On the Quality of a Social Simulation Model*

Claudio Cioffi-Revilla
Center for Social Complexity, Krasnow Institute for Advanced Study
George Mason University, Fairfax, Virginia 22030 USA
E-mail: ccioffi@gmu.edu

George Mason University

Abstract. Computational social science arises from several research traditions with roots in The Enlightenment and origins in Aristotle's comparative analysis of social systems. Extant standards of scientific quality and excellence have been inherited through the history and philosophy of science in terms of basic principles, such as formalization, testing, replication, and dissemination. More specifically, the properties of Truth, Beauty, and Justice proposed for mathematical social science are equally valid criteria for assessing quality in social simulation models. Helpful as such classic standards of quality may be, social computing adds new scientific features (complex systems, object-oriented simulations, network models, nonlinear dynamics) that require development as new standards of quality emerge. Social simulation models in particular (e.g., agent-based modeling) contribute further specific requirements for judging quality. This paper proposes and discusses a set of dimensions for discerning quality in social simulations, especially agent-based models.

Keywords: quality standards, evaluation criteria, social simulations, agent-based models, comparative analysis, computational methodology

1 Introduction: Motivation and Background

The field of social simulation in general, and agent-based modeling in particular, have begun to generate methodological proposals for assessing and promoting quality across diverse and related areas (Gilbert and Troitzsch, 2005; Taber and Timpone, 1996).[1] For instance, proposals exist in the area of communicating social simulation models (Cioffi and Rouleau, 2010; Grimm et al., 2005), comparing

* Prepared for the 2nd Annual Conference of the Computational Social Science Society of America, Santa Fe, NM, October 9–12, 2011. This paper was inspired by the First Workshop on Quality Commons, Maison de la Recherche, Paris, 28–29 January, 2010. Thanks to Petra Ahrweiler, Edmund Chattoe-Brown, Bruce Edmonds, Corinna Elsenbroich, Nigel Gilbert, David Hales, Dirk Helbing, Andreij Nowak, and Paul Ormerod for stimulating discussions, and the Center for Mathematics and Analysis in the Social Sciences, University of Paris Sorbonne, for local arrangements.

[1] This paper focuses on social simulations, so the broader field of computational social science (e.g., social data algorithms, complexity models, social networks, social GIS, and related areas of social computing) lies beyond the scope of this paper.

models (Rouchier et al., 2008; Cioffi, 2011; Cioffi and Gotts, 2003), and assessing complex projects that involve large interdisciplinary teams (Cioffi, 2010). Consensus on quality standards in social simulation has not yet emerged. However, the properties of "Truth," "Beauty," and "Justice" have been proposed and are widely used for discerning quality in *social science models* (Lave and March, 1993).

The three terms "Truth," "Beauty," and "Justice" (or "TBJ," for short) are labels for quality dimensions referring to fundamentally good—i.e., normatively desirable—features of social science modeling. Accordingly, the TBJ terms must be interpreted as labels, and not literally (Lave and March, 1993).

Truth refers to the empirical explanatory content of a model—its contribution to causal understanding—in the sense of positive theory. For example, truth is normally judged by internal and external validation procedures, corresponding to axiomatic coherence and empirical veracity, respectively (Kaplan, 1964; Sargeant, 2004). Truthfulness is the main classical criterion for evaluating empirical science (Hempel, 1965; Cover and Curd, 1998; Meeker, 2002), whether the model is statistical, mathematical, or computational. "Truth" must be a constituent feature in a social science model, or without it a model has no overall quality contribution.

Beauty refers to the esthetic quality of a model, to its elegance in terms of properties such as parsimony, formal elegance, syntactical structure, and similar stylistic features. Beauty is about art and form. For example, the mathematical beauty of some equations falls within this criterion, including features such as the style of a well-annotated system of equations where notation is clear, well-defined, and elegant. Unlike truth, beauty is not necessarily a constituent attribute, but is certainly a desirable scientific quality.

Justice refers to the extent to which a model contributes to a better world—to improvement in the quality of life, the betterment of the human condition, or the mitigation of unfairness. Justice is a normative criterion, unlike the other two that are positive or esthetic. For example, a model may improve our understanding of human conflict, inequality, refugee flows, or miscommunication, thereby helping to mitigate or improve social relations and well-being through conflict resolution, poverty reduction, humanitarian assistance, or improved cross-cultural communication, respectively. Policy analysis can be supported by modeling.

The Lave-March criteria of truth, beauty, and justice are useful for evaluating the quality of social simulation models. For example, in the classic Schelling (1971) model of segregation all three criteria are well-recognized. This is a fundamental reason why Schelling's model is so highly appreciated. Other examples that satisfy the Lave-March TBJ criteria might also include the Sugarscape model (Epstein and Axtell, 1996), the Iruba model (Doran, 2005), and Pick-a-Number (Hoffmann, 2002, 2005).

However, a further challenge exists because social simulations have features that render truth, beauty, and justice insufficient as criteria for assessing quality. This is because social simulation models are instantiated or rendered in code (a computer program in some language), so one can easily imagine a social simulation that would be of high quality in terms of truth, beauty, and justice, but fail in overall quality because simulation models pose additional challenges beyond other social science models (i.e., beyond the features of statistical or mathematical models).

As illustrated in Figure 1 (UML class diagram), social simulations have properties that are shared with all models in science generally and social science in particular, based on inheritance as a specialized class, in addition to having other features of their own. For example, the programming language of an agent-based model is a defining feature.

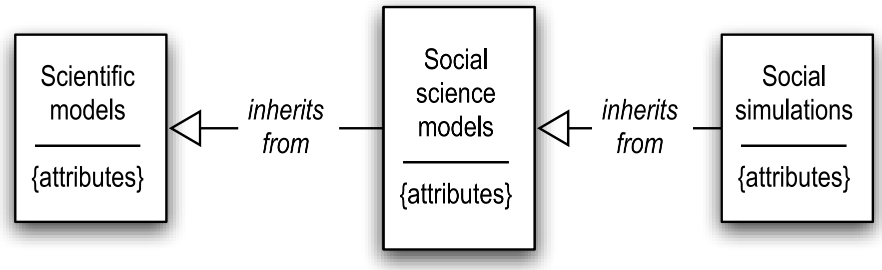

Fig. 1. UML class diagram illustrating the hierarchy of scientific models (left), social science models (center), and social simulations (right), each having increasingly specific standards for judging quality (left to right). *Source:* Prepared by the author.

The inheritance relation between social science models and social simulations readily suggests the specific features that distinguish the latter from the former, as illustrated in Table 1.

Table 1. Quality Criteria for Evaluating Models in Domains of Science

Models in ...	Truth	Beauty	Justice	Additional
Science	Yes	Yes	No	No
Social science	"	"	Yes	"
Social simulation	"	"	"	Yes

Source: Prepared by the author.

Additional criteria for social simulations—i.e., criteria beyond classical standards for social social models—should allow us to judge quality in terms of "The Good, The Bad, and The Ugly."

Common practices such as verification and validation are accepted quality control procedures for assessing the quality of scientific models in general (Cover and Curd, 1998). Verification and validation are insufficient criteria for assessing the quality of social science models and—specifically for social simulations. An important implication is that current emphasis on model verification and validation is warranted (Cioffi, 2010; Sargent, 2004), but *verification and validation are insufficient by themselves for judging the quality of a social simulation model* (agent-based or other).

Therefore, a key methodological question concerning quality is: Which additional criteria—i.e., beyond Truth, Beauty, and Justice—could or should be used to assess the quality of a social simulation model? The next section addresses this by proposing an initial set of dimensions for evaluating the quality of a given social simulation model.

2 Dimensions of Quality in Social Simulation Models

Arguably, there are two levels of quality assessment for computational social simulations, corresponding to the concepts of *a model* and *modeling*, respectively.

First, from a model's perspective, any set of quality dimensions for evaluating a social simulation must be based on its specific attributes or uniquely constituent features as a specific computational artifact in the sense of Simon (1996). Moreover, whether the overall quality of a given model should be an additive or a multiplicative function of individual qualitative features is less important than the idea that overall quality depends on a set of dimensions or desirable features beyond the Lave-March criteria, not on some single preeminent feature (e.g., simulation environment or programming language).

Second, from a modeling perspective, quality assessment should cover the broader modeling process as such, beyond the social simulation model that is produced in a narrow sense. This is because a computational model in final (i.e., committed) instantiated code is the result of a sequence of earlier modeling stages that precede the model itself, such as the critical stage of model design prior to implementation. Quality in design affects quality in the product of implementation, even when implementation *per se* is carrier out in a proper manner (i.e., competently, with effectiveness and efficiency).

The following set is proposed as a viable checklist of quality dimensions to consider, based on the preceding methodological principles for social simulation:

1. **Research question.** Is the research question or class of research questions clearly formulated? Every computational simulation model is designed to address a research question, so clarity is critical.
2. **Motivation.** Is the model properly motivated in terms of relevant extant literature? Or, is the simulation model the very first of its kind? If so, are there prior statistical or mathematical models in the same domain?
3. **Instantiation selection.** Does the code instantiate relevant social theory? Is the underlying social formal theory instantiated using a proper program or programming language?

4. **Code quality.** These may be collectively referred to as "Grimson-Guttag Standards:" Is the code well-written? Is the style safe/defensive? Is it properly commented? Can it be understood with clarity one year after it was written?
5. **Code class.** Is the model written in native code or using a toolkit? If toolkit, which, why, and how good is the application?
6. **Nuts and bolts.** Quality questions such as: What is the quality of the random number generator (RNG)? Think Mersenne Twister (Luke, 2011), MT19937, or other PRNG.
7. **Algorithm efficiency.** What is the implementation difficulty of the problem(s) being addressed by the model? How efficient is the code in terms of implementing the main design ideas?
8. **Computational efficiency.** How efficient is the code in terms of using computational resources? This differs from algorithm efficiency.
9. **Architectural design.** Is the code structured in a proper and elegant manner commensurate to the research question?
10. **Object ontology.** Does the model instantiate the object-based ontology of the focal system for the chosen level of abstraction?
11. **Network structures.** If networks are present and significant in the focal system, does the model exploit theory and research in social network analysis (Wasserman and Faust, 2005)?
12. **Complexity analysis.** Does the model facilitate analysis of complexity in the system of nonlinear interactions and emergent properties?
13. **GUI functionality.** Is the user interface of high quality according to the main users?
14. **Visualization analytics.** Is visualization implemented according to high standards of visualization analytics (Thomas and Cook, 2005)? This does not concern only visual quality (Tufte, 1990), but analytics for drawing valid inferences as well.
15. **Replicability.** What is the model's replication potential or feasibility?
16. **Overall effectiveness.** Does the model render what is necessary for answering the research question or class of research questions? This differs from efficiency.
17. **Simulation facilities.** Does the model possess the necessary functionality for conducting extensive computational analysis to answer the research questions?
18. **Experimental capacity.** How powerful is the model in terms of enabling critical or insightful experiments? For example, in terms of parameter exploration (evolutionary computation) and record-keeping.
19. **Communicative clarity and transparency.** Are useful flowcharts and UML diagrams of various kinds (class, sequence, state, use case) provided for understanding the model? Are they drawn with proper style (Ambler, 2005)?
20. **Pedagogical value.** Does the model teach well? I.e., does it teach efficiently and effectively?

21. **Curatorial sustainability.** How well is the model supported in terms of being easily available or accessible from a long-term perspective?
22. **Simulation infrastructure.** What is the quality of the infrastructure that renders the most effective simulation experience?
23. **Policy relevance.** Some social simulations are intended as policy analysis tools. Does the model add value to the overall quality of policy analysis?
24. **Other.** Additional criteria come to mind as a model is assessed by these and related dimensions...

3 Discussion

Given these criteria, the next challenge is to reduce this set of criteria to a cognitively reasonable number, such as seven or so dimensions. For example, some of these criteria may be viewed as utilitarian (e.g., based on resources), while others are non-utilitarian (based on style).

Quality must have dimensions because it is a latent concept, and not a directly measurable property. Therefore, proxies (i.e., measurable dimensions or attributes) are needed. The quality dimensions proposed in the preceding section provide a tentative framework and are a work-in-progress as social simulation develops as a field, not as a permanent set of fixed criteria.

Interestingly, Osgood's first dimension in cognitive EPA-space is Good-Bad (evaluation). This is why quality evaluation (good-bad-ugly) is essential (Osgood, May, and Miron, 1975). The proposed criteria should allow a classification of social simulations into categories of good, bad, or outright ugly.

As computational social scientists we need to better understand the micro-processes that compose the overall quality of social simulation:

- How is a problem chosen for investigation?
- How is the problem-space reduced by abstraction?
- How is the model designed?
- How well are the entities and relations understood?
- How is the simulation language chosen?
- How are verification and validation conducted?
- How are simulation runs actually conducted?
- Etc.

Requiring additional quality criteria for social simulation models is not an argument against the unity of science. It is a plea for greater specificity and more rigor in the evaluation of quality in the field of social simulation.

4 Summary

Computational social science arises from a number of research traditions that have roots in The Enlightenment and even earlier origins in Aristotle's comparative analysis of social systems. Therefore, our existing standards of scientific quality and excellence have been inherited through the history and philosophy of science in terms of basic principles, such as formalization, testing, replication, and dissemination.

More specifically, the properties of Truth, Beauty, and Justice proposed for mathematical social science (Lave and March, 1993) are equally valid quality criteria for assessing social simulation models. But useful as such classic standards of quality may be, social computing adds new scientific features (e.g., emphasis on understanding complex adaptive systems, object-oriented ontologies, network structures that can evolve in time, nonlinear dynamics) that require development as new standards of quality emerge. Social simulation models in particular (e.g., agent-based modeling) require further specific requirements for judging quality. This paper proposed and discussed a set of two dozen dimensions for discerning quality in social simulations, especially agent-based models. These criteria are offered as an initial heuristic framework to consider and develop as a work-in-progress, not as a finalized set of fixed criteria.

About the author

Claudio Cioffi-Revilla is Professor of Computational Social Science and Director of the Center for Social Complexity at George Mason University, Washington DC. He is co-founder and former President of the North American Association for Computational Social and Organizational Sciences NAACSOS and co-founder and former President of the Computational Social Science Society CSSS. His research interest focuses on agent-based and complexity-theoretic models of conflict in socio-natural systems (Asia, Africa), the origins of social complexity during the Neolithic age, and the methodology of computational social science, especially agent-based and network models. He also serves as a Jefferson Science Fellow of the National Academies (2006–2012) at the US Department of State.

Acknowledgements

Funding for this study was provided by the Center for Social Complexity of George Mason University and by ONR MURI grant no. N00014-08-1-0921. Thanks to members of the Mason-Yale Joint Project on Eastern Africa (MURI Team) for discussions and comments, and to Joseph Harrison for converting the earlier LaTeX article into LNCS style. The opinions, findings, and conclusions or recommendations expressed in this work are those of the author and do not necessarily reflect the views of the sponsors. This paper is dedicated to the memory of Auguste Comte (1798–1857), who lived at the Hotel de Saint-Germain-des-Pres, Paris, where this paper originated during the Quality Commons workshop.

References

1. Ambler, Scott W. 2005. *The Elements of UML 2.0 Style*. Cambridge and New York: Cambridge University Press.

2. Cioffi-Revilla, Claudio. 2008. Simplicity and Reality in Computational Modeling of Politics. *Computational and Mathematical Organization Theory* **15** (1): 26–46.

3. Cioffi-Revilla, Claudio. 2010. On the Methodology of Complex Social Simulations. *Journal of Artificial Societies and Social Simulations* **13** (1): 7. Available online.

4. Cioffi-Revilla, Claudio. 2011. Comparing Agent-Based Computational Simulation Models in Cross-Cultural Research. *Cross-Cultural Research* **45** (2): 1-23.

5. Cioffi-Revilla, Claudio, and Nicholas M. Gotts. 2003. Comparative Analysis of Agent-Based Social Simulations: GeoSim and FEARLUS Models. *Journal of Artificial Societies and Social Systems* **6** (4). Available online.

6. Cioffi-Revilla, Claudio, and Mark Rouleau. 2010. MASON RebeLand: An Agent-Based Model of Politics, Environment, and Insurgency. *International Studies Review* **12** (1): 31–46.

7. Cover, J. A., and M. Curd, eds. 1998. *Philosophy of Science: The Central Issues*. New York, NY: W. W. Norton.

8. Doran, James E. 2005. Iruba: An Agent-Based Model of the Guerrilla War Process. In *Representing Social Reality: Pre-Proceedings of the Third Conference of the European Social Simulation Association* (ESSA), ed. Klaus G. Troitzsch. Koblenz, Germany: Verlag Dietmar Foelbach. Pp. 198–205.

9. Epstein, Joshua M., and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, D.C.: The Brookings Institution.

10. Gilbert, Nigel, and Klaus Troitzsch. 2005. *Simulation for the Social Scientist*. Second edition ed. Buckingham and Philadelphia: Open University Press.

11. Grimm, Volker, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L. DeAngelis. 2005. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* **310**: 987–991.

12. Hempel, Carl G. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

13. Hoffmann, Matthew J. 2002. Entrepreneurs and the Emergence and Evolution of Social Norms. In *Proceedings of Agent-Based Simulation 3 Conference*, edited by Christoph Urban. Ghent, Belgium: SCS-Europe (2002): 32–37.

14. Hoffmann, Matthew J. 2005. Self-Organized Criticality and Norm Avalanches. In *Proceedings of the Symposium on Normative Multi-Agent Systems* AISB (2005): 117–125.

15. Kaplan, Abraham. 1964. *The Conduct of Inquiry*. San Francisco, CA: Chandler.

16. Lave, Charles A., and James G. March. 1993. *An Introduction to Models in the Social Sciences*. Lanham, MD: University Press of America.

17. Luke, Sean. 2011. Mersenne Twister. Available at: http://www.cs.gmu.edu/ sean/research/ Accessed May 30, 2011.

18. Meeker, Barbara F. 2002. Some Philosophy of Science Issues in the Use of Complex Computer Simulation Theories. In *The Growth of Social Knowledge: Theory, Simulation, and Empirical Research in Group Processes,* edited by Jacek Szmatka, Michael Lovaglia and Kinga Wysienska. Westport, CT and London: Praeger. Pp. 183–202.

19. Osgood, Charles E., W. H. May, and M. S. Miron. 1975. *Cross-Cultural Universals of Affective Meaning*. Urbana, Illinois: University of Illinois Press.

20. Rouchier, Juliette, Claudio Cioffi-Revilla, J. Gary Polhill, and Keiki Takadama. 2008. Progress in Model-to-Model Analysis. *Journal of Artificial Societies and Social Simulation* **11** (2–8).

21. Sargent, Rober G. 2004. Verification and Validation of Simulation Models. In *Proceedings of the 2007 Winter Simulation Conference*, eds. S. G. Henderson, B. Biller, M. H. Hsieh, J. Shortle, J. D. Tew and R.R. Barton. Piscataway, New Jersey: IEEE Press.

22. Schelling, Thomas C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* **1**: 143–186.

23. Simon, Herbert A. 1996. *The Sciences of the Artificial*. 3rd ed. Cambridge, MA: MIT Press.

24. Taber, Charles S., and Richard J. Timpone. 1996. *Computational Modeling*. Thousand Oaks, California, London and New Dehli: Sage Publications.

25. Thomas, James J., and Kristin A. Cook, eds. 2005. *Illuminating the Path.* Los Alamitos, CA: IEEE Computer Society.

26. Tufte, Edward R. 1990. *Envisioning Information.* Cheshire, CT: Graphics Press.

27. Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis.* Cambridge University Press.