Abstract

Computer simulations, one of the most powerful tools of science, have many uses. This paper concentrates on the benefits to the social science researcher. Based on our, somewhat paradoxical experiences we had when working with computer simulations, we argue that the main benefit for the researchers who work with computer simulations is to develop a mental model of the abstract process they are simulating. The development of a mental model results in a deeper understating of the process and in the capacity to predict both the behavior of the system and its reaction to changes of control parameters and interventions. By internalizing computer simulations as a mental model, however, the researcher also internalizes the limitations of the simulation. Limitations of the computer simulation may translate into unconscious constrains in thinking when using the mental model. This perspective offers new recommendations for the development of computer simulations and highlights the importance of visualization. The recommendations are different from the recommendations for developing efficient and fast running simulations; for example, to visualize the dynamics of the process it may be better for the program to run slowly.

Keywords:
Computer Simulations, Mental Models, Benefits of Simulations, Recommendations for Modeling


Advantages to modeling

2.1 Models are ubiquitous in science. Formalizing a theory into a model allows the researcher to describe her ideas in a precise, unambiguous way (Goldstone & Janssen 2005; Epstein 2008). Models are conceptually precise, their assumptions are clear; they allow formal deduction and an easy way to verify their internal validity (Timpone & Taber 1996). Last but not least they provide an unambiguous way to communicate within the scientific community.

2.2 Computational models - a category of models that is based on procedural (algorithmic) description of the theory - especially agent based models and simulations, bring further advantages. They allow the unpredictable - their outcomes may be counterintuitive and impossible to infer using linear, straightforward modeling (Helbing & Balietti 2011). This is due to the phenomenon of emergence, wherein rules at the system's micro level result in unpredictable patterns and behavior at the macro level. Therefore, computational models - implementing the rules and showing macro behavior - link different levels of description of collective behavior; they show how order may emerge spontaneously (Goldstone & Gureckis 2009).

2.3 Computational models are powerful when we need to implement and infer from several theoretical assumptions simultaneously. Such a model, especially when the relations between elements are non-linear, often cannot be solved by analytical methods. Thus, computational models allow for testing hypotheses that are only theoretically possible to verify through mathematical modeling (Timpone & Taber 1996).

2.4 Computational models (especially agent based simulations) allow for transparency (one agent - one individual) - they do not require any averaging techniques to relate the model to reality (Edmonds & Hales 2005). They are an abstraction of the target (modeled) system and the insight that they provide can be applied - through the interpretation of their results - back to the target system (Edmonds 2001).

2.5 At the same time computational models are by far more interactive than mathematical models; computer simulations provide an environment to develop, test and articulate theoretical propositions (Mollona 2008). They provide an experimental environment for researchers to play with symbolic representations of phenomena by changing the model's structure or varying its parameters.

2.6 Epstein (2008) summarizes these advantages to modeling by proposing 16 different uses for models besides prediction and they include among others: explanation, guiding data collection, suggesting dynamic analogies, discovering new questions, offering crisis options in near real time, challenging robustness of the prevailing theories through perturbations, revealing the apparently simple (complex) to be complex (simple), exposing the prevailing wisdom to be incompatible with data, and developing a scientific, open mindset, which he stresses to be the most important. However, it is the benefit of 'illuminating core dynamics' that comes closest to the thesis that the crucial effect of simulations is developing mental models.

New insights

3.1 What does simulation add to the advantages of modelling? Mathematical modelling does not require simulation while computational modelling is inevitably linked to computer simulations. Are these two equivalent then? For many purposes, the answer would be yes. Simulations allow for testing different hypotheses, verifying the interdependences between parameters, providing results and graphs of the models' dynamics, and similar. All these uses are naturally associated with the process of simulating a model. We argue here, however, that there is one more, crucial role of computer simulations that researchers rarely realize - simulations allow the researcher to build a mental model of the target system.

3.2 To understand the primary use of computer simulation for modelers, the distinction between the public and private contexts for computer simulations is crucial. In publicly communicated results of simulations researchers provide: graphs showing how the results change with changes of control parameters, statistics of aggregated data, significance tests etc. In the private context, for the researcher the process of interacting with the model and watching visualization is most revealing as it gives a subjective feeling of understanding of the process. An intuition develops that links the parameters of the model to its dynamics and the simulations may "be run mentally" by imagining the process, where no computer is needed. This kind of running the model mentally may even enable the researcher to predict - without actually running the simulation - the results of a particular parameter set. (Of course, it may happen that the intuition and the resulting mental model are wrong, as sometimes happens with formal models as well). The concept of mental models, developed in cognitive psychology, explains why repeated playing with the simulation program results in understanding of the process and the ability to mentally run the simulation model.

3.3 What are mental models, then? Johnson-Laird concisely summarized their key value: "To understand a phenomenon is to have a working model of it" (Johnson-Laird 1983, p. 4). Mental models are the mind's replicas of phenomena, upon which humans manipulate to learn the mechanism and workings of the surrounding environment. Mental models by definition have explanatory value (Craik 1943). However, they are necessarily incomplete - they are always simpler that what they represent - and sometimes completely wrong. Still, they enable the individual to make predictions, to understand phenomena, to decide about action, and to experience events by proxy. Moreover, mental models are finite and the mechanisms for constructing them are computable (Johnson-Laird 1983).

3.4 Individuals use mental models to explain and anticipate events and actions (Craik 1943; Johnson-Laird 1983; Rouse & Morris 1986; cf. Wilson & Rutherford 1989). With respect to social cognition, mental models are a platform for mental simulation of
people's actions and decisions. They provide "a mechanism whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states." (Rouse & Morris 1986, p. 360).

3.5 By manipulating mental models one can estimate what is possible and what is not possible in a given system, and how to achieve possible states (Johnson-Laird 2001). Because mental models represent cognitive tools used for understanding and prediction their form can shape human decisions and actions (Park & Gittleman 1995; Gentner 2002). For example, naïve market investors may build models of market crashes and rises and make their financial decisions accordingly.

3.6 Mental models are distinctly different from imagery. Mental images (or, more precisely, representations) are static while mental "videos" (called mental simulations by Craik) are dynamic but do not allow manipulation. Mental models always contain an element of such a mental "movie" (Johnson-Laird 1983) but they also contain relation-structure similar to the modeled process (Craik 1943). This provides them with explanatory value.

3.7 Mental models develop most accurately through interaction with the target system (Norman 1983). Moreover, they evolve and are elaborated through interaction as well. If interaction is impossible, they are often built upon an analogy with a process for which a model already exists - e.g. electricity might be modeled in analogy to river flow (Genuiler & Gentner 1983). The analogy may shape the model and may have impact on the operations on the model, which are afterwards available to the individual.

3.8 How do simulations relate to mental models? The researcher clearly has at least a natural language theory before she or he starts modeling. We can think of this theory as a mental pre-model as it is based on some experience or knowledge (eg. previous modeling attempts in similar areas, results of experimental studies or data analysis) but lacks the understanding of consequences of putting this knowledge to action. By implementing the theory in a computer code the researcher clarifies its assumptions and makes them unambiguous.

3.9 This step is the first crucial part in computational modeling, yet it is still not a mental model. For the construction of the mental model the next step is crucial - running the simulation in an interactive way. Through the interface of the simulation, the multiple runs with different parameters or "live" manipulation of variables the researcher is provided with an environment to interact with a replica of the system that they are modeling. This clearly is similar to the process of forming mental models of physical systems.

3.10 The mental model consists of both the implemented rules of the system as well as their results on system level - often emergent and unpredictable - of the simulation runs with different control parameters. Once its development is complete, running the simulation on a computer is needed as further reassurance and of course to produce appropriate aggregate results. In sum, running simulations in an interactive way enables the formation of a mental model, and thus enables understanding.

3.11 The next step - analysis of results and interpretation - is the second crucial part in computational modeling, enabling the relation between the modeled system and the target, modeled system (Edmonds 2001). However, for the mental model development it is the proverbial icing on the cake. In fact, once the mental model is developed through interacting with the simulation, the researcher already has the interpretation in her or his mind. There is no need to relay the results to the target system as the results are part of the mental model and the mental model is OF the target system. Like all mental models, the ones stemming from the simulations are necessarily simpler than the real systems but this fact does not hinder their explanatory value. Of course, presenting the results of simulation runs is still required to communicate the model properties to others and to allow them to check the claims of the researcher.

3.12 Simulations can also be used as way to elaborate an existing mental model. In such a case the researcher implements her model in the computational form and through interaction with the simulation may modify it and increase its precision.

3.13 The psychological notion of mental models nicely captures the process by which computational modeling with simulation enables researchers to deeply understand the target systems of their models. Linking the widely accepted role of computational models with a psychological concept leads to some clear consequences.

   a. Development of the mental model is achieved by internalization of the relation between simulation parameters and results. It comes, however, in a package deal. Simulation is sometimes so persuasive that it might occlude the unbiased perception of reality - the researcher internalizes all limitations and constraints of a particular implementation, which frame her further thinking about the real process in sometimes detrimental way.

   Usually without being aware of it the researcher internalizes the technical limitations of simulations such as the simulation size, or the way the parallel social processes are handled on a serial computer. For example, running simulations where 2,500 individual are arranged on a matrix of 50 by 50 allows the researcher to see emergent group level processes such as clustering. In most cases, however, the size of the simulation would be too small to allow the emergent clusters to be treated as higher (mezzo) level elements that can interact with each other. The limitations of the size of simulations can therefore prompt one to think of two-level: micro - macro level systems. On the other hand using fast computers and strong visualization tools makes it possible to run simulations of millions of agents. The mezzo level structure may be visible in the observed simulations and the mere possibility of observing them may provoke the researcher to explicitly introduce into the simulation rules for interactions between the emergent mezzo-level structures.

   b. From this perspective, the interface of the simulation is crucial. Some ways of visualizing might bring more insights and
understanding than others. The ability to interact with the program can be designed in various ways for the same system and the implementation, again, matters.

c. The language of implementation (code) shapes the mental model. In the traditional approach to computer science it is believed that the language in which the code is written is irrelevant; what counts is the algorithm, not its implementation. A mental model may contain the code of the program as well as the description of the model and visualization as the emergent dynamics. As a part of the model, the language of the program can thus make a difference. If the researcher models in matrices they will also “think” in matrices. And this way of thinking may be alien to another modeler that represents data in dynamic structures of C++ or Java objects or as a LOGO-turtle. If the code is written in a clear, concise way the mental model may be simpler and easier to manipulate than if the code is written in a language that is, for example, too low level (e.g. assembler). Similarly, declarative language will result in different mental models (on different level) than structural or object coding. It is also the case that the mental model may be difficult to form if the language is too high level and it occludes what is really happening in the program.

d. Finally, sharing the simulation capture or simulation code alone is not equal to sharing the mental model that results from and links both. Presenting a video of a simulation run or even a couple runs - even the most revealing of all - will provide the audience with a mental “movie” but not a mental model. Interaction with the simulation is critical for the development of a mental model on the basis of computer simulations.

When are mental models especially useful?

4.1 The capacity of the human mind to form accurate mental models is, however, limited in several ways. First of all, because mental models contain the effects of manipulations of the object, the capacity to interact with the object or to influence the process is critically important for developing an accurate mental model of it. Mental models of concrete objects and observable processes are built on the basis of observation. Building a mental model of an abstract concept of a process, on the other hand, is difficult, because abstract entities cannot be directly observed and are impossible to manipulate directly.

4.2 Furthermore, even when observation is possible, individuals can only discover the rules that govern the dynamics of the mentally modeled process if they are relatively simple. This is usually the case with linear dependencies between variables, if the number of variables is relatively small. Human intuition fails, however, with non-linear dependencies between variables. In cellular automata, for example (see Wolfram 2002) one set of rules leads to a very complex dynamics, another set, which differs just by one detail, to a fast trivial unification. No one, however, can tell, just by looking at the rules, which of the two set of rules will result in which type of dynamics. It is difficult for the human brain to model the effects of interactions of elements governed by even a small set of non-linear rules. This stems in a large part from the fact that we are almost incapable of extracting non-linear causality rules from observed behavior, even if we are able to manipulate the object. That is, if the results of manipulation do not follow linear causality we may deem them unrelated to the manipulation, while in fact they might be exhibiting a non-linear dependency. This is especially true if the non-linear dependencies follow a complex function or are composed of several, over-imposed rules.

4.3 Another difficulty in forming mental models comes from the fact that many observed processes take place in complex systems composed of many interacting elements. If the elements are few (for example an interacting dyad) we might be able, in some cases, to extract system level behavior from the rules governing the elements and thus form an accurate mental model of the whole system. However, even a small increase in the number of elements increases the complexity in such a way that, again, we are unable to relate in a causal way system level behavior to the behavior of its elements. Thus, from a human perspective, new, unexpected properties emerge on the level of the system that are not present on the level of elements. The modern notion of emergence refers to exactly this phenomenon.

4.4 The difficulties in constructing mental models grow substantially if these three issues are present at the same time. Thus, mental models of abstract, non-linear processes happening in complex systems are almost impossible to construct solely using individual cognitive capabilities.

4.5 Computer simulations provide exactly the means by which individuals can form mental models of such processes. In contrast to humans, computers have no problems with non-linearity of relations. Arguably, for modern computers size does not matter any longer either, or at least matters less and less with the current advances in computing capacities. Moreover, computers are not bothered by the abstract nature of the process. For building mental models computer simulations have a strong advantage over watching video visualizing the process because they offer the chance to interact actively with the model, even if they model an abstract process. The possibility of active interaction with the model, changing the model's parameters and observing the effects is critically important to be able to mentally represent the effects of various interventions.

4.6 Mental models acquired through computer simulations consist of the abstract rules (internalized fragments of the code of the program or internalized algorithms) and the visualization of the emergent properties at the level of the system. The capacity to predict the effects of various interventions depends on the ability of the simulation to produce different visualizations of the emergent properties as the rules change. In essence, computer simulations, along with their visualization, may be a means of developing mental models in place of observation and interaction with physical systems when such observation is impossible - as it often is with abstract, non-linear process in complex systems.
The goal of developing mental models offers a new perspective on the requirements for computer simulations. This perspective supports some earlier intuitions (e.g. Hegselman & Flache 1998) such as the importance of visualization. To our knowledge some of these requirements are, however, novel or even run contrary to the prevailing beliefs. Below we will discuss some of them, especially elaborating on the role of visualization.

Recommendations

5.1 The speed of simulation. It is usually believed that the faster the program runs the better. It is certainly the case if the simulation program is used for the many repetitions needed for statistics. However, from the point of view of the acquisition of a mental model there is some optimal speed for the program at which the dynamics of the model is best visible. If the program is too fast some aspects of the dynamics are lost; the user sees only the starting and the final configuration while the lost, intermediate dynamics might be exactly what is needed to understand how the parameters influence the results. If the program runs too slowly, on the other hand, the dynamics becomes too boring to observe and may be difficult to internalize, as the changes in time are too slow to extract the effects of the manipulation of parameters.

5.2 Dynamical minimalism. The strategy of dynamical minimalism (Nowak 2004) calls for constructing the simplest model that implements what we believe are the most basic rules for interaction between a system's elements and that can reproduce known dynamics of the system being modeled. The usefulness of computer simulations is proportional to the difference between simple rules of interaction between elements and the emergent complex properties at the system level. In other words, the more emergent properties, the more useful is the simulation model. The strategy of dynamical minimalism thus maximizes the usefulness of simulation models. The requirement of simplicity also has important consequences for the richness of the simulation interface. It is often tempting to put as much information on the visualization screen as possible following the belief that additional information does not hurt, as long as it fits onto the screen. If the screen is too rich with information it may be difficult to isolate the mental model of the phenomenon. For example, often programmers using Net Logo include in their articles screen shots that contain a lot of additional elements, not directly related to the visualization of the process which is being modeled.

5.3 The importance of visualization. From this perspective the way of visualizing the program is of critical importance - it affects the shape of our mental model. The crucial role of visualization is especially apparent in the simulation of social processes - it is indispensable at the stage of building a model but it is also very useful when trying to convey the modeler's understanding of the process to social scientists in publications and presentations of the results. Some researchers, especially those with a formal background, concentrate on publishing only the model's results (e.g. Stauffer, Sousa & Schulze 2004). Most researchers, however, chose a clear visualization paradigm for their simulation. Sometimes these are matrices of objects or simple agents, derived from the tradition of cellular automata (Hegselmann & Flache 1998; Flache & Hegselmann 2001), sometimes models resembling gas particles dynamics (Ballinas-Hernández, Muñoz-Meléndez & Rangel-Huerta 2011), other times typical agent based visualizations are often implemented in the Logo language which is the only language to incorporate visualization into its core definition (e.g. Kim & Hanneman 2011). When efficiency is at stake (for example, when the number of agents is very large) the Java language (Axtell et al 2002) or C++ (Millington, Romero-Calcerrada, Wainwright & Perry 2008) is chosen.

5.4 The evolution of the model of Dynamic Social Impact model (Nowak, Szamrej & Latané 1990; Lewenstein, Nowak & Latané, 1993) illustrates how different ways of visualizing by developing different mental models can lead to different shapes of the theory. The first generations of the models (Nowak et al, 1990) used the ASCII characters X and I to represent individuals located in a 240 by 240 grid. SITSIM (Nowak & Latané 1994) used the same size of the grid and smily faces either in normal print or inverted as characters. These characters in simulation visualization gave the impression of areas of different shades of gray. The mental model of authors at this stage represented a disordered mixture of opinions changing into large coherent clusters of likeminded individuals, or in some conditions unifying. If the model included noise (randomness) some individuals, mostly located on the borders of clusters, continued to change their opinion back and forth. If the noise was higher, the clusters become unstable, moved around and began to lose their coherence. The mental model and thus the research questions concerned the relation between the model parameters and whether the final outcome was clustering or unification.

5.5 In the next generation of the model (Szamrej & Nowak unpublished), individuals were represented as boxes of different color, which represented opinions, and different height, which represented persuasive strength. In initial versions of the model strengths was a random continuous variable. The resultant visualizations were so messy that nothing interesting could be observed. In the next version (Nowak, Lewenstein & Freijak 1996) the authors decided to use only two values of the strength parameter, dividing individuals to be either leaders (20 % of individuals with the strength equal to 10) or followers (80 % of individuals with the strengths equal to 1). Suddenly, a different configuration of leaders became visible and it was clear that the configuration of the leaders was critical for the dynamics of clusters of minority opinions. The research has subsequently concentrated on the basic configurations of the leaders, which was identified as: a single leader surrounded by the ring of followers, a ‘stronghold’ (a clustered group of leaders supporting each other) and ‘the wall’ (a group of followers guarded by the leaders located on the border of the cluster). The mental model of authors changed to boxes, and it started to represent possible configurations of the leaders and the resultant dynamics. The theory has shifted towards the effects of leadership.

5.6 Changing the scale of the model led to a very different visualization. This, in turn, has resulted in the change of researchers' mental model and led to a somewhat different shape of the theory of social influence. One version of visualizing social impact involved using a much larger simulation area of 600 by 600 (Nowak, Lewenstein, Szamrej 1993) and representing individuals by
pixels. At this resolution, what on a much smaller scale appeared as shapely clusters, started to look like small, somewhat interconnected randomly distributed patches. At this scale it was apparent that the model resulted in clusters of characteristic size, which looked very small and not very interesting at this resolution (see Fig. 1). Playing around with the model it was discovered that much less demanding assumptions than in the original model led to more visually interesting results. Assuming that the opinion in each simulation step changes as the result of interacting with a random sample of neighbors, rather than with all the neighbors as in the original model, had the effect of increasing the size of the clusters and making their shapes more interesting. Another new feature introduced into the simulation was that individuals discuss three topics with multiple, rather than binary possible opinions on each topic. The opinion on each topic was displayed as an intensity of one of the RGB components of color (Culicover, Nowak & Borkowski 2003). The combination of opinions on all the three topics was displayed as the composite color. The resulting visualization was very different than the previous ones. From a random dispersion of intermixed dots of different colors, growing clusters of various colors started to appear, where each color represented a particular combination of opinions. The number of colors appearing on the screen decreased with the progress of the simulation, as the number of combinations of opinions on all the three topics decreased. This process resembles formation of an ideology or culture, where the initially unrelated features become correlated as the result of social interaction. If we assume that each ‘opinion’ or characteristic of an individual that changes as the result of influence from others, represents a feature of a language, our result resembles a decrease in the number of languages that is observed in the world (Culicover, Nowak & Borkowski 2003). The new visualization was so different from the other ones that the researchers have formed a new mental model instead of extending the old model. These two models are not, however, conflicting. In thinking about social simulations it is possible to switch between the two models, choosing one or the other depending on the assumed theoretical model and the scale of simulations. Subjectively, it resembles two views of the same object. It is also possible formulate mental models that are the combination of the two theoretical models and imagine their dynamics.

5.7 The possibility and ease of playing with model's parameters. Mental models contain the effects of manipulations of the properties of the modeled process. One learns to manipulate the models by manipulating parameters of computer simulations. Therefore, the interface of the simulation - so far often ignored by many modelers - plays a vital part in simplifying the construction of mental models. The possibility of changing simulation parameters is critically important for the development of mental models. If one just observes the simulation as a movie it may result in the development of just a dynamical image, which is similar to a movie, rather than a model which one can manipulate.

5.8 Writing the code oneself A mental model contains both the representation of the rules as the code and the emergent properties
as the visualization. A perfect way to form a mental model is thus to write the code - it is either a creation of the basis of the mental model or an expression of the model that already exists. Writing the code oneself guarantees that the rules of the model exist in the mind of the scientist. If the code is written by someone else we may internalize the dynamics of the visualization, but its usefulness may be limited because the rules producing the emergent properties that are visualized do not exist in the mind. When working with the programmer, it is essential to remain in close contact with the person developing the code, and be aware of how the principles of the simulation are developed in code. To some extent it may be possible to build a mental model by running a program developed by someone else. Here the critical component is the ability to understand the exact rules implemented in the model.

Summary

6.1 In this paper, we have considered the benefits of computer simulations from the perspective of the author of the code, or a scientist closely collaborating with the programmer who develops the code. We have discussed computer simulations in the context of scientific insight. We have concentrated on issues that have not received much attention before. We have suggested several recommendations for maximizing the effectiveness of computer simulations as a tool for building mental models of abstract processes. Some of these recommendations will not apply or may be even reversed for other uses of computer simulations. For example, for visualizing the process it may be advisable to present computer simulations at a slow speed. For generating massive results that will serve as an input for statistical analysis it is better if the simulation programs run fast. Very different recommendations can be specified if the computer simulations are considered in the context of a proof. In that context hard, rigorous results are expected to result from computer simulations, rather than improvement of a scientist's intuition.

6.2 Computer simulations arguably represent the most powerful tool of modern science. They are the tools of choice for investigating emergent phenomena (e.g. Nowak 2004) and for investigating the properties of systems that cannot be investigated by analytical means. They provide a new formalism, the third symbol system, besides the natural language and mathematics for the description of objects and processes (Ostrom 1988). The benefits of computer simulations have been discussed elsewhere (see for example Epstein 2008; Hegselman & Flache 1998; Liebrand, Nowak & Hegselmann 1998). Computer simulations are used for many purposes and different requirements may influence their usefulness for various goals. In this paper we have argued that one of the important goals is for the researcher to build a mental model of an abstract process.

Acknowledgements

This work was partially supported by the Future and Emerging Technologies programme FP7-COSI-ICT of the European Commission through project QLectives (grant no.: 231200)

References


http://jasss.soc.surrey.ac.uk/14/3/2.html


http://jasss.soc.surrey.ac.uk/16/3/12.html


