces that we endow with empirical relevance by interpreting these mathematical abstractions to represent political policy spaces of the choice sets of society. We then go a step further by interpreting points in these spaces to represent either ideal points (most preferred outcomes) of individuals in society, or feasible outcomes in the policy space. And then, further, to treat distances in these spaces as representing, in some loss function, increasing in the distances between ideal points of individuals in society and feasible outcomes in the policy space. These ‘empirical’ interpretations of perfectly abstract mathematical entities may or may not seem reasonable but they are clearly arbitrary.

The second difficulty in empirical testing of deductively inferred theories is the infinite complexity of the phenomenal world. No set of observation can deterministically prove a theoretical claim. The set of all possible occurrences of any particular event is infinite. We only get to observe a small subset of these occurrences and the infinite number of events we never get to observe could always turn out to be very different. Statistics has gone around this problem by using probability theory to put very low

probabilities on events that recurrently fail to materialize. But to do this one needs to stick by the very problematic assumption of ‘*everything else being equal.*’<sup>1</sup> In other words, statistical inference allows us to infer that an event that fails to happen in repeated trials have a low probability to occur if the exact same conditions continue to hold. But in the phenomenal world conditions constantly change and many of these constantly changing conditions are out of the reach of scientific control,<sup>2</sup> excluding, maybe, laboratory experiments.<sup>3</sup> The second side of the same problem is the difficulty of ruling out alternative explanations. Every attempt to reject a hypothesis may fail because of so many different reasons unaccounted for. While visualizing data cannot possibly solve this problem, it is an explicit effort to leave the scientist and the critical reader with more information on the patterns that may be present in the data set studied. Visualizing tools are conscientiously used to report further information beyond point statistics and variance measurements. Such summary statistics often hide important information that is present in the data but does not show up in the statistical summaries of the data.

Finally, a failure in fit between empirical data and theories may reflect a problem with the theory, or the data, or the data analysis. We cannot know if the empirical test is flawed or the theory is wrong. But in the end, this should probably not be the main concern of science. The following sections develop an argument according to which the main concern of science is to match patterns in the phenomenal world with logically coherent structures that we use as explanations. Visualizing data is probably not the best tool to test hypotheses but this set of techniques allows us to better evaluate the fit between systematic patterns in the data and the structure proposed by theoretical speculation. In short:

*visualizing data is an attempt to do with the data more than just test or refute hypotheses, and test and refute hypotheses is probably not what we want to do with data to begin with.*

### ***3. Relativism and Falsificationism***

Centuries of philosophical and scientific explorations did not narrow the gap between the world of ideas and the phenomenal world.<sup>4</sup> In the 1970s two essentially negative approaches to the problem have emerged: ***Relativism*** and ***Falsificationism***. Most practitioners in the social sciences endorse a version of falsificationism associated with Popper's (1959) ***Logic of Scientific Discovery***, as a viable answer to the specter of relativism. In this section I explain that falsificationism is just as untenable as relativism. This argument is not new (e.g. Lakatos, 1978a: 35) but it is often overlooked in the professional circles of social scientists in general and political scientists in particular.

In the end, both sophisticated relativists (e.g. Feyerabend, 1975) and sophisticated falsificationists (Lakatos, 1978a) converged to a common view emphasizing proliferation and consistency as the fundamental criteria for scientific progress and viable alternatives to the search of 'truth' and the reliance on hypothesis testing. While traditional hypothesis testing is most appropriately associated with falsificationism, visualizing data techniques seem more appropriate to serve the criteria of proliferation and consistency that replaced 'critical tests' (Hempel, 1996: 25-8) and hypothesis testing in the eyes of the leading contemporary philosophers of science (Lakatos, 1978a; Feyerabend, 1975).

In ***Reason, Truth and History***, Putnam (1981) makes a useful distinction between *metaphysical realism* and *total relativism*. ***Metaphysical realism*** holds that "the world consists of a fixed totality of mind-independent objects" and that "there is exactly one true and complete description of the way the world is" (Putnam, 1994: 429). ***Total relativism***<sup>5</sup> holds that there exist no way to distinguish a theory as 'truer' than another. Both views are untenable. Metaphysical realism cannot account for the world of ideas unless it reduces them to phenomena in the physical world. Not only is this reductionism not helpful and

untenable, it also begs the question of the relationship between ideas and facts. Total relativism is self-defeating as it is easy to see that if no theory is truer than another, relativism itself has no merit either, other than just, yet another, theory. As a viable middle ground between these two objectionable extremes, Putnam (1981) suggests his ‘internal realism’ that he later summarized in the following terms (Putnam, 1994: 439-440):

Let me conclude by saying a little more about my own picture, for I do have a picture. I don’t think it is bad to have pictures in philosophy. What is bad is to forget they are pictures and to treat them as “the world.” In my picture, objects are theory-dependent in the sense that theories with incompatible ontologies can both be right. Saying that they are both right ...is not saying that there are both absolute space-time points and points which are mere limits. It is saying that various representations, various languages, various theories, are equally good in certain contexts. In the tradition of James and Dewey, it is to say that devices that are functionally equivalent in the context of inquiry for which they are designed are equivalent in every way we have a “handle on.”

It is precisely this notion of theory-dependence of objects that makes falsificationism untenable. Given the unwarranted, yet widespread acceptance of a ‘dogmatic’ version of ‘Popper’s<sup>6</sup> falsificationism in the social sciences in general and in political science in particular, I quote Lakatos’ (1978a: 14-16) rebuttal of this approach at some length in the hope of bringing it to the attention of practitioners in the field (all *italics* are in the original text):

Dogmatic falsificationism, however, is untenable. It rests on two false assumptions and on a too narrow criterion of demarcation between scientific and non-scientific.

The *first assumption* is that there is a natural, *psychological* borderline between theoretical or speculative propositions on the one hand and factual or observational (or basic) propositions on the other...

The *second assumption* is that if a proposition satisfies the psychological criterion of being factual or observational (or basic), then it is true...

These two assumptions secure for the dogmatic falsificationist's deadly disproofs an empirical basis from which proven falsehood can be carried by deductive logic to the theory under test.

These assumptions are complemented by a *demarcation criterion*: only those theories are 'scientific' which forbid certain observable states of affairs and therefore are factually disprovable. *Or, a theory is 'scientific' if it has an empirical basis.*

But both assumptions are false. Psychology testifies against the first, logic against the second, and finally, methodological judgment testifies against the demarcation criterion. I shall discuss them in turn.

- (1) A first glance at a few characteristic examples already undermines the *first assumption*. ...For there are and can be no sensations unimpregnated by expectation and therefore *there is no natural (i.e. psychological) demarcation between observational and theoretical propositions.*
- (2) But even if there were such a natural demarcation, logic would still destroy the *second assumption* of dogmatic falsification. For the truth-value of the 'observational' proposition cannot be indubitably decided: *no factual proposition can ever be proved from an experiment.* Propositions can only be derived from other propositions, they cannot be derived from facts...If propositions are unprovable then they are fallible. If they are fallible then clashes between theories and factual propositions are not 'falsifications' but merely inconsistencies. ...*Thus, we cannot prove theories and we cannot disprove them either.*
- (3) Finally, ...*exactly the most admired scientific theories simply fail to forbid any observable state of affairs.*

Thus, Falsificationism and its core premise that science progresses by rejecting false hypotheses, is untenable. *How does science progress then and what do visualization techniques have to do with it?* Kuhn (1962), and other relativists, chose to argue that we cannot assert progress in science. For Kuhn (1962), scientific paradigms are not a whole lot different than stylistic fashions. They appear, following impressive exemplars that are copiously imitated by practitioners in scientific fields (Kuhn, 1962: Chapter II). They are abandoned when new attractive exemplars catch the attention of scientific communities

(Kuhn, 1962: chapters VI-VII) in the constantly evolving (but not necessarily progressing) march of scientific paradigms. Like fashion shows, scientific paradigms have their internal coherence that Kuhn (1962: chapter III) referred to as the practice of ‘normal science’ or ‘puzzle solving’ (Kuhn, 1962: IV) guided by scientific paradigms (Kuhn, 1962: chapter V). Thus progress is obvious within the practice of normal science but is all too illusive when we try to attribute it to a series of successive scientific paradigms (Kuhn, 1962; chapter XIII).

Kuhn’s position was rejected, as it seems that paradigmatic scientific shifts do lead humanity further in reaching better understanding of the phenomenal world. It has by now become clear that science progresses not by the rejection of unwarranted or improbable hypotheses but through the progress of sequentially improved theories. Every so often scientific disciplines experience paradigmatic shifts. These discontinuous leaps lead to new perceptions, previously impossible to even imagine. While these new paradigms are disconnected from the old scientific perceptions of the world, they do not abandon the achievements by their predecessor. Instead, they seem to explain old findings and suggest new ones by proposing new mixes of theories and evidence to advance knowledge further.

Lakatos (1978a, 1978b) provided the most systematic account of this new philosophy of science. The next section is a short summary of his argument. Yet, in my mind, Lakatos (1978a) provided an excellent account of the notion of scientific proliferation and progress, but came short of explaining what exactly is meant by the *consistency criterion* that is so central to his argument. In section five, I use a forgotten insight of Wittgenstein, the notion of ‘*world-pictures*,’ to suggest a tangible consistency criterion. This criterion is central to my argument in favor of using visualizing tools in empirical research in the social sciences. I conclude by explaining how visualizing techniques may be helpful in endowing this criterion with essence and substance.

#### ***4. Lakatos' Methodology of Scientific Research Programmes***

In his *Methodology of Scientific Research Programmes* (1978a: 47-8), Lakatos argues that science advances through the evolution of research program(me)s:

I have discussed the problem of objective appraisal of scientific growth in terms of progressive and degenerating problem shifts in series of scientific theories.

The most important such series in the growth of science are characterized by a certain continuity [that] connects their members. This continuity evolves from a genuine research programme adumbrated at the start.

...All scientific research programmes may be characterized by their 'hard core'.

The negative heuristic of the programme forbids us to direct the *modus tollens* at this 'hard core'. Instead, we must use our ingenuity to articulate or even invent 'auxiliary hypotheses', which form a protective belt around this core, and we must redirect the *modus tollens* to these.

The hard core of a research program includes basic epistemological assumptions, methodological imperatives and definitions of what should be the focus of research. The protective belt of a research program is made of cumulative achievements of the program. As the protective belt grows wider and richer in its achievements, the belt buffers the hard core of the program. A research program degenerates when its protective belt withers under successful attacks, but the program's hard core never stands 'direct trial'. For the core consists of primitives, axioms, premises or assumptions that need not and cannot withstand the test of either external verification or falsification. The program collapses only when its protective belt is eroded to become so thin and poor that it can no longer withstand systematic, scientific criticism.

Lakatos (1978a) made an important step forward by pointing out something that most practitioners of science in general and the social sciences in particular are keenly aware of. Critical tests do not prove or disprove anything, nor do we rush to reject

theories because one or even several of their derivative hypotheses were rejected in empirical testing. Rare are the instances in which empirical data lead researchers to reject a theoretically meaningful hypothesis. Instead, researchers make concrete efforts to accumulate hypotheses that they make every effort to ‘fail to disprove.’ This practice was labeled by Lakatos (1978a: 25) as ‘methodological falsificationism’ and explained as follows (*Italics in the original*):

*The methodological falsificationist separates rejection and disproof, which the dogmatic falsificationist had conflated. He is a fallibilist but his fallibilism does not weaken his critical stance: he turns fallible propositions into a ‘basis’ for a hard-line policy. On these grounds he proposes a new demarcation criterion: only those theories – that is, non-‘observational’ propositions – which forbid certain ‘observable’ states of affairs, and therefore may be ‘falsified’ and rejected, are ‘scientific’: or, briefly, a theory is scientific (or ‘acceptable’) if it has an ‘empirical basis’.*

This new found ‘empirical basis’ is of crucial importance as it goes around the problem that made dogmatic falsificationism untenable. “...probabilistic theories may qualify now as ‘scientific’: although they are not falsifiable they can be easily made ‘falsifiable’ by an *additional (third type) decision* which the scientist can make by specifying certain rejection rules which may render statistical interpreted evidence ‘inconsistent’ with the probabilistic theory.” (Lakatos 1978a: 25).

But this ‘empirical basis’ should not be taken too literally. Lakatos (1978a: 32) draws a sharp distinction between what he calls ‘naïve methodological falsificationist’ and ‘sophisticated methodological falsificationist’:

For the naïve falsificationist, a theory is *falsified* by a[n] ... ‘observational’ statement which conflicts with it (or which he decides is conflicting with it). For the sophisticated falsificationist a scientific theory T is *falsified* if and only if another theory T’ has been

proposed with the following characteristic: (1) T' has excess empirical content over T: that is, it predicts *novel* facts, that is facts improbable in light of T, or even forbidden by T; (2) T' explains the previous success of T, that is all the unrefuted content of T is included ...in T'; and (3) some of the excess content of T' is corroborated.

This seems to me to be the exact point where visualizing data may play a crucial role. Traditional hypothesis testing tends to limit itself to the report of the parameters and test statistics used in refuting the null hypothesis. The approach of visualizing data puts forwards a variety of unabridged descriptions of data sets without the reductionist tendency of describing the data in terms of a set of parameters that measure the goodness of fit of the data with a more or less adequate functional form summary of it. Theories are not corroborated or falsified through the success or failure of an attempt to reject null hypotheses. Consecutive series of theories are either *progressive* or *degenerating* (Lakatos, 1978a: 33-5):

Let us take a series of theories,  $T_1, T_2, T_3, \dots$  where each subsequent theory results from adding auxiliary clauses to ...the previous theory in order to accommodate some anomaly, each theory having at least as much content as the unrefuted content of its predecessor. Let us say that such a series of theories is *theoretically progressive* ... if it predicts some novel, hitherto unexpected fact. Let us say that a theoretically progressive series of theories is also *empirically progressive* ... if some of this empirical content is also corroborated, that is, if each new theory leads us to the actual discovery of some *new fact*. Finally let us call a problemshift [series of theories] *progressive* if it is both theoretically and empirically progressive, and *degenerating* if it is not.

...Contrary to naïve falsificationism, *no experiment, experimental report, observation statement or well-corroborated low-level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory.*

Science, then, progresses through improving the precision, scope and depth of theoretical explanations. Forcing theorists and empiricists to constantly expand and detail their theories and empirical analysis to find and make room for more and more

facts. Such improvements can only be tested against richer and richer descriptions of the empirical world. Traditional hypothesis testing consists of answering the question of whether a particular pattern or relationship that is deduced from theory is present in the data. The null hypothesis states that such relationship is not found in the data. If the null hypothesis is shown to be very unlikely we ‘reject’ it and interpret it as a reason to continue and hold the hypothesis derived from our theoretical speculation. Visualizing data supplements this technique with providing visual evidence to the effect that the theoretically postulated pattern can be seen in visual inspection of the data. In addition, by providing more detailed descriptions of the data it allows us to spot those *new facts* that escape the analysis in the process of reducing it to functional forms that are presumed to best summarize it.

In principle, every interesting fact that is spotted in the visualization of the data can then be tested using significant tests and other tests of goodness of fit. That is, visualizing data provides a richer description of the data, showing not only the summary statistics and their significance but multiple other information such as the exact shape of the distribution of the data points, sorted by different variables, schematic patterns visible in the data, in the residuals or just fancy descriptive that are often time more telling than summaries of postulated relationships. Resorting to such techniques often times makes it harder to interpret the data in traditional statistical terms but these techniques prove to be more challenging to theorists and empiricists alike as they force them to make sense of real patterns in the data rather than postulated summaries of it.

But visualizing data represents not just an incremental improvement on traditional hypothesis testing, but a more dramatic departure from conventional empirical analysis. The next section proposes an elegant rationale for this radical departure.

## ***5. Wittgenstein's Pictures of the World: Beyond Hypothesis Testing***

Many producers and users visualizing data techniques may see in it just a better way to present the full import of the data. In this section I suggest a more radical view of what is achieved through visualizing data drawing from Wittgenstein's *Tractatus Logicus Philosophicus* (1922). The short aphorisms of the text have left generations of philosophers guessing and double guessing arguments and notions that in many ways remain enigmatic to this very day. One of these notions is Wittgenstein's idea of 'world pictures.' I wish not to enter the philosophical debate surrounding this notion because I believe it is remarkably pedantic and of little help. I believe that Wittgenstein (1922) has made himself quite clear in his presentation of his view and I wish to draw directly from this presentation (pp.39-43):

- 2.1 We make ourselves pictures of facts.
- 2.11 The picture presents the facts in logical space, the existence and non existence of atomic facts.
- 2.12 The picture is a model of reality.
- 2.13 To the objects correspond in the picture the elements of the picture
- 2.131 The elements in the picture stand, in the picture for the objects.
- 2.14 The picture consists in the fact that its elements are combined with one another in a definite way.
- 2.15 That the elements of the picture are combined with one another in a definite way, represents that the things are so combined with one another. This connexion of the elements of the picture is called its structure and the possibility of this structure is called the form of the representation of the picture.
- 2.151 The form of representation is the possibility that the things are combined with one another as are the elements of the picture.
- ....
- 2.16 In order to be a picture a fact must have something in common with what it pictures.
- 2.17 What the picture must have in common with reality in order to be able to represent it after its manner – rightly or falsely – is its form of representation.
- ....
- 2.18 What every picture, of whatever form, must have in common with reality in order to be able to represent it at all – rightly or falsely -- is the logical form, that is, the form of reality.
- ....
- 2.19 The picture has the logical form of representation in common with what it pictures.

It seems to me that the best way to interpret and use visual tools in representing data is to think of the exercise in those Wittgensteinian terms. The theorist draws a logical picture while the empiricist constructs a picture of the data. The theoretical picture is true if it shares some basic logical aspects

or what Wittgenstein calls a ‘logical form’ with the empirical pictures of the data about which the theory sets itself to speculate. We are yet to fully comprehend the process by which the goodness of fit between the theoretical pictures and the empirical pictures may be assessed. And yet most works in this field use some systematic measures to determine the goodness of fit between the logical form of the theory and the logical form of the picture of the data set.

Lakatos (1978a) made sense of how scientific progress is achieved by explaining the logic of scientific programs. But he left us with doubts as to how one is supposed to assess the empirical corroboration of consecutive theories within a research program so as to determine whether the program is *progressive* or *degenerating*. What is the practical meaning of ‘consistency’ and how do we go about assessing it? Wittgenstein seems to have pointed to a solution to this problem about five decades earlier. What Wittgenstein seems to suggest is that the extent to which empirical evidence supports new (and old) theories can be assessed by the extent to which the logical form of the theoretical argument conforms with the logical form of the visual representation of the data. I tend to believe that the fact that this important insight was overlooked for so many years has something to do with the fact that until recently we did not use visual tools in theory or in empirical analysis. But current practices in scientific research have changed that.

Practitioners of all sciences tend to use more and more visuals, in the presentation of the theory and in the reports of the empirical evidence. What remains to be clarified is how exactly do we go about the process of assessing the consistency or conformity of the structure of ‘empirical pictures’ with ‘theoretical pictures’ of the world. In the next section I provide three examples of how we may proceed in this process. In the concluding remarks that follow I discuss briefly the urgent need to formalize and systematize this practice with more precise and reliable tools.

## **6. *Three Examples: Constructing Pictures of the World*** (section under construction)

In this section I use three examples to illustrate my point. Starting from the simplest regression model of the effect of newspaper coverage on electoral success, through a generalized event count model of U.S. congressional reforms in the last century, and concluding with an analysis of general trends in electoral behavior in the last two decades in Israel. The point of this section is that different theoretical argument warrant different types of 'pictorials' to test their validity and empirical relevance. Therefore, the issue is not what statistical method is better, but what statistical model best fit to provide a picture of the empirical data that may help us assess the validity and consistency of the theoretical argument with empirical evidence in our possession. From a practical point of view the lesson to be drawn is that instead of finding the right hypothesis to test, which, as we all know, is, often times, a rather awkward exercise and often times besides the point, students of social sciences need to spend time finding the right visual presentation of the data that best support (or may question) the theoretical argument at hand.

***Example 1: There is nothing Wrong with Using and Interpreting Regression (Achen, 1982)***

***Example 2: Congressional Reform: A Generalized Event Count Model***

***(Ainsworth and Sened, 2000)***

***Example 3: Constructing a Bayesian picture of a spatial model of electoral behavior***

***(Quinn, Ofek and Sened, 1998)***

## ***7. Conclusion***

For decades, political scientists made remarkable efforts to summarize complex political phenomena in linear parametric statistics. In the last decade, modern estimation techniques surfaced to allow us to move away from the linearity assumptions but retain the elegant appearances of parametric statistics that allow a summary of the data into a few point and distribution statistics. But in this essay I suggest that summarizing the data into point and distribution parameters may not be such a good idea to begin with. The world of politics is a complex one. No one summarizes the structure of the DNA in a functional form? We don't think that trying to compress Italian politics, St. Louis politics or the lending trends in the global markets to simplistic functional forms makes much more sense than to try and inform the public on the structure of the DNA by its functional form.

The central argument of this essay is that we must abandon naïve falsificationism as a philosophy of social science. Instead, we could adopt the appealing Wittgensteinian philosophy of science that searches to establish the relationship between analytical theory and empirical evidence on the similarity in the logical structure of the theoretical and the empirical pictures of the world. The role of research is to highlight the structure of the theoretical picture and then demonstrate that a sensible picture of reality is characterized by the same logical structure. To do this parametric statistics will have to return to its natural role in assessing the strength, the validity and the reliability of 'localized' relationships within the data and leave more room for visual tools to explore the intricate way in which these local relations relate to the bigger picture. More importantly, the next task is to improve on existing measures and find new techniques to assess the goodness of fit between the structure of theoretical arguments and the structure of empirical representation of the world.

## **Reference List**

- Achen, Christopher H. 1982. *Interpreting and Using Regression*, London: Sage Publications.
- Ainsworth, Scott and Itai Sened. 2000. "Institutional Stickiness and Congressional Reform," Unpublished Manuscript, Washington University in St. Louis.
- Austen-Smith, David. 1983. "The Spatial Theory of Electoral Competition: Instability, Institutions and Information," *Environment and Planning*, 1:439-59.
- Bacon, Francis. (1900 (1620)). *Novum Organum*, Translated by R Ellis and James Spedding, London: George Routledge and Sons.
- Buchanan, James. 1986. *Liberty, Markets and the State*, New York: New York University Press.
- Calvert, Randall. L. 1995. "Rational Actors, Equilibrium and Social Institutions" in Knight and Sened [Eds]: *Explaining Social Institutions*, Ann Arbor: The University of Michigan Press.
- Descartes, Rene. 1637. *Discourse on the Method of Rightly Conducting the Reason and Seeking the Truth in the Sciences*. Edinburgh; London: W. Blackwood & Sons.
- Doron and Sened, 2001. *Political Bargaining: Theory Practice and Process*, London: Sage Publication.
- Feyerabend, Paul. 1975. *Against Method*, London: New Left Books.
- Hempel. G. Carl. 1966. *Philosophy of Natural Science*, Englewood Cliffs: Prentice-Hall.
- Hume, David. 1748. *An Enquiry Concerning Human Understanding*, Chicago: Regency Company (1956).
- Kuhn, Thomas, S. 1962. *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press.
- Lakatos, Imre. 1978a. *Methodology of Scientific Research Programmes, Philosophical Papers*, Volume I, Edited by J Worrall and G.P. Currie, Cambridge: Cambridge University Press.
- Lakatos, Imre. 1978b. *Mathematics, Science and Epistemology, Philosophical Papers*, Volume II, Edited by J Worrall and G.P. Currie, Cambridge: Cambridge University Press.
- Malcolm, Norman. 1977. "Moore and Wittgenstein on the sense of 'I know'" in N. Malcolm, *Thought and Knowledge*, Ithaca: Cornell University Press.
- Ofek, Dganit, Kevin M. Quinn and Itai Sened. 1998. "Voters, Parties and Coalition Formation in Israel: Theory and Evidence," Unpublished manuscript, Washington University in St. Louis.
- Popper, Karl. R. 1959. *The Logic of Scientific Discovery*, London: Hutchinson.
- Putnam, Hilary. 1994. "A Defense of Internal Realism," in Paul Horowich, [Ed.] *Theories of Truth*, Brookfield: Dartmouth.
- Putnam, Hilary. 1981. *Reason, Truth and History*, New York: Cambridge University Press.
- Ricketts, Thomas. 1996. "Pictures, Logic, and the Limits of Sense in Wittgenstein's Tractatus," in H. Sluga and D.G. Stern [Eds.] *The Cambridge Companion to Wittgenstein*, Cambridge: Cambridge University Press.
- Wittgenstein, Ludwig. 1922 *Tractatus Logicus Philosophicus*. London: Routledge.

## *End Notes*

---

<sup>1</sup> See Lakatos, 1978a: 27.

<sup>2</sup> This leads Lakatos (1978a: 11) to remark that: “under very general condition all theories have zero probability, whatever the evidence.” “*All theories*” he concludes (*italic in the original*) “*are not only equally unprovable but also equally improbable.*” On this important point, see also, Lakatos 1978b: 154, ff; 162, fn. 2, and 164.

<sup>3</sup> This is certainly one of the reasons why experimental research has become so prominent in social sciences in the last two decades. The important achievements of this growing sub-field in the social sciences notwithstanding, a constantly nagging question persists with regards to the amount of control that scientists can master even in the laboratory experimental environment, without rendering their research trivial.

<sup>4</sup> For an excellent discussion of G.E. Moore and Wittgenstein on this matter, see N. Malcolm (1977). I return to Wittgenstein later in the text but a serious discussion of G.E. Moore is beyond the scope of this essay. Also relevant here is Bertrand Russel’s so called ‘Multiple relation theory of judgment’ that is largely due to G.E. Moore’s metaphysics and is a main target of criticism in Wittgenstein’s *Tractatus Logico-Philosophicus* (1922, see Ricketts, 1996) to which I return later in the text.

<sup>5</sup> That Putnam (1981) associates with Kuhn (1962) and Feyerabend (1975).

<sup>6</sup> Most scholars of Popper agree that Popper himself rejected this naïve or dogmatic interpretation of his ‘falsificationist’ philosophy of science (e.g. Lakatos, 1978a: 31).