COMPUTATIONAL SOCIOLOGY

DISCUSSION AND RESEARCH

Antonio Aguilera Ontiveros
Norma Leticia Abrica Jacinto
COORDINATORS

Computational Sociology
Discussion and Research

COLECCIÓN INVESTIGACIONES

**COMPUTATIONAL
SOCIOLOGY
DISCUSSION AND RESEARCH**

ANTONIO AGUILERA ONTIVEROS

NORMA LETICIA ABRICA JACINTO

COORDINATORS

EL COLEGIO
DE SAN LUIS

Esta obra fue dictaminada por evaluadores externos a El Colegio de San Luis por el método de doble ciego

Primera edición en formato digital: 2021

© Por la coordinación: Antonio Aguilera Ontiveros, Norma Leticia Abrica Jacinto

© Todos los textos son propiedad de sus autores

D.R. © El Colegio de San Luis
Parque de Macul 155
Fracc. Colinas del Parque,
San Luis Potosí, S.L.P. 78294
<https://libreria.colsan.edu.mx/>

E-ISBN: 978-607-8794-38-6

Diseño de la portada: Natalia Rojas Nieto

Diseño de ePub: Ernesto López Ruiz

Hecho en México

A REFLECTION ON MODELING AND THE NATURE OF KNOWLEDGE

JORGE ZAZUETA GUTIÉRREZ

ABSTRACT

We reflect on the criticism that the methods in the natural sciences have been an object of and argue that these very methods, seen from the right perspective, are crucial to the progress of social science. We propose a knowledge generation framework that sets the minimum requirements for a claim to be considered scientific knowledge and sketch an ontology of the objects of study. We present science as an accumulating yet self-revising process and present examples in the social sciences.

INTRODUCTION

The reputation of modeling as a formal methodology in the social sciences has fluctuated from being essential to reductionist and even restrictive to theory development. Rather than taking a side, we discuss the role of modeling as a unifying methodology across schools of thought, arguing that the clarity required by the modeling process provides theoretical advantages beyond the specific results of a given prototype by reducing ambiguity and allowing for generalization. An explicit model provides not only structure but a common language and a set of tools to advance the development of social theory by facilitating interdisciplinary discussion and providing a mechanism to understand the impact of our propositions.

Our discussion is driven by developing a framework that relates modeling and knowledge, or theory building, at different levels. The framework emphasizes falsifiability, generalization, and accumulation as the basic tenets of knowledge generation while recognizing various levels of mind-independent reality, from concrete physical objects to abstract social constructs that might be unobservable but detectable.

Throughout our discussion, we classify our models into conceptual, and formal and further subdividing formal models in terms of their output as (1) closed-form, (2) qualitative and (3) numerical, placing formal constructs higher in a hierarchy while recognizing the conceptual component as a necessary condition for model building and computation as, increasingly, the tool of choice to instantiate and explore a given theoretical or analytical construct; noting that the purpose of modeling in this context is not constrained to prediction, but, just as importantly, it aims to explain and to demonstrate plausibility.

Social theory is represented by a range of worldviews whose proponents are not always in line with the idea of modeling as a unifying research methodology. While recognizing that not all of sociology is quantifiable, we do suggest that, to the extent that a social process can be defined in sufficiently unambiguous terms to be modeled, analytical and computational approaches can contribute significantly not only to the understanding of specific social phenomena, but through the necessary clarity required by modeling, contribute to the identification of structurally similar processes in distinct settings leading to a more integrated, and less divided, general theory.

BASE FRAMEWORK

We adopt a working definition of *knowledge* as a claim about an *object* or a category of objects and thus, only meaningful in the context of the object that is being known. For us to make such a claim, we need a *method* that allows us to gain insights into the object. The relationship among knowledge, method and object is depicted in Figure 1, stressing the interdependence of the three elements.

The recursive nature of this process is the key to what we call *accumulation*. By accumulation, we mean that knowledge is not acquired by a series of independent and isolated acts, but that it grows from an established basis of facts and theories and this process includes the development and improvement of methods as well as the continuous refining of such facts and theories.

FIGURE 1
KNOWLEDGE MODELING FRAMEWORK

Accumulation is crucial for the development of any significant body of knowledge. In order to claim that dogs are omnivores, it is necessary that we already have the concepts of dog and omnivore. Without accumulation we are doomed to paralysis and any criticism will have to deal with explaining what criticism is without assuming previous knowledge of the term.

Language, culture, religion and even skepticism are a result of accumulation. To be fruitful, this process needs not only to incorporate new elements to our body of knowledge, but to refine and even prune those in need of revision.

Legitimacy of Knowledge

Without basic quality control, claims can be problematic. If I make an assertion about an object, say, that the earth is spherical, you can always make a contradictory statement and assert that it is flat. We would both be making conflicting claims about an object, presumably supported by some research method.

To resolve the conflict, we need our process to display some essential features: (1) the claim is falsifiable, (2) the method is replicable and rationally linked to the claim, and (3) the object under study is measurable. The claim that the earth is flat is certainly falsifiable and has been falsified both directly as observed from a satellite feed or indirectly inferred by looking at a ship “sinking” as it recedes in the distance, contradicting flatness. These and other observations can be replicated by other researchers to measure curvature either binarily or precisely, making every claim provisional and open for revision.

While these essential features legitimize knowledge, they do not necessarily make it relevant or transcendental. We propose a continuous hierarchical characterization based on qualities of the process (see Figure 1). A claim gains status to the extent that it is general and fruitful. Knowing that my dog barks lies lower in the hierarchy than knowing that all dogs bark. Similarly, knowing how an apple falls from a tree is not as impressive as knowing how any object is attracted to any other, anywhere.

A similar criterion applies to the method. Is it specific to the object under study or can it be applied elsewhere? The universality of the object is a more controversial issue and we will spend more time discussing it in the next sections as we delve deeper into the features of our proposed framework.

The Nature of Knowledge

The idea of knowledge has been—and still is—at the center of many philosophical debates, and pretty much any definition is bound to be challenged. There is some agreement on classifying knowledge in three main types: (1) *acquaintance knowledge*, (2) *knowledge-how or procedural knowledge* and (3) *propositional knowledge*. Scientific knowledge is usually tagged as a subtype of the latter, but many agree that procedural knowledge is also an important aspect of science. Our discussion will hover around the loosely defined idea of scientific knowledge, without diving into semantic subtleties. For a more comprehensive discussion, the reader is referred to [1].

Our central focus will be on explanation. While prediction and control are appropriately regarded as critical features of science, explanation enables both. A good discussion on the role of explanation in science can be found in [2].

The Case for Falsifiability

One afternoon in the seventeenth century, Galileo allegedly dropped three balls of steel from the Tower of Pisa. The balls had different sizes and weights and, contrary to common knowledge at the time, they all landed simultaneously.

Even today, there is a variety of classroom experiments and demonstrations to falsify the—still—common belief that heavier objects fall faster than lighter ones.

Competing ideas contribute to the accumulation of knowledge because they are falsifiable, if there were not, a mounting body of contradictory factoids will clutter our understanding and impede progress altogether. This is not a philosophical position, but a pragmatic one.

The claim that heavy objects fall faster than lighter ones is an appealing one and is apparently consistent with common sense. Were we to allow this view to share the same factual status as the one claiming that heavy objects will fall just as fast as lighter ones, absent of any external forces, knowledge derived from these premises will be plagued with logical paradoxes and disastrous practical consequences.

As we are allowing a claim to be true and false simultaneously, we can validate all sorts of outlandish conclusions. In a calm and windless afternoon, a light person can jump off a tall building and a proportionally heavier one could jump after her and catch on, given that he is heavier and falls faster than she does. Once he catches on, because it is also true that both fall with the same speed, they can choose to maintain the same velocity and chat. If they are in a hurry, they could hug to speed up, or not. From the perspective of the heavier person, the lighter one could fly up at any time, but also stay still relative to her. This kind of dynamics seem only possible for visitors from Krypton.

Fortunately, both claims are falsifiable, and we can consistently replicate Galileo's experiment to choose one claim over another. The implications to aviation, ballistics and engineering demonstrate the pragmatic importance of this claim and the developments from the time of Galileo illustrate the principle of accumulation.

The Fruitfulness of Some Ideas

While any falsifiable claim is by our definition a candidate to the body of scientific knowledge, some ideas are more influential than others and occupy a higher level in the hierarchy. Stating that my coffee mug is blue, is a falsifiable claim and would likely be validated by countless independent methods but is of little relevance beyond my immediate reality. The idea of evolution, advanced by Charles Darwin in *The Origin of Species* [3], on the other hand, is one of the most influential theories of all times. Not only as the key to modern evolutionary biology, but as inspiration to theories outside biology such as game theory, linguistics and economics.

Not every "fact" in an established theory remains untouched. Often, evidence falsifies aspects of a theory and adjustments need to be made. Far from a weakness, this is a perfect example of the tentative nature of scientific theories and the self-updating process of science. In this sense, we can say that the current theory of evolution is the best version available and it represents our current beliefs in that matter.

This is a critical point in our discussion, as accumulation and self-actualization tend to receive less attention in the social sciences. If I want to learn about evolution or classical mechanics, I look for a modern text that incorporates the accumulated knowledge and present the ideas of the discipline from an integrated and current perspective. I would not rely on the original work of Darwin or Newton. Those are important elements of the history of the discipline, and should be understood from that perspective, but not as contemporary guides for a practicing scientist. Why then, should we still hold to an exuberant variety of schools of thought in the social sciences? Is it because the principle of accumulation and self-updating in the natural sciences simply does not apply? We sustain than it does, and that failing to apply it when constructing social theory can have dire consequences, the least of them being stagnation.

Levels of Reality

Does an objective world exist independently from the observer? If so, can it be comprehended? To quote Erwin Schrödinger, we assume that "the display of Nature can be understood" [4]. Because if not, the pursuit of scientific knowledge is simply futile. Granted that we are not claiming that this is irrefutably the case, we are simply stating that otherwise the conversation is over. Different views about reality have been advanced by philosophers and social scientists. Karl Popper proposed a three-world ontology [5, 6]. In this view, reality occupies three worlds—which he creatively called world 1, world 2 and world 3, see Figure 2. World 1 comprises the physical universe. This includes the living creatures and humans, who evolved from it. As humans experience the physical world, a new psychological world evolves (world 2), from which, the world of the "products of the human mind" is engendered.

FIGURE 2
POPPER'S THREE WORLDS

Within this ontology, the products of the human mind can either be conceptual (*i.e.*, theories or claims) or physical (*i.e.*, helicopters or cupholders). These three worlds interact among each other with the physical world providing experiences to the human mind and thus enriching world 2, which influences world 3 by generating new products that modify world one (see Figure 2).

All objects in the three worlds are considered objective by Popper. He explains that they exist independently from the observer and that they clearly exert influence on each other. As objects in world 2 and world 3 can and do modify the physical universe, they can be regarded as objectively real. These ideas have been compared with the Platonic view, except that Popper does not claim that the objects in world's 2 and 3 are perfect. One source of ambiguity is that physical objects that are created by men, belong to both world 1 and 2. The distinction of a concept and the physical object that holds it is not crystal clear and has been criticized by some, questioning this idea of reality [7].

In the spirit of accumulation, we will build on Popper's idea to construct a modified ontology that will serve as a backdrop for our discussion of scientific inquiry.

First, we will not become entangled in philosophical reflections about reality and existence and will use these terms casually in order to avoid unnecessary linguistic complications when referring to objects in our ontology, whose only purpose is to provide some structure to what we consider the object of scientific inquiry.

Second, we redefine the content of the three worlds to avoid intersections between them and to clarify their relationships.

Figure 3 depicts our proposed model. The physical universe represents all objects that adjust to our common definitions of physicality, either natural or man-created and we, unapologetically, assume its existence as independent from the observer's mind. This world is continuously changing through the effects of natural processes, such as erosion or volcanic activity, as well as man-made processes, such as the building of dams and bridges.

FIGURE 3
MODIFIED ONTOLOGY

An important subset of the physical universe is life. Especially intelligent life. Intelligent beings, such as humans, have a conscious, subjective experience in relation to the physical world they inhabit. These experiences constitute the “conscious experience” universe and are the epitome of interpreted reality or subjectivism. However, these experiences and states of mind trigger into existence what we call the “conceptual universe”, in which all human concepts and theories reside.

The three universes interact continuously. The physical universe evolves both naturally and by the addition of man-created objects. The conscious experience world evolves through new interactions with the physical world, but also from conceptual constructs. Consider racism or religion and how they trigger emotions and feelings that go beyond the basic perceptions of our senses. The conceptual universe is critical to scientific knowledge, since it is itself a mancreated concept. Any idea or concept resides in this circle and they can and do directly affect the physical universe. A key difference with Popper’s view is that only the concepts belong to this circle and—in our view—to exist, they need to be anchored to an object in the physical universe. In this sense, a theory is an object on its own, and is independent of the observer much as Popper suggested, but it can only exist if enabled by a media. Say, a book, in the physical universe where the theory is documented.

The ontology is comprehensive in the sense that any object exists in one of the three circles, yet not every object is knowable scientifically in the sense that any claim we make about it can be falsifiable. Notice, also, that it is not necessary for an object to exist or a concept to be truthful for it to affect the world (e.g. religion, monsters). So, effect is not a sufficient condition of existence.

We say, then, that the subset of elements within this ontology that are knowable is our object of study and what we call science is not a product but a process that generates claims, assembles theories and most importantly revises itself as claims are falsified. Figure 4 illustrates this process. The box represents the state of the art. A network of claims as they have accumulated up to a specific point in time. We think of it as a network since claims are related to each other and can group in clusters to form theories. Claims from these theories are still related to others and can contribute to the network in different ways. This “body of knowledge” grows from within as new ideas are inferred from others but can also receive completely novel ideas.

FIGURE 4
THE PROCESS OF SCIENCE

As the network of claims grows exponentially from these inputs, it also prunes itself from claims that are found to be false. In this sense, the “most current edition” of science is the most relevant, because not only holds the claims that we consider true, but because it has discarded those that are not. Most of the growth come from within and while any theory is always provisional, it gains credibility in terms of its fruitfulness or the amount of explanations that can be inferred from it.

To Measure or not to Measure

The purpose of studying objects in any of the universes depicted in Figure 3 is to make falsifiable claims about their properties, which implies assigning an objective and intersubjective value to such claims. Representation Theory of Measurement [8] provides a theoretical framework that, at a very basic level, consists on defining a homomorphism between an empirical relational system and a formal one, such that the relationships between the empirical objects are represented congruently by the measurement system defined by the morphism. This representation is consistent with the familiar normal, ordinal, interval and ratio scales introduced by Stevens [9] and is reasonably well accepted today, conceptually, despite some criticism to its applicability in actual research. Frigeiro, for example, discusses its limitations and proposes a refinement that is consistent with current metrological practices and definitions [11]. For our purposes, the general notion of measure as the assignment of a formal value to a property of the system under measurement will suffice.

Without a measurement, claims cannot be falsified and hence do not belong to our framework. However, that does not mean that measurements must be infinitely precise. To the contrary, the difference between a measurement and the value of the property being measured is an expected characteristic of it. The level of aggregation at which we study reality, together with the claim we want to make, determine the required accuracy and precision of our measure. Say that we are investigating whether a certain breed of dog can distinguish the colors displayed by a stoplight. We only need a measure that distinguishes red, yellow, green, and “other”. Further precision would not only be unnecessary, but confusing. With a finer metric, different stoplights will show a wide variety of “reds” that will not add any new useful information but instead will complicate our problem (was it magenta or crimson?). All that we need is to clearly distinguish red from the other options. All philosophical rants about your and my red are not significant within our definition and a measurement system that is consistent across all varieties of stoplights. On the other hand, if our prospective claim is about the human limits to detect different colors, we might want to measure frequencies in the light spectrum.

Measuring is in our view a pragmatic device. It helps us establish claims in a way that they are falsifiable and intersubjective, and it is intrinsically related to the level of reality we are investigating. Scientific knowledge is hierarchical. Imagine trying to explain everything from interactions of elementary particles. In principle, everything from chemistry to psychology and political economy is explainable in terms of quantum mechanics, but that path will not take us very far. Instead, we create different categories and look for explanations —and metrics— at that level. We do not think of societies as groups of atoms or theories as the result of complex interactions among neurons. Reality is constructed at the level necessary for us to measure and understand. Furthermore, measurement needs not to be direct. Many objects in either universe cannot be gauged directly, but their effects can be quantified, as is the case of black holes or the bending of light due to a large gravitational pull. In a way, our earlier assessment of red as a category, is simply a quantification of the effect of a range of wavelengths.

Local vs. Universal

Both the objects of study and our claims about them can be local, universal or anything in between. Generality plays a role in our framework as a measure of hierarchy, as the more general claims explain the more particular. If I study the mug I use to drink coffee every morning, I could conclude that it has a handle and my claim is falsifiable, so it can be regarded as scientific knowledge. However, I cannot say anything about other mugs. My knowledge is decidedly local. If, on the other hand I postulate a theory that claims that all mugs, by

definition, have a handle, this is also falsifiable and more general. It explains my mug as well as all others. Some skeptics would say that a claim of universality is one of unfunded vanity, but without generality, progress is challenging. This is a practical issue. Our knowledge that people have two legs and two feet allows us to design pants and cars and shoes and create games like football. Generality spawns knowledge, enables technology and it has undeniably changed all three universes in our framework.

THE METHOD

We have established a working definition of knowledge and sketched an ontology of our objects of study. But how do scientists go about making that connection? Contrary to popular belief, there is not a uniquely defined research procedure known as the “scientific method”. As Steven Weinberg put it [12]:

The fact that the standards of scientific success shift with time does not only make the philosophy of science difficult; it also raises problems for the public understanding of science. We do not have a fixed scientific method to rally around and defend.

To understand the process by which a scientist develops a claim implies understanding inspiration, creativity and even luck among other elements influencing the generation of a hypothesis that are beyond the scope of this paper. We confine the term method to the mechanism that connects knowledge to the object known: (1) measuring the value of a property, (2) making an inference about an object from measurements of its effects, or (3) arriving at a conclusion through logical reasoning from established facts.

For a method to be considered scientific within our framework, it needs to be replicable and provide a rationale for how the results of its application are tied to the underlying claim. Consider the witch scene from Monty Python’s and the Holy Grail [13] where Sir Bedevere, quite socratically, helps the villagers conclude that an accused woman is a witch because she weighs the same as a duck, and since ducks are made out of wood, she must therefore be a witch! While the method proposed by Sir Bedevere is clearly replicable, the rationale by which he links its result to the nature of the accused is ludicrously flawed. The range of applicability of a method is supplementary to its validity and lends it hierarchical status.

Modeling as a Method

In a sense, all methods are models. As we formulate or apply any method to generate knowledge about an object, we are running a model. The question is how explicit it is. Making a model explicit, forces us to clearly define the variables involved and the interactions among them. There is less room for interpretation and the model itself serves as a mechanism to reduce any uncertainty about its workings. In the words of Joshua Epstein [14]:

The choice, then, is not whether to build models; it is whether to build explicit ones. In explicit models, assumptions are laid out in detail, so we can study exactly what they entail. On these assumptions, this sort of thing happens. When you alter the assumptions that is what happens. By writing explicit models, you let others replicate your results.

Explicit modeling, therefore, not only affords clarity and replicability to the process, but is a medium of exploration that can yield results well beyond its initial purpose. The logistic function, for instance, was introduced by Verhulst in the mid-19th century to model the population growth of France, Belgium, Essex and Russia [15]. Since then, it has found applications in economics, chemistry, biology, and sociology to name just a few fields. A Google Scholar search for “application of logistic function” returned 1, 410, 000 results.

Models can be conceptual or formal. The former, specify the elements of the model and its interactions, without providing a detailed or operational description of such interactions. The figures in this paper are examples of this. The latter describe, in an unequivocal and precise manner, these interactions. They are often classified in terms of the primary toolset as mathematical, statistical, and computational. This can be simplistic, as many models have characteristics from more than one of these categories (Think of stochastic differential equations or Markov Chain Monte Carlo), and, of course, all of them are mathematical.

A useful classification of models can be made in terms of their solutions, as (1) closed form (2) qualitative and (3) numerical. Closed-form solutions are typically represented as mathematical formulas. $K(t) = K_0 e^{at}$, for instance, is a solution to the model $dK(t)/dt = aK(t)$ that is both specific —as it can give us a precise answer given the right set of parameters— and general —as it describes a broad range of cases. However, closed-form solutions are not always available and might not even exist. Qualitative solutions provide information about the behavior of the system we are modeling, answering questions about existence, uniqueness, and general behavioral patterns such as fixed points or bifurcations. They can be quite general, but do not provide concrete answers. Numerical solutions on the other hand are specific, providing a concrete answer for given parameter values. They are realizations of the model. They are frequently approximations and often, the only ones we have. Modeling works well as a research method as it forces us to define the objects under study and provides a rationale connecting its results to the property under study. Furthermore, it affords a natural mechanism for inquiry and exploration.

FRAMEWORK APPLICATIONS

As we defined our framework and its accompanying ontology, we relied heavily on examples and reflections from the physical and biological sciences. This is natural, as mature, and well-established models abound in these specialties. However, the social sciences present a great opportunity to apply these ideas and considerable success has been achieved in several fields, including our views on positivism, feminism, and language.

Positivism and Critical Realism

Both positivism and critical realism have been criticized for assuming that social phenomena can be explained from simple premises and challenged by feminists of having a masculine orientation of sociological knowledge. Our framework doesn't make any assumptions about the complexity of explanations nor assumes that everything can be explained, but instead it rescues falsifiability of claims and replicability of the method as the basic demarcation criteria between what we call scientific knowledge and other kinds of knowledge. Without these features the process of science lacks a mechanism to purge itself and evolutionarily select the best available explanation and be prepared to revise itself as new evidence becomes available. Our framework defines an ontology from which the objects of study are available, but it doesn't impose a formal notion of reality in a deep philosophical sense, but instead, leverages that ontology along with the process of science as a framework for exploration, discovery and revision. While we do propose a demarcation of science, we do not advance any value judgement and recognize that knowledge generated or acquired outside the scope of science, can and do influence scientific inquiry. We also detach from any ideological background and focus only on the process.

Feminism

As women have been oppressed both physically and ideologically for the better part of modern civilization, suspicion of the establishment are natural reactions as we become more aware of the oppressing circumstances that they have been victims of through history, resulting on much of the discussion to naturally become political. In the words of Michelle Meagher: "feminist theory in the late twentieth and early twenty first centuries is not limited to thinking about the lives of men and women and all those in between, but rather, offers explanations of how gender shapes the entire social world." [16]. In this light, we argue that progress can be made beyond dialectics. A recent paper by Clifton [17] introduces a mathematical model that explores the role of bias and homophily in the progression of women in various professional hierarchies, suggesting policies to achieve gender parity. Another model by Zazueta and Accinelli [18] explores the evolution of the gender gap in the labor market by explicitly modeling cross-bias by gender. The authors' conclusions are falsifiable and their (modeling) method replicable. Their arguments are free from ideology and open to be challenged or supported, informing policy from a non-partisan point of view.

Language

Language is especially suited for scientific analysis and mathematical methods have been extensively applied to study its structure and to understand its messages with minimal human intervention. Applications include measuring political influence on Twitter [19], uncovering what the Chinese government is censoring online [20], identifying authorship based on style [21] or exploring the ideological structure of political parties [22]. As these examples illustrate, language modeling is a powerful research tool with applications well beyond language itself.

CONCLUSIONS

We propose a pragmatic research framework that sets minimal requirements for knowledge to be considered scientific and sketch an ontology to help us reflect on the potential objects of research that builds from Popper's original proposal, but refrains from giving reality an objective meaning. Not all objects within the ontology are knowable in a scientific sense, but those that are, can be studied through our framework. The process of science is a feedback loop that accumulates new knowledge, but also purges its content through falsifiability.

By demarcating what we believe is scientific knowledge, we are not making any value judgments and fully recognize that there are many legitimate avenues to knowledge that do not fit our definition of scientific. Furthermore, the creative process by which researchers structure new claims, is to a large extent non-scientific. Is only when we want to label a new claim as scientific, that we need to abide by the minimal rigor defined in this paper.

¹ The dynamics of falling objects is a bit more complicated than this, as they are accelerating towards earth according to Newton's laws. But, as we will point out, approximated claims are valid within the context of the problem at hand. Just as Newton's laws can be seen as an approximation to Einstein's special relativity.

² Typical depictions of Popper's ontology, show the three worlds as non-intersecting. Our depiction stresses the intersection between world 1 and world 3 because, in our mind, is a source of ambiguity.

³ We say man-created because we are the only intelligent species that we are aware of, but the same idea will apply to other intelligent beings from this or other planets, organic or mechanical.

⁴ A good review of the evolving theories of measurement is given in [10] which, incidentally, is a good example of the process of science as described in Figure 4.