



Mechanistic models in computational social science

Petter Holme^{1*} and Fredrik Liljeros²

¹ Department of Energy Science, Sungkyunkwan University, Suwon, South Korea, ² Department of Sociology, Stockholm University, Stockholm, Sweden

Quantitative social science is not only about regression analysis or, in general, data inference. Computer simulations of social mechanisms have an over 60 years long history. They have been used for many different purposes—to test scenarios, to test the consistency of descriptive theories (proof-of-concept models), to explore emergent phenomena, for forecasting, etc... In this essay, we sketch these historical developments, the role of mechanistic models in the social sciences and the influences from the natural and formal sciences. We argue that mechanistic computational models form a natural common ground for social and natural sciences, and look forward to possible future information flow across the social-natural divide.

Keywords: computational social science, mechanistic models, simulation, complex systems, interdisciplinary science

OPEN ACCESS

Edited by:

Yamir Moreno,
University of Zaragoza, Spain

Reviewed by:

Ladislav Kristoufek,
Institute of Information Theory and
Automation (AS CR), Czech Republic
Carlos Gracia-Lázaro,
University of Zaragoza, Spain

*Correspondence:

Petter Holme,
Department of Energy Science,
Sungkyunkwan University,
Suwon 440-746, South Korea
holme@skku.edu

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 30 June 2015

Accepted: 01 September 2015

Published: 17 September 2015

Citation:

Holme P and Liljeros F (2015)
Mechanistic models in computational
social science. *Front. Phys.* 3:78.
doi: 10.3389/fphy.2015.00078

Background

In mainstream empirical social science, a result of a study often consists of two conclusions. First, that there is a statistically significant correlation between a variable describing a social phenomenon and a variable thought to explain it. Second, that the correlations with other, more basic, or trivial, variables (called *control*, or *confounding*, variables) are weaker. There has been a trend in recent years to criticize this approach for putting too little emphasis on the mechanisms behind the correlations [1–3]. It is often argued that regression analysis (and the linear, additive models they assume) cannot serve as causal explanations of an open system such as usually studied in social science. A main reason is that, in an empirical study, there is no way of isolating all conceivable mechanisms [4]. Sometimes authors point to natural science as a role model in the quest for mechanistic models. This is somewhat ironic, since many natural sciences, most notably physics, traditionally put more emphasis on the unification of theories and the reduction of hypotheses [1]. In other words, striving to show that two theories could be more simply described as different aspects of a single, unified theory. Rather than being imported from natural or formal sciences, mechanistic modeling has evolved in parallel in the social sciences. Maybe the most clean-cut forms of mechanistic models are those used in computer simulations. Their past, present and future, and the flow of information regarding them across disciplines, are the themes of this paper. Before proceeding, other authors would probably spend considerable amounts of ink to define and discuss central concepts—in our case “mechanism” and “causal.” We think their everyday usage in both natural and social sciences is sufficiently precise for our purpose and recommend [3] to readers with a special interest of details.

In practice, establishing the mechanisms behind a social phenomenon takes much more than simulating a model. Mechanistic models can serve several different purposes en route to establishing a mechanistic explanation. We will make a distinction of *proof-of-concept modeling*, *discovery of hypotheses* and *scenario testing* (described in detail below). There are of course other

ways, perhaps also better, to characterize mechanistic models. These categories are not strict either—they could be overlapping with regard to a specific model. Nevertheless, we think they serve a point in our discussion and that they are fairly well defined.

The idea of proof-of-concept modeling is to test the consistency of a verbal description, or cartoon diagram, of a phenomenon [5]. It is in general hard to make an accurate verbal explanation, especially if it involves connecting different levels of abstraction, such as going from a microscopic to a macroscopic description. A common mistake is to neglect implicit assumptions, some that may even be the convention of a field. With the support of such proof-of-concept models, a verbal argument becomes much stronger. Then one has at least firmly established that the constituents of the theory are sufficient to explain the phenomenon. The individual-based simulations of the Anasazi people (inhabiting parts of the American West millennia ago) by Joshua Epstein, Robert Axtell and colleagues [6] are blueprints of proof-of-concept modeling. In these simulations, the authors combined a multitude of conditions along with anthropological theories to show that they could generate outcomes similar to the archeological records.

The most common use of mechanistic models is our second category—to explore the possible outcomes of a certain situation, and to generate hypotheses. We will see many examples of that in our essay. As a first example, consider Robert Axelrod's computer tournaments to find optimal strategies for the iterated prisoner's dilemma [7]. The prisoner's dilemma captures a situation where an individual can choose whether or not to cooperate with another. If one knows that the encounter is the last one, the rational choice is always not to cooperate. However, if the situation could be repeated an unknown number of times, then it might be better to cooperate. To figure out the way to cope with this situation, Axelrod invited researchers to submit strategies to a round-robin tournament. The winning strategy ("tit-for-tat") was to start cooperating and then do whatever your opponent did the previous step. From this result, Axelrod could make the hypothesis that a tit-for-tat-like behavior is common among both people and animals, either because they often face a prisoner's dilemma or at that such situations, once you face them, tend to be important.

Mechanistic models forecasting social systems are less frequent than our previous two classes. One reason is probably that forecasting open systems is difficult (sometimes probably even impossible) [4]; another that non-mechanistic methods (machine learning, statistical models, etc...) are better for this purpose. A model without any predictive power whatsoever is, of course, not a model at all, and under some conditions all mechanistic models can be used in forecasting, or (perhaps more accurately) scenario testing. One celebrated example is the "World3" simulation popularized by the Club of Rome 1972 book *The Limits to Growth* [8] where an exponentially growing artificial population faced a world of limited resources. Maybe a sign of the time, since several papers from the early 1970s called for "whole Earth simulations" [9, 10]. Echoes of this movement were heard recently with the proposal of a "Living Earth Simulator" [11].

In this essay, we will explore mechanistic models as scientific explanations in the social sciences. We will give an overview of the development of computer simulations of mechanistic models (primarily in the social sciences, but also mentioning relevant developments in the natural sciences), and finally discuss if and how mechanistic models can be a common ground for cross-disciplinary research between the natural and social sciences. We do not address data-driven science in the interface of the natural or social, nor do we try to give a comprehensive survey of mechanistic models in the social sciences. We address anyone interested in using simulation methods familiar to theoretical natural scientists to advance the social sciences.

Influence from the Natural and Formal Sciences

As we will see below, the development and use of computer simulations to understand social mechanisms has happened on quite equal terms as in the natural and formal sciences. It will, however, be helpful for the subsequent discussion to sketch the important developments of computer simulations as mechanistic models in the natural sciences. This is of course a topic that would need several book volumes for a comprehensive coverage—we will just mention what we regard the most important breakthroughs.

The Military Origins

Just like in social science, simulation in natural science has many of its roots in the military from the time around the Second World War. The second major project running on the first programmable computer, ENIAC, started April 1947. The topic, the flow of neutrons in an incipient explosion of a thermonuclear weapon [12], is perhaps of little interest today, but the basic method has never ran out of fashion—it was the first computer program using (pseudo) random numbers, and hence an ancestor of most modern computer simulations. Exactly who invented this method, codenamed Monte Carlo, is somewhat obscure, but it is clear it came out of the development of the hydrogen bomb right after the war. The participants came from the (then recently finished) Manhattan project. Nicholas Metropolis, Stanislaw Ulam and John von Neumann are perhaps most well-known, but also Klara von Neumann, John's wife [12]. It was not only the first program to use random numbers, it was also the first modern program in the sense that it had function calls, and had to be fed into the computer along with the input. As a curiosity, the random number generator in this program worked by squaring eight-digit numbers and using the mid eight digits as output and seed to the next iteration. Far from having the complexity of modern pseudo random number generator (read Mersenne Twister [13]), it gives random numbers of (at least in the authors' opinion) surprisingly good statistical quality.

The first Monte Carlo simulation was not an outright success as a contribution to the nuclear weapons program. Nevertheless, the idea of using random numbers in simulations has not fallen out of fashion ever since, and the Monte Carlo method (nowadays referring to any computational model based on random numbers) has become a mainstay of numerical methods.

Another very significant step for the natural sciences, especially chemistry and statistical physics, by the Los Alamos group was the Metropolis–Hastings algorithm—a method to sample configurations of particles, atoms or molecules according to the Boltzmann distribution (connecting the probability of a configuration and its energy). The radical invention was to choose configurations with a probability proportional to the Boltzmann distribution and weighting them equally, rather than choosing configurations randomly and weighing them by the probability given by the Boltzmann distribution [14]. Hastings name was added to credit his extension of the algorithm to general distributions [15]. Today, this algorithm is an indispensable simulation technique to generate the probability distributions of the state of a system both in natural and social sciences (usually called Markov Chain Monte Carlo, MCMC).

The Monte Carlo project and the MCMC method did not immediately lead to fundamental advances in science itself. Deterministic computational methods, on the other hand, did, and (not surprisingly) post-Manhattan-project researchers were involved. Enrico Fermi, John Pasta, and Stanislaw Ulam (and, like the Monte Carlo project, with undercredited help by a female researcher, Mary Tsingou [16]) studied vibrations of a one-dimensional string with non-linear corrections to Hooke's law (that states that the force needed to extend a spring a certain distance is proportional to the distance). They expected to see the non-linearity transferring energy from one vibrational mode (like the periodic solution of the linear problem) to all other modes (i.e., thermal fluctuations) according to the equipartition theorem [17]. Instead of such a “thermalization” process, they observed the transition to a complex, quasi-periodic state [18] that never lost its memory of the initial condition. The FPU paradox was the starting point of a scientific theme called *non-linear science* that also, as we will see, has left a lasting imprint on social science.

Complexity Theory

Non-linear science has a strong overlap with chaos theory, another set of ideas from natural sciences that influenced social science. Chaos is summarized in the vernacular by the “butterfly effect”—a small change (the flapping of a butterfly's wings) could lead to a big difference (a storm) later. One important early contribution came from Edward Lorenz's computational solutions of equations describing atmospheric convection. He observed that a small change in the initial condition could send the equations off into completely different trajectories [19]. Just like for the FPU paradox, the role of the computational method in chaos theory has largely been to discover hypotheses that later have been corroborated by analytical studies. This line of research has not been directly aimed at discovering new mechanisms; still, ideas and concepts from chaos theory have also reached social sciences [20].

Another natural science development largely fueled by computer simulations, which has influenced social sciences, is that of fractals. Fractals are mathematical objects that embody self-similarity—a river can branch into tributaries, that branch into smaller tributaries, and so on, until the biggest rivers are reduced to the tiniest creeks [21]. At all scales, the branching looks the same. Fractals provide an

analysis tool—the fractal dimension—that can characterize self-similar objects. There are many socioeconomic systems that are self-similar—financial time series [22], the movement of people [23], the fluctuations in the size of organizations [24], etc... Quite frequently, however, authors have not accompanied their measurement of a fractal dimension with a mechanistic explanation of it, which is perhaps why fractals have fallen out of fashion lately.

Fractals are closely related to power-law probability distributions, i.e., the probability of an observable x being proportional to $x^{-\alpha}$, $\alpha > 0$. Power-laws are the only self-similar (or “scale-free”) real-to-real functions in the sense that, if e.g., the wealth distribution of a population is a power law, then a statement like “there are twice as many people with a wealth of 10X than 15X” is true, no matter if X is dollars, euros, yen or kronor [25]. The theories for such power-law phenomena date back to Pareto's lectures on economics published 1896 [26]. Fractals and power-laws are also connected to phase transitions in physics—an idea popularized in Hermann Haken's book *Synergetics* [27].

Next step in our discussion is the studies of artificial life. The central question in this line of research is to mechanistically recreate the fundamental properties of a living system, including self-replication, adaptability, robustness and evolution [28]. The origins of artificial life can be traced to John von Neumann's self-replicating cellular automata. These are configurations of discrete variables confined to an underlying square grid that, following a distinct set of rules, can reproduce, live and die [29]. The field of artificial life later developed in different directions, both toward the more abstract study of cellular automata and to more biology-related questions [28]. It is also strongly linked to the study of adaptive systems (systems able to respond to changes in the environment) [30] and has a few recurring ideas that also are related to social phenomena. The first idea is that simple rules can create complex behavior. The best-known model illustrating this is perhaps Conway's game of life. This is a cellular automaton with the same objectives as that of von Neuman, but with fewer and simpler rules [28]. The second idea (maybe not discovered by the field of artificial life, but at least popularized) is that of *emergence*. This refers to the properties of a system, as a whole, coming from the interaction of a large number of individual subunits. A textbook example is that of murmurations of birds (flocks of hundreds of thousands of e.g., starlings). These can exhibit an undulating motion, fluctuating in density, that in no way could be anticipated from the movement of an individual. Another feature of emergence, exemplified by bird flocks, is that of *decentralization*—there is no leader bird. These topics are common to many disciplines of social science (emergence is similar to the micro-to-macro-transition in sociology and economics). These theories have spawned its own modeling paradigm—agent based models [31–34]—that is similar to what was simply called “simulation” in early computational social science. One first sets up rules for how units (agents) interact with each other and their surroundings. Then one simulates many of them together (typically on a two-dimensional grid) and let them interact. We note that the concept of emergence has also been influential to cognitive, and subsequently behavioral, science. The

idea of cognitive processes being emergent properties of neural networks—connectionism [35]—is nowadays fundamental to our understanding of computational processes in nature [36].

In the 1980's, artificial life, adaptive systems, fractals and chaos were grouped together under the umbrella term complexity science [37]. This was in many ways a social movement gathering researchers of quite marginalized research topics (the Santa Fe Institute, and some similar centers, acted as hubs for this development). Many of the themes within complexity science could probably just as well be categorized as mutually independent fields. This is perhaps best illustrated in that there is no commonly accepted definition of “complexity.” Instead, there are a number of common, occasionally (but not always) connected, themes (like the above-mentioned, emergence, decentralized organization, fractals, chaos, etc...) that together defines the field. On the other hand, there is a common goal among complexity scientists to find general, organizational principles that are not limited to one scientific field. In spirit, this dates back to, at least, von Bertalanffy's general systems theory [38]. The diversity of ideas and applications has not necessarily been a problem for complexity science; on the contrary, it has encouraged many scientists of different backgrounds (including the authors of this paper) to try collaborating, despite the transdisciplinary language barriers.

Game Theory

Game theory is a mathematical modeling framework for situations where the state of an individual is jointly determined by the individual's own decisions and the decisions of others (who all, typically, strive to maximize their own benefit) [39]. Vaccination against infectious diseases is a typical example. If everyone else were vaccinated, the rational choice would be to not get vaccinated. The disease could anyway not spread in the population, whether or not you are vaccinated. Moreover, vaccines can, after all, have side effects, and injections are uncomfortable. If nobody were vaccinated, and the chance of getting the disease times the gravity of the consequences outweighs the above-mentioned inconveniences, then it would be rational to get vaccinated. This situation could, mathematically, be phrased as a minority game [40]. The emergent solution for a population of rational, well-informed and selfish individuals is that a fraction of the agents would get vaccinated and another fraction not. This example is, at the time of writing, the background to a controversy where people getting vaccinated see people resisting vaccination as irresponsible to the society [41].

Game theory has been an especially strong undercurrent in economy and population biology. We note that a special feature of game theory, compared to similarly interdisciplinary theories, is that the various fields using it seem rather well informed about the other fields' progress and not so many concepts have been reinvented. Game theory itself is not a framework for mechanistic models, and especially in population biology (where an individual usually represents a species or a sub-population) it is not clear that is its main use. Nevertheless, there are many mechanistic models in economy and population biology that uses game theory as a fundamental ingredient [42].

Network Theory

Just like complexity and game theory, network theory is a great place for information exchange between the natural and social sciences. Its basic idea is to use networks of vertices, connected pairwise by edges, as a systematic way of simplifying a system. By studying the network structure (roughly speaking, how a network differs from a random network) one can say something about how the system functions as a whole, or the roles of the individual vertices and edges in the system [43, 44]. The multidisciplinary nature of network theory is reflected in its overlapping terminology—vertices and edges are called nodes and links in computer science, sites and bonds in physics and chemistry, actors and ties in sociology, etc...

Many ideas in network theory originated in social science, and for that reason it may not fit in a section about influences from natural science. Nevertheless, as mentioned, it is a field where ideas frequently flow from the natural and formal sciences to social sciences. Centrality measures like PageRank and HITS were, for example, developed in computer science [43], as were fundamental concepts of temporal network theory (where information about the time when vertices and edges are active is included in the network) [45].

Early Computer Simulations to Understand Social Mechanisms

In this section, we will go through some developments in the use of mechanistic models in social science. We will focus on early studies, assuming the readers largely know the current trends. This is by no means a review (which would need volumes of books), but a few snapshots highlighting some differences and similarities to today's science in the methodologies and the questions asked.

Operations Research

Just like the computer hardware, the research topics for simulation and mechanistic models have many roots in military efforts around the Second World War. Perhaps the main discipline for this type of research is operations research, which is usually classified as a branch of applied mathematics. The objective of operations research is to optimize the management of large-scale organizations—managing supply chains, scheduling crews of ships, planes and trains, etc... The military was not the only such organization that interested the early computer simulation researchers. Harling [46] provides an overview of the state of computer simulations in operation research in the late 1950's. As a typical example, Jennings and Dickins modeled the flow of people and buses in the Port Authority Bus Terminal in New York City during the morning rush hour [47]. They modeled the buses individually and passengers as numbers of exiting, not transferring, individuals. The authors tried to simultaneously optimize the interests of three actors—the bus operators, the passengers, and the Port Authority (operating the terminal). These objectives were mostly not conflicting—in principle it was better for all if the passenger throughput was as high as possible. A further simplifying factor was that the station was the terminus for all buses. The challenge was that

buses stopping to let off passengers could block other buses, thus creating a traffic jam. To solve this problem, the paper evaluated different methods to assign a bus stop to an incoming bus.

Political Science

Although rarely cited today, simulation studies of political decision processes were quite common in the 1950s and 1960s. Crecine [48] reviews some of these models. One difference from today is that these models were less abstract, often focusing on a particular political or juridical organization. The earliest paper we are aware of is Guetzkow's 1959 investigation of the use of computer simulations as a support system for international politics [49]. However, many studies in this field credit de Sola Pool et al.'s simulation of the American presidential elections 1960 and 1964 as the starting point [50]. In their work, the authors gathered a collection of 480 voter profiles that they could use to test different scenarios (with respect to what topics that would turn out to be important for the campaign). Eventually they predicted the outcome of the elections with 82% accuracy.

In their Ph.D. theses, Cherryholmes [51] and Shapiro [52] modeled voting in the House of Representatives by: First, dividing members into classes with respect to how susceptible they were to influence. Second, modeling the influence process via an interaction network where people were more likely to communicate (and thus influence each other) if they were from the same party, state, committee, etc...Cherryholmes and Shapiro also validated their theories against actual voting behavior (something rarely seen in today's simulation studies of opinion spreading [53]). Other authors addressed more theoretical issues of voting systems, such as Arrow's paradox [54, 55] (which states, briefly speaking, that a perfect voting system is impossible for three or more alternatives).

There was also a considerable early interest in simulating decision making within an organization. Apparently the Cuban missile crisis of 1962 was an important source of inspiration. De Sola Pool was, once again, a pioneer in this direction with a simulation of decision-making in a developing, general crisis with incomplete information [56]. Even more explicitly, Smith [57] based his simulation on the personal accounts of the people involved in solving the Cuban missile crisis. Clema and Kirkham proposed yet a model of risks, costs and benefits in political conflicts [58]. Curiously, as late as 2007 there was a paper published on simulating the Cuban missile crisis [59]. However, this paper explores mechanistic modeling as a method of teaching history, rather than the mechanisms of the decision making process itself.

Another type of political science research concerns the evolution of norms. A classic example is Axelrod's 1986 paper [60] where he investigated norms emerging as successful strategies in situations described by game theory. Axelrod let the norms evolve by genetic algorithms (an algorithmic framework for optimization inspired by genetics). In addition to norms, Axelrod also studied metanorms—norms that promote other norms (by e.g., encouraging punishing of people breaking or questioning the norms). Axelrod interpreted the results of the simulation in terms of established social mechanisms supporting

the existence of norms (domination, internalization, deterrence, etc...).

Linguistics

In linguistics, the first computer simulation studies appeared in the mid-1960s. A typical early example is Klein [61] who developed an individual-based simulation platform for the evolution of language. Just like Cherryholmes and Shapiro (above), Klein assumed that the communication was not uniformly random between all pairs of individuals—spouses were more likely to speak to, and learn from, one another, as were parents and children. In multilingual societies, speakers were more likely to communicate to another speaker of the same language (Klein allowed multilingual individuals). A language was represented by a set of explicit grammatical rules (with explicit word classes: nouns, verbs, etc...). Communication reinforced the grammatical rules between the speakers. Klein incremented the time by years and simulated several generations of speakers, but was not entirely happy with the results as communities tended to lose the diversity of their grammar quickly or diverge to mutually incomprehensible grammars. In retrospect, we feel like it was a still a great step forward, where the negative results helped raising important questions about what mechanisms that were missing. More modern models of language evolution have considered much simpler problems [62]. One cannot help thinking that this is to avoid the complexities of reality, and more models in the vein of Klein's 1966 paper would be more important. Later, Klein focused his research on more specific questions like the evolution of Tikopia and Maori [63]. The goal of these early simulation studies was to create something similar to a sociolinguistic fieldwork study. Thus, these were proof-of-concept studies on a more concrete level than today's more theoretically motivated research.

Geography

Demography and geography were also early fields to adopt computer simulations. One notable pioneer was the authors' compatriot Torsten Hägerstrand whose Ph.D. thesis used computer simulations to investigate the diffusion of innovations [64]. His model was similar to two-dimensional disease-spreading models (but probably developed independently of computational epidemiology, where the first paper was published the year before [65]). Hägerstrand used an underlying square grid. People were spread out over the grid according to an empirically measured population distribution. At each iteration of the simulation, there was a contact between two random individuals (where the chance of contact decayed with their separation). If the one of the individuals had adopted the innovation, and the other had not, then the latter would (with 100% probability) adopt it. A goal of Hägerstrand's modeling was to recreate a "nebula shaped" distribution of the innovation (this is further developed in Hägerstrand [66]). To this end, Hägerstrand introduced a concept (still in use) called *mean information field* representing the probability of getting the information (innovation) from the source.

A technically similar topic to information diffusion is that of migration (as in moving one's home). This research dates

back to Ravenstein's 1885 paper "The laws of migration" which is very mechanistically oriented [67]. He listed seven principles for human migration such as: short-distance migration is more common than long-distance; people who migrate far have a tendency to go to a "great centre of commerce or industry." Computer simulation lends itself naturally to exploring the outcomes such mechanisms in terms of demographics. One such example is Porter's migration model where agents were driven by the availability of work and the availability of work was partly driven by where people were. If there was an excess of workers, workers would move to the closest available job opportunity; if there was an excess of vacancies, the closest applicant would be offered the job [68].

The study of human mobility (how people move around both in their everyday lives and extreme situations, such as disasters) is an active field of research. It has even been revitalized lately by the availability of new data sources (see e.g., [23]). One common type of simulation study, involving human mobility data, aims at predicting outbreaks of epidemic diseases. To model potentially contagious contacts between people, one can use more or less realism. However, even for the most realistic and detailed simulations, there is a choice of using the real data to calibrate a model of human mobility [69] or run the simulation on actual mobility data (perhaps with simulations to fill in missing data) [70].

Economics and Management Science

There were many early computational studies in economics that used simulation techniques for scenario testing [71, 72]. A typical question was to investigate the operations of a company at many levels (overlapping with the operations-research section above). Evidently, the researchers saw a future where every aspect of running a business would be modeled—marketing, human resource development, social interaction within the company, the competition with other firms, adoption of new technologies, etc... To make progress, the authors needed to restrict themselves. Birchmore [72], for example, focused on forest firms. Much of his work revolved around a forestry firm's interaction with its resource and the many game theoretical considerations that arise from the conflicting time perspectives of short- and long-time revenues and the competition with other companies. Birchmore only used one or a few combinations of parameter values, rather than investigating the parameter dependence like modern game theory would do. Finally, we note that economics and management science were also early to address questions about validation and other epistemological aspects of computer simulations [73].

Anthropology and Demographics

Anthropology was also early to embrace simulation techniques, especially to problems relating to social structure, kinship and marriage [74]. These are perhaps the traditional problems of anthropology that has the most complex structure of causal explanations, and for that reason are most in need for proof-of-concept-type computer simulations. Gilbert and Hammel [75], for example, addressed the question: "How much, and in what ways, is the rate of patrilineal parallel cousin marriage influenced

by the number of populations involved in the exchange of women, by their size, by their rules of postmarital residence, and by degree of territorially endogamic preference?" To answer these questions, the authors constructed a complex model including villages of explicit sizes, individuals of explicit gender, age and kinship, and rules for how to select a spouse. The model was described primarily in words, in much detail and length. A modern reader would think that pseudocode would make the paper more readable (and certainly much shorter). Probably the anthropology journals of the time were too conservative, or the programming literacy too low, for including pseudocode in the articles.

In a study similar to Gilbert and Hempel, one step closer to demographics, May and Heer [76] used computer simulations to argue that the large family sizes in rural India (of that time) were rational choices for the individuals, rather than a consequence of ignorance and indecision. Around the same time, there were studies of more general questions of human demographics [77], highlighting a transition from mechanistic models for scenario testing to proof-of-concept models and hypothesis discovery.

Cognitive and Behavioral Science

In cognitive science (sometimes bordering to behavioral science), researchers in the 1960s were excited about the prospects of understanding human cognition as a computer program.

Abelson and Carroll [78], for example, proposed that mechanistic simulations could address questions like how a person can reach an understanding ("develop a belief system") of a complex situation in terms of a set of consistent descriptive clauses (encoding, for example causal relationships). Several researchers proposed reverse engineering of human thinking into computer programs as a method to understand cognitive processes [79]. Some even went so far as to interpret dreams as an operating system process [80]. These ideas were not without criticism. Frijda [81] argued that there would always be technical aspects of computer code without a corresponding cognitive function. History seems to give the author right since few studies nowadays pursue replicating human thinking by procedural computer programs. There were of course many other types of studies in this area. For example, early studies in computational neuroscience influenced the behavioral-science side of cognitive science [82].

Sociology

Simulation, in sociology, has always been linked to finding social mechanisms. Even before computer simulations, there were mathematical models for that purpose [83, 84]. As an example of mathematical model building, we briefly mention Nicholas Rashevsky and his program in "mathematical biophysics" at University of Chicago [85, 86]. Trained as a physicist, Rashevsky and his group pioneered the modeling of many social (and biological) phenomena such as social influence [87], how social group structure affect information flow [88], and fundamental properties of social networks [89]. However, Rashevsky and colleagues operated rather disconnected from the rest of academia—mostly publishing in their *Bulletin of Mathematical Biophysics* and often not building on empirical results available.

Perhaps for this reason (even though his contemporaries were aware of his work [90]) is Rashevsky et al.'s direct impact on today's sociology rather limited.

Even though there were stochastic models in sociology in the early 1960's (e.g., [91]), these were analyzed analytically and early sociological computer simulations were off to a rather late start. Coleman [92], Gullahorn and Gullahorn [93, 94] gave the earliest discussions of the prospects of computer modeling in sociology that we are aware of. Coleman discussed both abstract questions about relating social action and social organization, and more concrete ones like using simulation to test social-contagion scenarios of smoking among adolescents. The Gullahorns were more interested in organization and conflict resolution, typically in the interface of sociology and behavioral science. McGinnis [95] presented a stochastic model of social mobility that he analyzed both analytically and by simulations. "Mobility," in McGinnis work, should be read in an extremely general sense, indicating change of an individual's position in any sociometric observable (including physical space).

Markley's 1967 paper on the SIVA model is another early simulation study of a classic sociological problem [96], namely what kind of pairwise relationships that could build up a stable organization. The letters SIVA stands for four aspects of such relationships in an organization facing some situation that could require some action to be taken—Strength (the ratio of how important the two individuals are to the organization), Influence (describing how strongly they influence each other), Volitional (the relative will to act with respect to the situation) and Action (quantifying the joint result of the two actors). These different aspects are coupled and Markley used computer simulations to find fixed points of the dynamics. For many parameter values, it turned out that the SIVA values diverged or fluctuated—which Markley took as an indication that one would not observe such combinations of parameter values in real organizations.

A model touching classical sociological ground that recently has received exceptional amounts of attention is Schelling's segregation model [97]. With this model, Schelling argued that a strong racial segregation (with the United States in mind) does not necessarily mean that people have very strong opinions about the race of their neighbors. Briefly, Schelling spread individuals of two races on a square grid. Some sites were left vacant. Then he picked an individual at random. If this individual had a lower ratio of neighbors of the same race than a threshold value, then he or she moved to a vacant site. It turned out that the segregation (measured as the fraction of links between people of the same race) would always move away from threshold as the iterations converged. Segregation, Schelling concluded, could thus occur without people actively avoiding different races (they just needed to seek similar neighbors), and spatial effects would make a naïve interpretation of the observed mixing overestimating the actual sentiments of the people. The core question—what are the weakest requirements (of tolerance to your neighbors ethnicity) for something (racial segregation) to happen—was a hallmark of Schelling's research and probably an approach that could be fruitful for future studies. We highly recommend Schelling's popular science book *Micromotives and Macrobehavior* [98] as a bridge between the methodologies of natural and social science.

Discussion and Conclusions

The motivation for the use of mechanistic models in social science is often to use them as proof-of concept models. "[I]t forces one to be specific about the variables in interpersonal behavior and the exact relation between them" [93, 99, 100]. The way computer programming forces the researchers to break down the social phenomena into algorithmic blocks helps identifying mechanisms [93, 101]. Other authors point out that with computational methods, the researchers can avoid oversimplifying the problem [50]. Another point of view is that simulation in social sciences is primarily for exploring poorly understood situations and phenomena as a replacement for an actual (in practice impossible to carry out) experiment [48, 102–104]. Such models are obviously closest to hypothesis generators in our above classification. Crane [105] and Ostrom [106] think of computer simulations that, alongside natural languages and mathematics, could describe social sciences. Going a bit off topic, other authors went so far as to using, or recommending to use, computer programs as representations of human cognitive processes [79, 80, 107].

The history of computational studies in social science—as illustrated by our examples—has seen a gradual shift of focus. In the early days, it was, as mentioned, often regarded as a replacement for empirical studies. Such mechanistic models for scenario testing still exists in both natural and social science. However, nowadays it is much more common to use computational methods in theory building—either one uses it to test the completeness of a theoretical framework (proof-of-concept modeling), or to explore the space of possible mechanisms or outcomes (hypothesis discovery).

It is quite remarkable how similar this development has been in the natural and social sciences. At least since mid-1950s, it is hard to say that one side leads the way. This is reflected in how the information flows between disciplines. Looking at the interdisciplinary citation patterns [108] found that out of 203,900 citations from social science journals, 33,891 were to natural science journals, and out of 10,080,078 citations from natural science journals 35,199 were to social science journals. If citations were random, without any within-field bias there would be around 201,000 interdisciplinary citations in both directions, which is 5.9 times the number of social science citations to natural science and 5.7 times the number of natural science citations to social science. In this view, there is almost no inherent asymmetry in the information flow between the areas, only an asymmetry induced by the size difference.

Even though social scientists do not need to collaborate with natural scientists to develop mechanistic modeling, we do encourage collaboration. The usefulness of interdisciplinary collaborations comes from the details of the scientific work. It can help people to see their object system with new eyes. One discipline may, for example, care about the extreme and need input from another to see interesting aspects of the average (cf. phase transitions in the complexity of algorithms [109]). Interdisciplinary information flow could help a discipline overcome technical difficulties. The use of MCMC techniques in the social sciences may be a good example of this. It is,

however, important that such developments come from a need to understand the world around us and not just because they have not been done before.

A major trend at the time of writing is “big data” and “data science.” This essay has intentionally focused on the other side of computational social science—mechanistic models. In practice, these two sides can (and do) influence each other. If it cannot predict real systems at all, a mechanistic model is quite worthless in providing a causal explanation [110, 111]. Modern, large-scale data sets provide plenty opportunities to validate models [112–114]. Another use of big data is in hybrid approaches where one combines a simulation and an empirical dataset, for example simulations of disease spreading on temporal networks of human contacts [45].

As a concluding remark, we want to express our support for social scientists interested in exploring the methods of natural science and natural scientists seeking applications for their methods in the social sciences. To be successful and make most out of such a step, we recommend the social scientist to spend a month to learn a general programming language (Python, Matlab, C, etc...). There is not shortcut (like an integrated modeling environment) to learning the computational subtleties and trade-offs of building a simulation model, and simulation papers often do not mention them. Furthermore, if a social scientist leaves this aspect to a natural scientist,

then she also leaves parts of the social modeling to the natural scientist—collaboration simply works better if the computational fundamentals need not be discussed. To the theoretical natural scientists that are used to simulations, we recommend spending a month reading popular social science books (e.g., [98, 102, 115]). There are too many examples of natural scientists going into social science with the ambition to use the same methods as they are used to—only replacing the natural components by social—and ending up with results that are unverifiable, too general to be interesting, infeasible or already known. While reading, we encourage meditating the following question—why do social scientists ask different questions about society than natural scientists do about nature?

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2011947).

Acknowledgments

The authors thank Martin Rosvall for the citation statistics.

References

- Woodward J. Scientific explanation. In: Zalta EN, editor. *The Stanford Encyclopedia of Philosophy*. Available online at: <http://plato.stanford.edu/archives/win2014/entries/scientific-explanation> (2014). doi: 10.1093/acprof:oso/9780195145649.003.0002
- Salmon W. *Causality and Explanation*. Oxford, UK: Oxford University Press (1998). doi: 10.1093/0195108647.001.0001
- Hedström P, Ylikoski P. Causal mechanisms in the social sciences. *Annu Rev Sociol.* (2010) **36**:49–67. doi: 10.1146/annurev.soc.012809.102632
- Sayer A. *Realism and Social Science*. New York, NY: Sage (2000). doi: 10.4135/9781446218730
- Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, Stern CA, et al. Not just a theory: the utility of mathematical models in evolutionary biology. *PLoS Biol.* (2014) **12**:e1002017. doi: 10.1371/journal.pbio.1002017
- Epstein JM, Axtell R. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, DC: Brookings Institution Press (1996).
- Axelrod R. *The Evolution of Cooperation*. New York, NY: Basic Books (1984).
- Meadows DH, Meadows DL, Randers J, Behrens WW. *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. New York, NY: Universe Books (1972).
- Patterson P. World simulation: a logical extension. *Simulation* (1970) **15**:63–4. doi: 10.1177/003754977001500211
- Rau D. World simulation: the need, the feasibility, and a way to start. *Simulation* (1970) **15**:64–5. doi: 10.1177/003754977001500212
- Paulucci M, Kossman D, Conte R, Lukowicz P, Argyrakakis P, Blandford A, et al. Towards a living Earth simulator. *Eur Phys J Spec Top.* (2012) **214**:77–108. doi: 10.1140/epjst/e2012-01689-8
- Haigh T, Priestley M, Rope C. Los Alamos bets on ENIAC: nuclear Monte Carlo simulations, 1947–1948. *IEEE Ann Hist Comput.* (2014) **36**:42–63. doi: 10.1109/MAHC.2014.40
- Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans Model Comput Simul.* (1998) **8**:3–30. doi: 10.1145/272991.272995
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* (1953) **21**:1087–92. doi: 10.1063/1.1699114
- Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* (1970) **57**:97–109. doi: 10.1093/biomet/57.1.97
- Dauxois T. Fermi, Pasta, Ulam, and a mysterious lady. *Phys Today* (2008) **6**:55–7. doi: 10.1063/1.2835154
- Landau LD, Lifshitz EM. *Statistical Physics, Part 1*. 3rd ed. Oxford, UK: Pergamon Press (1980).
- Fermi E, Pasta J, Ulam S. *Studies of Nonlinear Problems. I*. Los Alamos National Laboratory, Los Alamos, NM (1955). Los Alamos Report LA-1940.
- Lorenz EN. Deterministic nonperiodic flow. *J Atmos Sci.* (1963) **20**:130–41.
- Kiel LD, Elliott EW (eds.). *Chaos Theory in the Social Sciences: Foundations and Applications*. Ann Arbor, MI: University of Michigan Press (1996).
- Mandelbrot BB. *The Fractal Geometry of Nature*. London, UK: Macmillan (1983).
- Mantegna RN, Stanley HE. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, UK: Cambridge University Press (1999). doi: 10.1017/CBO9780511755767
- Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. *Nature* (2006) **439**:462–5. doi: 10.1038/nature04292
- Mondani H, Holme P, Liljeros F. Fat-tailed fluctuations in the size of organizations: the role of social influence. *PLoS ONE* (2014) **9**:e100527. doi: 10.1371/journal.pone.0100527
- Newman MEJ. Power laws, Pareto distributions and Zipf's law. *Contemp Phys.* (2005) **46**:323–51. doi: 10.1080/00107510500052444
- Pareto V. *Cours d'Économie Politique. Professe à l'Université de Lausanne*. Lausanne: F. Rouge (1896).
- Haken H. *Synergetik*. Berlin: Springer (1982). doi: 10.1007/978-3-642-96663-7
- Langton CG. *Artificial Life: An Overview*. Cambridge, MA: MIT Press (1998).
- von Neumann J. *Theory of Self-Reproducing Automata*. Urbana, IL: University of Illinois Press (1966).

30. Miller JH, Page SE. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press (2009).
31. Šalamon T. *Design of Agent-Based Models: Developing Computer Simulations for a Better Understanding of Social Processes*. Repin: Bruckner Publishing (2011).
32. Epstein JM. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press (2006).
33. Carley KM, Wallace WA. *Computational Organization Theory*. Berlin: Springer (2001). doi: 10.1007/1-4020-0611-x_143
34. Hedström P, Manzo P. Recent trends in agent-based computational research: a brief introduction. *Sociol Methods Res.* (2015) **44**:179–85. doi: 10.1177/0049124115581211
35. Dawson MRW. *Minds and Machines: Connectionism and Psychological Modeling*. Hoboken, NJ: John Wiley & Sons (2008).
36. Flake GW. *The Computational Beauty of Nature*. Cambridge, MA: MIT Press (1988).
37. Mitchell M. *Complexity: A Guided Tour*. Oxford, UK: Oxford University Press (2011).
38. von Bertalanffy L. *General System Theory: Foundations, Development, Applications*. New York, NY: George Braziller (1968).
39. Hofbauer J, Sigmund K. *Evolutionary Games and Population Dynamics*. Cambridge, UK: Cambridge University Press (1998). doi: 10.1017/CBO9781139173179
40. Challet D, Marsili M, Zhang YC. *Minority Games: Interacting Agents in Financial Markets*. Oxford, UK: Oxford University Press (2005).
41. Honigsbaum M. Balancing unreason: vaccine myths and metaphors. *Lancet* (2015) **385**:763. doi: 10.1016/S0140-6736(15)60423-8
42. Rasmussen E. *Games and Information: An Introduction to Game Theory*. Hoboken, NJ: John Wiley & Sons (1989).
43. Newman MEJ. *Networks: An Introduction*. Oxford, UK: Oxford University Press (2010). doi: 10.1093/acprof:oso/9780199206650.001.0001
44. Barabási A-L. *Network Science*. Cambridge, UK: Cambridge University Press (2015).
45. Holme P, Saramäki J. Temporal networks. *Phys Rep.* (2012) **519**:97–125. doi: 10.1016/j.physrep.2012.03.001
46. Harling J. Simulation techniques in operations research: a review. *Oper Res.* (1958) **6**:307–19. doi: 10.1287/opre.6.3.307
47. Jennings NH, Dickins JH. Computer simulation of peak hour operations in a bus terminal. *Manage Sci.* (1958) **5**:106–20. doi: 10.1287/mnsc.5.1.106
48. Crecine JP. Computer simulation in urban research. *Public Adm Rev.* (1968) **28**:66–77. doi: 10.2307/973586
49. Guetzkow H. A use of simulation in the study of inter-nation relations. *Behav Sci.* (1959) **4**:183–91. doi: 10.1002/bs.3830040302
50. De Sola Pool I, Abelson RP, Popkin S. *Candidates, Issues, and Strategies: A Computer Simulation of the 1960 And 1964 Presidential Elections*. Cambridge, MA: The MIT Press (1965).
51. Cherryholmes CH. *The House of Representatives and Foreign Affairs: A Computer Simulation of Roll Call Voting*. Ph.D. thesis, Northwestern University (1966).
52. Shapiro MJ. *The House and the Federal Role: A Computer Simulation of Roll-Call Voting*. Ph.D. thesis, Northwestern University (1966).
53. Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. *Rev Mod Phys.* (2009) **81**:591–646. doi: 10.1103/RevModPhys.81.591
54. Klahr D. A computer simulation of the paradox of voting. *Am Polit Sci Rev.* (1966) **60**:384–90. doi: 10.2307/1953365
55. Tullock G, Campbell CD. Computer simulation of a small voting system. *Econ J.* (1970) **80**:97–104. doi: 10.2307/2230441
56. Kessler AR, de Sola Pool I. Crisiscom: a computer simulation of human information processing during a crisis. *IEEE Trans Syst Sci Cybern.* (1965) **1**:52–58. doi: 10.1109/TSSC.1965.300061
57. Smith RB. Presidential decision-making during the Cuban missile crisis: a computer simulation. *Simul Gaming* (1970) **1**:173–201.
58. Clema J, Kirkham J. CONSIM (Conflict Simulator): risk, cost and benefit in political simulations. In: *ACM '71 Proceedings of the 1971 26th Annual Conference*. New York, NY (1971). p. 226–35. doi: 10.1145/800184.810488
59. Stover WJ. Simulating the Cuban missile crisis: crossing time and space in virtual reality. *Int Stud Perspect.* (2007) **8**:111–20. doi: 10.1111/j.1528-3585.2007.00272.x
60. Axelrod R. An evolutionary approach to norms. *Am Polit Sci Rev.* (1986) **80**:1095–111. doi: 10.2307/1960858
61. Klein S. *Dynamic Simulation of Historical Change in Language Using Monte Carlo Techniques*. Technical Report SP-1908. Santa Monica, CA: System Development Corporation (1964).
62. Perfors A. Simulated evolution of language: a review of the field. *J Artif Soc Soc Simul.* (2002) **5**:4. Available online at: <http://jassoc.soc.surrey.ac.uk/5/2/4.html>
63. Klein S, Kuppin MA, Meives KA. Monte Carlo simulation of language change in Tikopia & Maori. In: *Proceedings of the International Conference on Computational Linguistics*. Stroudsburg, PA (1969). p. 699–729. doi: 10.3115/990403.990424
64. Hägerstrand T. *Innovationsförlöppet ur Korologisk Synpunkt*. Lund: Gleerup (1953).
65. Abbey H. An examination of the Reed-Frost theory of epidemics. *Hum Biol.* (1952) **24**:201–33.
66. Hägerstrand T. A Monte Carlo approach to diffusion. *Arch Eur Sociol.* (1965) **6**:43–67. doi: 10.1017/S0003975600001132
67. Ravenstein EG. The laws of migration. *J Stat Soc Lond* (1885) **48**:167–235. doi: 10.2307/2979181
68. Porter R. Approach to migration through its mechanism. *Geogr Ann.* (1956) **38**:317–43. doi: 10.2307/520255
69. Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. *Nature* (2004) **429**:180–4. doi: 10.1038/nature02541
70. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA.* (2009) **106**:21484–9. doi: 10.1073/pnas.0906910106
71. Cohen KJ. Simulation of the firm. *Am Econ Rev.* (1960) **50**:534–40.
72. Birchmore MJ. *A Review of Planning and Evaluation Models as a Basis for the Simulation of a Forest Firm*. M.Sc. thesis, University of British Columbia (1970).
73. Naylor TH, Finger JM. Verification of computer simulation models. *Manage Sci.* (1967) **14**:92–101. doi: 10.1287/mnsc.14.2.B92
74. Coult AD, Randolph RR. Computer methods for analyzing genealogical space. *Am Anthropol.* (1965) **67**:21–9. doi: 10.1525/aa.1965.67.1.02a00020
75. Gilbert JP, Hammel EA. Computer simulation and analysis of problems in kinship and social structure. *Am Anthropol.* (1966) **68**:71–93. doi: 10.1525/aa.1966.68.1.02a00070
76. May DA, Heer DM. Son survivorship motivation and family size in India: a computer simulation. *Popul Stud.* (1968) **22**:199–210. doi: 10.1080/00324728.1968.10405535
77. Barrett JC. A Monte Carlo simulation of human reproduction. *Genus* (1969) **25**:1–22.
78. Abelson RP, Carroll JD. Computer simulation of individual belief systems. *Am Behav Sci.* (1965) **8**:24–30.
79. Newell A, Simon HA. Computer simulation of human thinking. *Science* (1961) **134**:2011–7. doi: 10.1126/science.134.3495.2011
80. Newman EA, Evans CR. Human dream processes as analogous to computer program clearance. *Nature* (1965) **206**:534. doi: 10.1038/206534a0
81. Frijda NH. The problems of computer simulation. *Behav Sci.* (1967) **12**:59–67. doi: 10.1002/bs.3830120109
82. Green BF. Computer models of cognitive processes. *Psychometrika* (1961) **26**:85–91. doi: 10.1007/BF02289687
83. Edling CR. Mathematics in sociology. *Annu Rev Sociol.* (2002) **28**:197–220. doi: 10.1146/annurev.soc.28.110601.140942
84. Coleman JS. *Introduction to Mathematical Sociology*. New York, NY: Free Press (1964).
85. Cull P. The mathematical biophysics of Nicolas Rashevsky. *BioSystems* (2007) **88**:178–84. doi: 10.1016/j.biosystems.2006.11.003
86. Abraham TH. Nicolas Rashevsky's mathematical biophysics. *J Hist Biol.* (2004) **37**:333–85. doi: 10.1023/B:HIST.0000038267.09413.0d
87. Rashevsky N. Mathematical biology of social behavior. *Bull Math Biophys.* (1949) **11**:105–13. doi: 10.1007/BF02477497

88. Rapoport A. Spread of information through a population with a sociostructural bias: I. Assumption of transitivity. *Bull Math Biophys.* (1953) **15**:523–33. doi: 10.1007/BF02476440
89. Solomonoff R, Rapoport A. Connectivity of random nets. *Bull Math Biophys.* (1951) **13**:107–17. doi: 10.1007/BF02478357
90. Karlsson G. *Social Mechanisms*. Glencoe, IL: The Free Press (1958).
91. White H. Chance models of systems of causal groups. *Sociometry* (1962) **25**:153–72. doi: 10.2307/2785947
92. Coleman JS. The use of electronic computers in the study of social organization. *Eur J Sociol.* (1965) **6**:89–107. doi: 10.1017/S0003975600001156
93. Gullahorn JT, Gullahorn JE. A computer model of elementary social behavior. *Syst Res.* (1963) **8**:354–62. doi: 10.1002/bs.3830080410
94. Gullahorn JT, Gullahorn JE. Some computer applications in social science. *Am Sociol Rev.* (1965) **30**:353–65. doi: 10.2307/2090716
95. McGinnis R. A stochastic model of social mobility. *Am Sociol Rev.* (1968) **33**:712–22. doi: 10.2307/2092882
96. Markley OW. A simulation of the SIVA model of organizational behavior. *Am J Sociol.* (1967) **73**:339–47. doi: 10.1086/224481
97. Schelling TC. Models of segregation. *Am Econ Rev.* (1969) **59**:488–93.
98. Schelling TC. *Micromotives and Macrobehavior*. New York, NY: W. W. Norton (1978).
99. Hare AP. Computer simulations in small groups. *Behav Sci.* (1961) **6**:261–5.
100. Hartman JJ, Walsh JA. Simulation of newspaper readership: an exploration in computer analysis of social data. *Soc Sci Q.* (1969) **49**:840–52.
101. Dutton JM, Briggs WG. Simulation model construction. In: Dutton JM, Starbuck WH, editors. *Computer Simulation of Human Behavior*. New York, NY: John Wiley & Sons (1971). p. 103–26.
102. Simon HA. *The Sciences of the Artificial*. Cambridge, MA: MIT Press (1969).
103. Fleisher A. The uses of simulation. In: Beshers JM, editor. *Computer Methods in the Analysis of Large-Scale Social Systems*. Cambridge, MA: MIT Press (1965). p. 144–6.
104. Naylor TH, Wertz K, Wonnacott TH. Spectral analysis of data generated by simulation experiments with econometric models. *Econometrica* (1969) **37**:333–52. doi: 10.2307/1913541
105. Crane D. Computer simulation: new laboratory for the social sciences. In: Philipson MH, editor. *Automation*. New York, NY: Vintage Press (1962). p. 339–53.
106. Ostrom TM. Computer simulation: the third symbol system. *J Exp Soc Psychol.* (1988) **24**:381–92. doi: 10.1016/0022-1031(88)90027-3
107. Colby KM. Computer simulation of change in personal belief systems. *Syst Res.* (1967) **12**:248–53. doi: 10.1002/bs.3830120310
108. Rosvall M, Bergstrom CT. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS ONE* (2011) **6**:e18209. doi: 10.1371/journal.pone.0018209
109. Moore C, Mertens S. *The Nature of Computation*. Oxford UK: Oxford University Press (2011). doi: 10.1093/acprof:oso/9780199233212.001.0001
110. Watts DJ. Common sense and sociological explanations. *Am J Sociol.* (2014) **120**:313–50. doi: 10.1086/678271
111. Hindman M. Building better models: prediction, replication, and machine learning in the social sciences. *Ann Am Acad Pol Soc Sci.* (2015) **659**:48–62. doi: 10.1177/0002716215570279
112. Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, et al. Computational social science. *Science* (2009) **323**:721–3. doi: 10.1126/science.1167742
113. Holme P, Huss M. Understanding and exploiting information spreading and integrating information technologies. *J Comput Sci Technol.* (2011) **26**:829–36. doi: 10.1007/s11390-011-0182-3
114. Pentland A. *Social Physics: How Good Ideas Spread: Lessons From a New Science*. Brunswick: Scribe (2014).
115. Watts DJ. *Everything is Obvious: Once You Know the Answer*. London: Atlantic Books (2012).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Holme and Liljeros. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.