

**Simulating Society:
The Tension Between Transparency and Veridicality¹**

Kathleen M. Carley

October 2002

Citation: Kathleen M. Carley, forthcoming 2002, “Simulating Society: The Tension Between Transparency and Veridicality” Proceedings of Agents 2002, Chicago IL.

¹ This work was supported in part by National Science Foundation IGERT in CASOS, the Office of Naval Research (ONR), United States Navy Grant No. N00014-97-1-0037, Army Research Labs, NSF ITR/IM IIS-0081219, NSF KDI IIS-9980109, and Darpa. Additional support was provided by the NSF IGERT for research and training in CASOS and by the center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University (<http://www.casos.ece.cmu.edu>). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Department of Defense, the National Science Foundation or the U.S. government.

Computational analysis is dramatically re-shaping the way we think and reason about society and social processes. Everything from the impact of information technology to the fundamentals of cooperation and altruism are being addressed using computational models. Computational models, often in the form of virtual worlds, are used in social, technological and engineering policy domains to address via what-if analysis, how different technologies, decisions and organizational and government policies influence the performance, effectiveness, flexibility, adaptiveness and survivability of complex social and organizational systems. Computational models are being increasingly used in the classroom to demonstrate social processes and the impact of change to undergraduate and graduate students. New programs are rapidly springing up in which computational modeling and analysis plays a role.

Essentially, the nascent field of computational social and organizational science has been born. The focus of this field is the study of societies and organizations as computational entities. Organizations and societies are viewed as inherently computational as they are complex adaptive information processing systems incorporating search engines. As noted by Gasser and Carley (1999) computational organizations are seen as taking two complementary forms – natural and artificial. The natural or human organization or society is universally informatted; that is, filled with, continually acquires, manipulates, produces, and disseminates information. It is a multi-agent system in which information acquisition, dissemination, processing and search is carried out by the joint, and interlocked activities of people and automated information technologies which are embedded in a specific organizational design. The artificial organization or society is composed of multiple distributed heterogeneous socially intelligent adaptive agents. Each of these agents has organizational properties such as the need to act collectively, a task assignment, a set of knowledge, skills or abilities, and constraints on which others agents they can interact with when and about what. These agents can mutually influence, constrain and support each other as they try to manage and manipulate the knowledge, communication and interaction networks in which they are embedded.

Computational analysis is used to develop a better understanding of the fundamental principles of “sociality”; i.e., of organizing, coordinating, adapting, and managing

multiple information processing agents (whether they are humans, corporations, WebBots, or robots) and the fundamental dynamic nature of groups, organizations, institutions and societies. Indeed, computational analysis plays a ubiquitous role in theory building, data collection, data analysis, education and policy analysis. For example, a combination of model development, simulation, and virtual experiments are used to develop a better understanding of the fundamental principles of organizing multiple information processing agents and the nature of organizations as computational entities. Overall, the aim of research in this area is to build new concepts, theories, and knowledge about organizing and organization, coordination and linkage, communication and technology; to develop tools and procedures for the validation and analysis of computational organizational models; and to develop computational organization tools that can be used as educational and management aids. Importantly, computational analysis is not simply in service to organizational and social theorizing; rather, computational theorizing about human phenomena is actually pushing the research envelope in terms of computational tools and techniques. Research in this area has resulted in a large number of models, empirically grounded theory of organizational design and adaptation, better management tools, and a more complete understanding of the way in which social, organizational, and knowledge networks inter-link to affect effective, robust, and adaptive organizational designs. A number of edited volumes (e.g., Carley and Prietula, 1994; Prietula, Carley and Gasser, 1998) and the journal *Computational and Mathematical Organization* carry research in this area.

Computational organization science is a new scientific field whose roots are interdisciplinary. Despite differences in training, the researchers in this area share a common methodological orientation to formal modeling, which due to the complex and non-linear nature of organizations often results in the use of computational models. The formal models in this field thus include computational (e.g., simulation, emulation, expert systems, computer-assisted numerical analysis) and mathematical (e.g., formal logic, matrix algebra, network analysis, discrete and continuous equations) with many researchers using whichever is appropriate to the research question being addressed. However, the community is not in agreement about the relationship between models, theory, and reality; moreover, the community is not in agreement about the fundamental

bases for judging the value, importance, or quality of a computational model. As the growth in use of simulation grows in the social sciences so does the debate over evaluation of models.

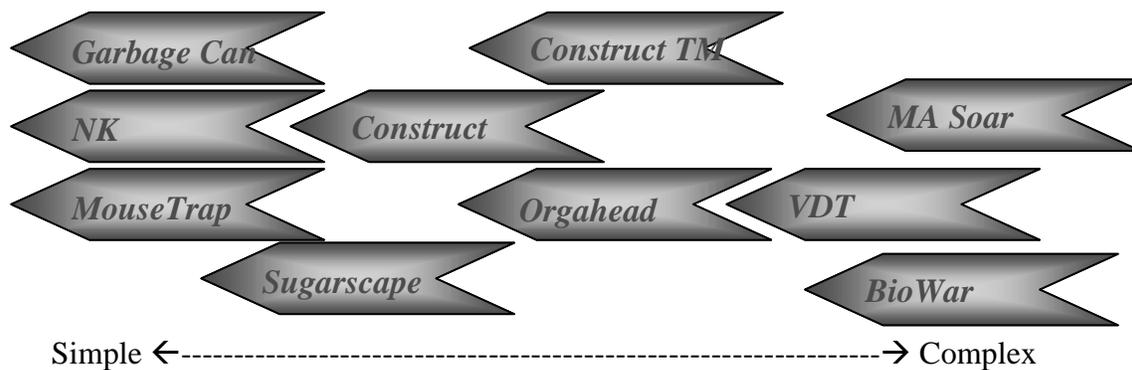
Formal models are used to develop and test theory. Some members of this community take the strong computational stance that the theory (i.e., the simulation model) should do the task it seeks to explain. In this case, the models can actually take the place of agents (human, group, or organization) in an experimental setting. Due to the use of computational modeling, computational organization science is an important component of the curriculum in distributed artificial intelligence (Carley and Gasser, 1999). In this case high veridicality is called for. Some members of the community take the stance that the model is the theory. Some members of the community take the stance that many models can reflect a theory. Formal models are expected to bear some relation to reality.

The relation of computational models to reality is complex. Underlying all the diverse ways in which data can be linked to models is a fundamental tension – accuracy versus simplicity. In this paper, this tension and how it plays out in the computational social and organizational sciences is discussed. Findings from behavioral and cognitive psychology are used to explain the basic way people respond to computational models. On the one hand, there is a belief in simplicity. The basic argument is that, if they are to be explanatory, models should be a reduction of reality. Apply Ockham's razor and find the simplest explanation. On the other hand, there is a belief in accuracy. The basic argument is that, if they are to be accurate, models should provide a match to the real world at a sufficient enough detail for the problem at hand. Apply validation tests and find the satisfactory explanation that enables you to make decisions, set policies, etc., with minimal risk. Immediately, it should be obvious to the reader that the problem is a socio-psychological one – that is simple and satisfactory are in the eye of the beholder. This tension is often played out in terms of arguments over transparency and veridicality.

Transparency means that it is “obvious” to the viewer how things work. The basic analogy is the idea of a glass clock where the transparent face lets you see in to the mechanism. In other words, transparency implies “I understand it”. Veridicality means

that the model works like the real world; i.e., it portrays truth. This can be thought of as “I observe a match between the model and the real world.”

Within the computational social and organizational sciences, models run the gamut from the very simple models to the complex detailed models. For simple models the authors often argue for the value of transparency. For the more complex, the authors often argue for the value of veridicality.



The research in computational organization science spans all aspect of social and organizational science. In each domain examples of simple and complex models exist. To anchor the discussion, consider two such models. The first is the Garbage Can Model of Organizational Choice by Cohen, March and Olsen (GCM). This is a classic and very simple model. It is readily re-programmed in a couple of weeks by undergraduates in computational modeling courses. The second is BioWar, a very detailed complex model that to date has taken 5 people years to develop.

The purpose of the GCM was to illustrate that choice and energy lead to an organizational situation where not all decisions get made. This was in fact an argument against optimization and rational behavior and for satisficing and boundedly-rational behavior. The purpose of BioWar is to enable policy makers to evaluate privacy restrictions, containment policies, facilitate detection, etc. for weaponized biological attacks in cities. BioWar is a city-scale multi-agent network model of weaponized biological attacks linked to census data, school district information etc. and capable of generating insurance claim reports, absenteeism, etc.

The basic tension transparency and veridicality is not unique to the social and organizational sciences. However, the state of the computational field here, the level of mathematical training, and the relative paucity of computation leads to a different balancing act than in engineering, physics, and chemistry. The basic difference is this: *In engineered systems people don't assume they know how things work, but trust the "math" of "the physical world."* *In social systems people assume they know how things work, and don't trust the "math" of "the social world."* The "physical and engineering" sciences extensively utilize simulation. In part, the greater acceptance of computation is due to their being older science and so as a result having a greater understanding of the phenomena being studied. They are also a "wealthier" science with greater budgets from foundations, funding agencies, industry. Which means, more work has been done. And, very importantly, they are relatively simpler sciences mathematically. That is the phenomena being studied are less complex (fewer interacting parts), the fundamental entities don't "learn", and as a result less data is needed for validation of a model than in the social and organizational sciences. Whereas, the social and organizational sciences are newer, less financially well off, and the phenomenon of study more complex than are the physical sciences. When you couple this with the fact that there is relatively less mathematical and computer sciences training in the social and organizational sciences, it should be obvious that there is a problem. Moreover, I would argue, the fundamental nature of human cognition exacerbates this problem leading to extended debates and possibly poor choices regarding transparency and veridicality.

The upshot is that in the social and organizational sciences, there is essentially both a "physics envy" and a distrust of mathematics. Consequently, social and organizational scientists are prone to equate transparency with simplicity. If the model is simple I assume it is transparent. Moreover, many people go on and suggest that if a model is transparent it has achieved sufficient accuracy with the real world and it is a meaningful model. In other words, basic human nature really means that transparency is not "I understand it"; but, "I think I understand it". Thus, the field is filled with people who look at a simple model and simply assume that they understand it. Transparency is not transparency but perceived transparency.

In terms of veridicality, additional forces come to play. The lack of personnel and finances, as well as the relative youth of the field, means that there is relatively little data on the phenomenon of study. Consequently, trust often replaces proof. Highly veridical models, which are of necessity reasonably complex, are generally perceived with distrust. Essentially, the general distrust of math engenders a lack of trust in computational models. This is then exacerbated by the paucity of data, which leads to both divergent expectations due to inability to completely map the landscape of possibilities and minimal levels of validation. This lack of trust often leads to arguments of the form – “this model (the highly veridical one) does not provide insight” or “if you had a good theory you wouldn’t need this level of complexity.” Both arguments are specious. However, they are made, in part, due to a lack of education and in part due to a lack of agreement with the model. The latter often is due to people thinking – “well my data doesn’t agree with your model.” Thus, what veridicality really means is “*I believe there is a match between the model and the real world.*”

Rapid advances have been possible due to a unified approach to information processing, explicit attention to the findings of contingency theory, the use of canonical tasks, and the use of social network representation schemes and measures. This unified approach is beginning to payoff in that researchers models are now building on each other and the models can be docked one to the other. Nevertheless, the problem just described and with the rapid increase in graduate students interested in computational modeling is likely to affect many individual’s careers. To understand the ramifications of this problem I take a socio-cognitive perspective and explore how the basic tenets of human behavior affect the modeling community.

The Psychology of Perception

Within cognitive science and behavior psychology a number of findings have emerged in the last few decades about the nature of the human mind. Let us consider a few of these:

- People automatically create interpretations of visual images. Whether we are talking about blobs or networks, when faced with a picture – people are uniformly able to create a story, an interpretation of what they see. Simpler pictures lead to simpler stories. However, the relationship between commonality of story and the

picture is unclear. In other words, pictures create meaning but not necessarily shared meaning.

- Chunking facilitates learning. Basically if you can take a complex story and break it down into self contained segments it will be easier to learn. Average attention span, age, gender, and countless other factors contribute to the size of the chunk that can be learned at once. Many educators suggest 15 minutes as the temporal size for a chunk. If it takes longer than 15 minutes – divide it up.
- People learn many things by experience. Moreover, with feedback, the greater the experience the better the performance. This is the typical Bush-Mosteller learning. The result is essentially an s-shaped learning curve.
- When we have nothing else to go on, we assume others are like us. Essentially, people generalize a lot and use analogical reasoning a lot. These coupled with experiential learning result in people knowing themselves best. If people decide that they are “alike” then they will assume that they will behave in the same way and know the same things and share the same values.
- People are overconfident in decisions even when have little data. Basically, people base the likelihood of things on their own experience without taking in to account actual data. If there are more red cars in my neighborhood than blue, then I assume that everywhere there are more red cars than blue.
- People’s beliefs are a function of their social information processing (social influence). The basic idea is simple, you are more likely to believe something I tell you if we are friends. Similarly, you are more likely to share the same beliefs as your friends. In affecting a change of opinion, therefore, social influence can be as or more important than the influence of facts.

Now let us consider the implications of these findings for the modeling community. First consider the implications of the fact that people create interpretations of visual images automatically. This means that visual images are being interpreted. Further, it means that people have their own interpretation. In fact, since you cannot understand something unless it relates to something you already know, interpretations will vary widely when you have a group with a wide variety of experience. So how do we know which interpretations are accurate? When does it even occur to us that our interpretation is not shared?

Accuracy of interpretation, at least in science in the United States is typically judged by consensus. However, as human we appeal to the will of the majority only when in doubt. The higher the complexity of the visual image the less likely it is to be completely processed by the viewer. People tend to be aware of their processing. Thus, the higher the complexity of the model and/or the higher the complexity of the visual aid, the more likely it is that people will be aware of not having processed everything. As such, it is more likely that people will think they don't understand the complex model. This is exacerbated by chunking, which is discussed next. This means that even though the viewer is interpreting both the simple and the veridical model they are more likely to be aware that they are making an interpretation and to be less confident in their interpretation for highly veridical models.

Thus, simplicity facilitates visualization. Visualization increases perceived transparency. Thus simple models are viewed as transparent. People think they understand them and that there is no room for interpretation. There is no call for consensus as it is assumed. This in turn engenders extensive claims of applicability as each viewer interprets and so applies the model in their own substantive context. In contrast, veridicality leads to either more complex visual images or to simple images containing proportionally less information (than the same complexity of image for a simple model). When a complex image is used people are more aware they don't understand things. As a result, they are less likely to trust the model. Whereas when a high-level but simple image is used people are more likely to think they understand the model and that there is no room for interpretation. However, they may be wrong.

Chunking exacerbates this process. The idea, again, is that people learn in shot contained chunks. Simplicity facilitates short presentations. There just isn't that much to say. Consequently, simpler models, which are perceived as transparent, should be easier to learn. Simple models can be "learned" in fewer lessons than highly veridical models. From this, a common inference is likely to be that transparency promotes learning. This would, however, be a somewhat fallacious inference as a) it is only perceived transparency and b) it is the simplicity that is the core cause. For highly veridical models, chunking implies that the model must be modularized in order to present it. Since veridical models just have more to them, this means that they require more and/or longer presentations. Now, if we add the fact that most people are busy, this means that the chance of being present to learn all of model is higher for simple than complex models. In addition, the chance of learning the model if present is higher for simple than complex models because there would be fewer or smaller chunks. Additionally, given the limited number of contact hours we have with students, an implication is that educators will be unlikely to teach veridical models in total; but may try to teach multiple simple models. As a result, there should over time be a broader community of scholars who think they understand the simple models, have their own interpretation of it, and don't question it. And, there should be a smaller community of scholars who think they understand or have even been exposed to the more veridical models. This can lead to a wide-spread view of the lack of utility of such models at the same time as a smaller group of insiders develop who fully believe in, and have found validation for, these same models. It also suggests that the more veridical models would be taught at very few institutions; most likely, only at those where the developers teach.

Next, consider the role of social learning. Here there are three findings that need to be considered at the same time: 1) experiential learning – "I live in the real world therefore I have "learned" how it works"; 2) others are like me – "My interpretation of how things work is shared", and 3) overconfidence – "I am right about how things work even though I am reasoning from an experience of 1." These factors come together to suggest that the accuracy of a model is judged not by "objective" shared data, but by subjective experience.

For simple models, social learning means that simple models are perceived as transparent. Basically, people look at the simple model and go through an exercise like the following.

“I think I understand the model. I think my understanding is shared by everyone. I don’t expect the model to match the real world. I think my perception of what the model has to say about the real world is shared by others. I am right and so I do not need to check my facts. Moreover, because there is a common understanding of what the model has to say and its limitations; we can use this model to set policy, to make decisions, and to educate. “

Since simple models are easily learned, taught, and communicated there will be many people acting this way. Which means that such models will be used to set policy without the users ever confirming that it really matches the real world or that their interpretation is shared by everyone. This means that policy setting will be based on story telling with the simulations used as a device for creating scenarios from which to reason. This is not meant to imply that this is a bad way to set policy or make managerial decisions. However, it is meant to suggest that human socio-cognitive behavior will lead simple models to be used even when they don’t match the real world (are not veridical). Moreover, since as you increase the simplicity of the model you often increase the number of interpretations; this use of simple models under the guise of unspoken agreement means that the group of decision makers may be acting on a presumed consensus, which, in fact, does not exist.

Veridical models are difficult to learn, teach, and communicate, so there will be little consensus and little social error checking. The implications of social learning for veridical models depend on whether the model is presented at a high level and so with perceived transparency, or in all its detail. When veridical models are perceived as transparent, people will look at the simple model and go through an exercise like the following:

“I think I understand the model (I’m not sure as I know stuff is being left out). I think my understanding is shared by everyone. I expect the model to match the real world (after all the developers claim the model is veridical). I think my perception of what the model has to say about the real world is shared by others. I am probably right. I don’t

really need to check with others, but if there is any disconfirming evidence I will be ready to change my mind.”

This line of reasoning means that acceptance of the model will hinge on the interpretation that people make of the model. If the interpretation of the model does not match the user’s view of the real world, then the model would be wrong; regardless, of whether or not it matched any actual data on the real world. The reason that data would not outweigh opinion is because the model is sufficiently detailed that there is not sufficient data to validate all aspects of the model. This means that accurate models, at least models that are more accurate than opinion, may not be used to set policy. In contrast, if the user’s interpretation of the model does match the user’s view of the real world, then the model will be viewed as accurate, regardless of the force of evidence. In this case, decision makers may act overconfident in the model’s predictions.

When veridical models are not perceived as transparent then the story will change as follows:

“I know I don’t understand the model. My lack of understanding is shared by everyone. I expect the model to match the real world (after all the developers claim the model is veridical). I think my perception of what the model has to say about the real world is shared by others. Again I am not sure as the model is complex, so I may or may not be right.”

Again, acceptance of the model will hinge on the interpretation that people make of the model. If the user thinks his or her interpretation of the model does not match his or her view of the real world, then the user thinks that the model is “probably wrong”, regardless of the evidence, although evidence could be amassed to change the user’s opinion. As a result accurate models may be distrusted. On the other hand, if the interpretation of the model does match the user’s view of the real world, then the model is viewed as “probably right.” However, the tentativeness of this conclusion may lead to a lack of confidence in the model’s predictions. Since, veridical models may take many “lessons” to be learned. Imagine that what you first learn of the model is the high level simple and so transparent version. Then the details follow. This social learning process may lead to the problem that as people learn more about a veridical model their belief and confidence in model decreases.

Finally, the research on beliefs demonstrates that beliefs are a function of the individual's previous beliefs, the facts/information, and the beliefs of others I interact with. In addition, the impact of new information is a function of who sent the information, whether the information agrees with my current belief, and the weight/frequency of the information.

Social influence leads to simple and veridical models being believed, used, and thought of in very different ways. When a simple model is presented it is likely to be perceived as transparent; which is not to imply that it is transparent. People then decide if they believe it. Because the model is simple there are not multiple presentations; thus, there is a low flow of information. Consequently there are few opportunities to change opinion. Due to the factors discussed earlier, people tend to assume that other's share their understanding and interpretation. Thus, they do not seek information from others but assume that others have the same belief. A consequence is that people not only have beliefs about simple models but they have very strong beliefs. Due to the lack of presentations and the lack of information seeking, people rarely get contradictory information. Since people have strong beliefs they require a huge amount of contradictory information to change their beliefs. Thus, simple models, are likely to win or lose the day based purely on whether they are presented to a sympathetic audience.

In contrast, when a veridical model is presented, if it is presented in a high level fashion it is likely to be perceived as transparent. People then decide if they believe it. Since it is a complex model, people will recognize that they might not completely understand it so this will be a weak belief. Further, since it is a complex model, there are likely to be multiple presentations meaning that there is a high flow of information and many opportunities to change one's belief. The result, at least initially, is an increase in people's uncertainty about the model. Which leads to a general assumption that other's don't understand the model or interpret it in the same way. Consequently people are likely to seek information from each other, and this likelihood should increase at least initially. Since beliefs are initially weak it takes little information to change beliefs. Now, if people have access to more people than information about the model they will quickly come to take on the opinion of others. Thus, veridical models are likely to win or

lose the day based on whether there are more people in your social group or you have greater access to information about the model.

The Value of Real Transparency and Actual Veridicality

The application of the findings from cognitive science and behavior psychology thus suggest that the use of, belief in, and acceptance of models has more to do with social and cognitive processes than with the scientific process and the weight of evidence. This brings to the fore the question of whether or not there is any value to real transparency and actual veridicality.

There are in fact a number of benefits of real transparency. If models were truly transparent they would be easier to teach, learn and recode. Moreover, it should take less time to and space to explain as no discussion of interpretations would be needed. However, just because a model is transparent it does not guarantee comparability of the original and recoded results, due in part to both compiler issues and due to the fact that typically results presented for a model are the result of post-processing the model's results and such post-processing is rarely presented. Finally, transparency does enable theory building.

There are also many benefits of veridicality. Veridicality actually is valuable in explaining the model to decision makers or policy makers as you can appeal to the match with the real world. The closer the match the more the decision makers are able to reason within the model. However, actual veridicality leads to an increase in time for learning the model, and an increase in the amount of time and space to explain the model. The more veridical the model the more specific the predictions it generates. As such veridicality enables both policy analysis and managerial decision-making. Further, the more veridical the "easier" it is to validate in the sense that fewer assumptions need to be made about which real world data should be used to match the model. However, it is more difficult to validate in the sense that more data is needed. Interestingly, veridicality also creates transparent claims of applicability. Finally, veridicality enables theory construction.

The difference between the use and belief in models (as a function of the social and psychological processes) and the true value of veridicality and transparency leads to a great irony. Simple models are perceived as transparent and require little data to validate.

However, they generate only generic predictions with a plethora of interpretations and so are difficult to falsify. However, people don't recognize this morass of interpretations and so it is more likely that there will be greater belief in the truth of the simple models. Moreover they are likely to be viewed as having great utility and as improving theory.

In contrast, veridical models are perceived as being difficult to validate. However, they actually generate very specific predictions and are consequently more falsifiable. Yet, though they fit better in to the scientific process they are typically perceived as having less utility by basic researchers and as being further removed from theory and theory construction.

Conclusion

Human psychology coupled with the state of the social sciences has led to a misplaced trust in simple models. This is retarding the development of social and organizational engineering. This could have serious social and political consequences particularly if such simple models are used to set policy. This lack of trust in more veridical models is not shared by non social scientists. Consequently, they are more likely to develop the complex social and organizational models. Since they are subject to the same "naïve sociologist bias" and since they are unaware of findings they are likely to generate intuitive but inaccurate models. However, if policy makers and managers suffer the same "physics envy" as social scientists, these complex models built by non social scientists are likely to be believed simply by virtue of the discipline of the author.

So how do we solve the problem? Basically we need a shared infrastructure for social and organizational models. We need shared tool kits, shared data sets, databases linking papers, models, algorithms and data. We need in addition increased mathematical and computational training – not just statistics – in the social sciences. We need courses and textbooks on validations and analysis. Increased training on how to read and present models and model results is also called for. We need tools for visualizing highly veridical models. Moreover, we need more on-line journals with links to models.

Model simplicity and complexity is an axis of tension. This tension plays out in complex ways since simple models are often perceived to be transparent, whereas complex models are often argued to be more veridical. Science needs both transparency and veridicality. However, fundamental social and cognitive processes lead to model

development and use being based more on perceived transparency and believed veridicality rather than the actual transparency and veridicality of the model. Fundamental advances need to be made before the community can climb out of this quagmire.