A NEW RESEARCH ARCHITECTURE FOR THE SIMULATION ERA

Martin Ihrig
The Sol C. Snider Entrepreneurial Research Center
The Wharton School of the University of Pennsylvania
418 Vance Hall, 3733 Spruce Street, Philadelphia, PA 19104

KEYWORDS
Simulation, modeling, research, methodology, business, management, theory, social sciences.

ABSTRACT
This paper proposes a novel research architecture for social scientists that want to employ simulation methods. The new framework gives an integrated view of a research process that involves simulation modeling. It highlights the importance of the theoretical foundation of a simulation model and shows how new theory-driven hypotheses can be derived that are empirically testable. The paper describes the different aspects of the framework in detail and shows how it can help structure the research efforts of scholars interested in using simulations.

INTRODUCTION

Business and management researchers are increasingly interested in exploring phenomena that are emergent and/or one of a kind and in studying complex and non-repeatable processes. Simulation modeling is the appropriate methodological approach for this kind of research. Harrison, Lin, Carroll and Carley (2007: 1229) consider simulation modeling to be a "powerful methodology for advancing theory and research on complex behaviors and systems", while Davis, Eisenhardt and Bingham (2007: 480) point out that "the primary value of simulation occurs in creative and systematic experimentation to produce novel theory". Simulation research results in theory-driven frameworks and hypotheses that would be difficult to obtain from empirical analyses alone. In this paper, we propose a novel research architecture, a framework that gives an integrated view of a research process that involves simulation modeling. In what follows, we describe the different aspects of the framework and show how it can help structure the research efforts of scholars interested in using simulations.

AN 'EXTENDED' LOGIC OF SIMULATION AS A METHOD

Carley (2002: 254) gives a detailed explanation for "why so many social and organizational scientists and practitioners are turning to computational modeling and analysis as a way of developing theory and addressing policy issues". Among the many reasons she advances is that "social and organizational systems are complex non-linear dynamic systems; hence, computational analysis is an appropriate technology as models can have these same features." (Carley, 2002: 254) Harrison, Lin, Carroll and Carley note that "the academic field of management has been slow to take advantage of simulation methods" (2007: 1229). But looking at the increased number of recent articles based on simulation research in management journals and of simulation-specific workshops and papers at management conferences, it seems that management theorists are finally discovering the benefits of simulation methods. Davis, Eisenhardt and Bingham (2007) and Harrison, Lin, Carroll and Carley (2007) give guidelines for simulation research in the field of management. Gilbert and Troitzsch (2005) put forward the following framework (Figure 1) to explain the logic of simulation as a method in their authoritative book on simulation for the social scientist.

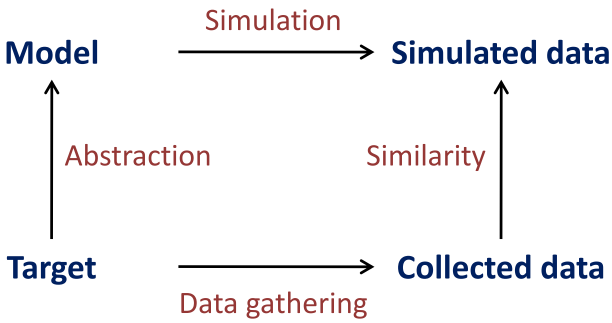

Figure 1: The logic of simulation as a method (Gilbert & Troitzsch, 2005: 17)

Starting at the bottom left, the real world 'target' under study is modeled by abstracting characteristics from it, and then a computer model is used to run simulations in order to produce simulated data. This data can then be compared with data collected in the 'real' social world.

This framework nicely illustrates the core logic of simulation as a method that underlies all the different kinds of simulation approaches that Gilbert and Troitzsch (2005) review in their book. But does it capture the entire simulation research process that scholars encounter, especially for complex agent-based modeling approaches? What is the role of existing theory, insights and frameworks from the literature, when it comes to modeling? Where do simulation environments, software toolkits that help researchers create, run, and analyze simulation models, come into play? To account for those questions and provide a more detailed picture of the simulation research life-cycle, we constructed an expanded framework (Figure

Proceedings 26th European Conference on Modelling and Simulation ©ECMS Klaus G. Troitzsch, Michael Möhring, Ulf Lotzmann (Editors)
ISBN: 978-0-9564944-4-3 / ISBN: 978-0-9564944-5-0 (CD)

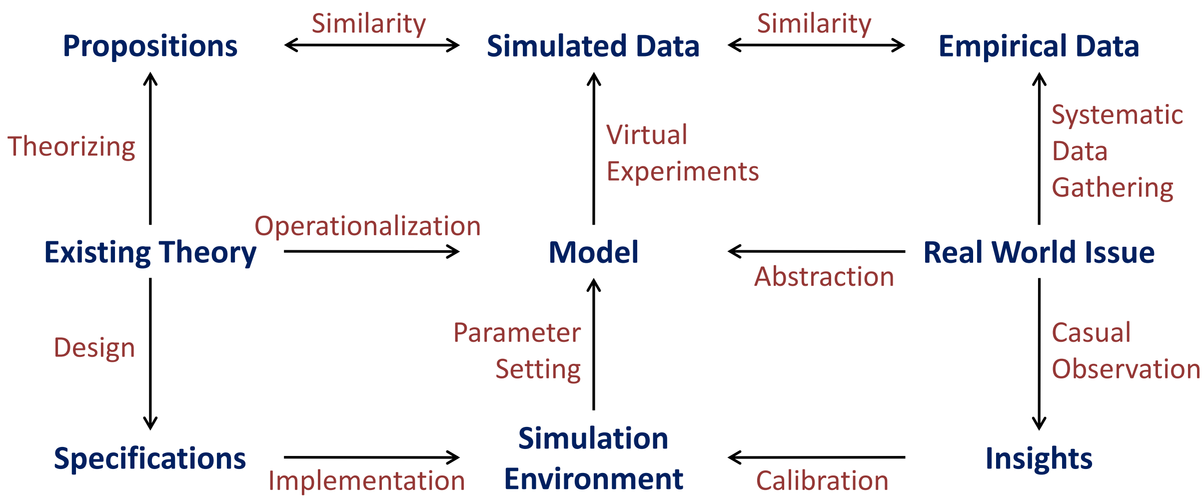

Figure 2: A new research architecture

2) to extend the logic of Gilbert and Troitzsch's (2005) basic model. This framework is meant to provide future simulation researchers with a research architecture that can assist them in their modeling efforts. Simulation methods are used for studying complex processes; simulation research, being a complex process itself, should therefore be described by an appropriately complex process model that depicts all the relationships between its building blocks. The following sections explain the different components of this research architecture.

FROM THEORY TO SIMULATION MODEL

In contrast to Gilbert and Troitzsch (2005), we propose to start the simulation research 'adventure' in the bottom half of the framework (Figure 3) on the left with a particular *theory*. The bottom left square describes the process of implementing theory with software, whereas the bottom right square depicts the real world grounding of the model.

Implementing Theory with Software. In a very general definition, a theory is "a statement of what causes what, and why" (Christensen, Carlile, & Sundahl, 2003). In order to build simulation software, one has to *design* and write a detailed document with *specifications* that mirror the theory in computer algorithms. This is both a conceptual and technical task, and requires the simulation researcher to be very specific about how "what causes what" is based on theoretical considerations ("why"). The next step is to *implement* the concrete specifications in computer code, and program the *simulation environment* as executable software. With the software up and running, finding the right *parameter settings* for the generic simulation environment can allow an application specific *model* to be created. The resulting model will *operationalize* the specific theory on which the simulation environment is based.

Real World Grounding. The purpose of simulation research is not necessarily to model a theory in the

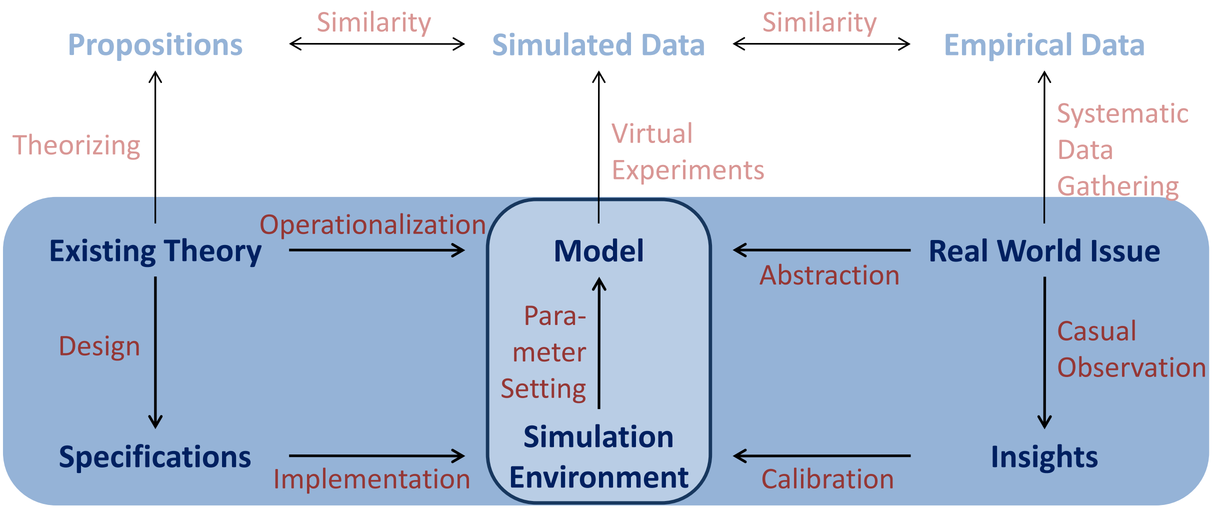

Figure 3: From theory to simulation model

abstract. The researcher wants to study a *real world issue* (what Gilbert and Troitzsch (2005) call a 'target') so as to produce new theory. *Casual observations* of the social world will yield helpful *insights* about the real world issue that will help *calibrate* the parameter space in the simulation environment: thus the final model will be an *abstraction* of the real world issue under study.

What distinguishes the simulation approach described here from many other simulation exercises is that a fully-fledged theory lies behind the simulation environment that is built. In the case of agent-based models and simulations (Gilbert, 2008), most modelers endow their virtual agents with only a couple of simple rules as an abstraction from real social agents (individuals, companies, etc.). Therefore, the entire bottom half of our research framework (and especially the left side) disappears, because the modeling lacks a theoretical underpinning. This is reflected in Gilbert and Troitzsch's (2005) simpler framework shown above. More simulation models are needed that have a stronger theoretical foundation, because "the advance of multi-agent techniques provides social scientists, who are used to thinking about agency, the ability to reason in the terms of their theories" (Carley, 2002).

SIMULATION RESEARCH IN ACTION

Once the researcher has built an application-specific model informed by the real world and grounded in some theory, the actual simulation research can start – the top part of our framework (Figure 4). The top left quadrant describes the process of new theory generation, and the top right links the simulation work to empirical follow-up studies.

The basic component of simulation research concerns running the simulation *model* and conducting many *virtual experiments* by varying the parameter space. The resulting *simulated data* can then be compared to both theoretical and empirical assessments, which is the first step in generating novel theory. The researcher will have *theorized* about the subject under study and will have come up with *propositions* based on the underlying *theory* used in the simulation effort. The simulated data can be evaluated in the light of these theoretical analyses and propositions, and the researcher can learn from studying *similarities* and *differences*. This exercise will result in theory-driven hypotheses that are empirically testable. Subsequently, the simulated data can be compared to *empirical data*. In an empirical follow-up study, the *real world issue* that is being investigated can be further examined by obtaining empirical data through *systematic data gathering* on a basis that is informed by the previous simulation research.

CONVENTIONAL RESEARCH APPROACH

The research architecture described above is more comprehensive than conventional approaches that do not employ simulation tools, and so is better suited for studying complex phenomena and obtaining new theoretical insights.

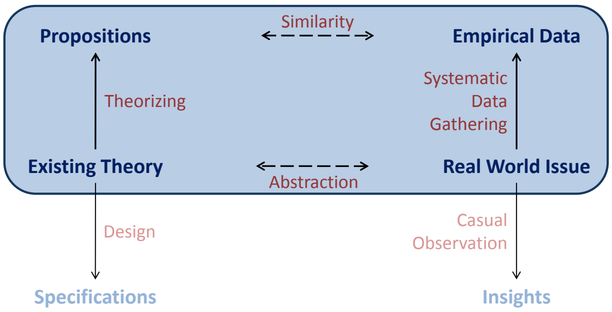

Figure 5: Conventional research approach

The conventional research approach can be depicted by connecting the left- and right-hand ends of our framework (Figure 5). Predictions and analyses are made based on existing theories, and the empirical data gathered on real world issues is compared to these theoretical accounts or propositions. What is lacking is

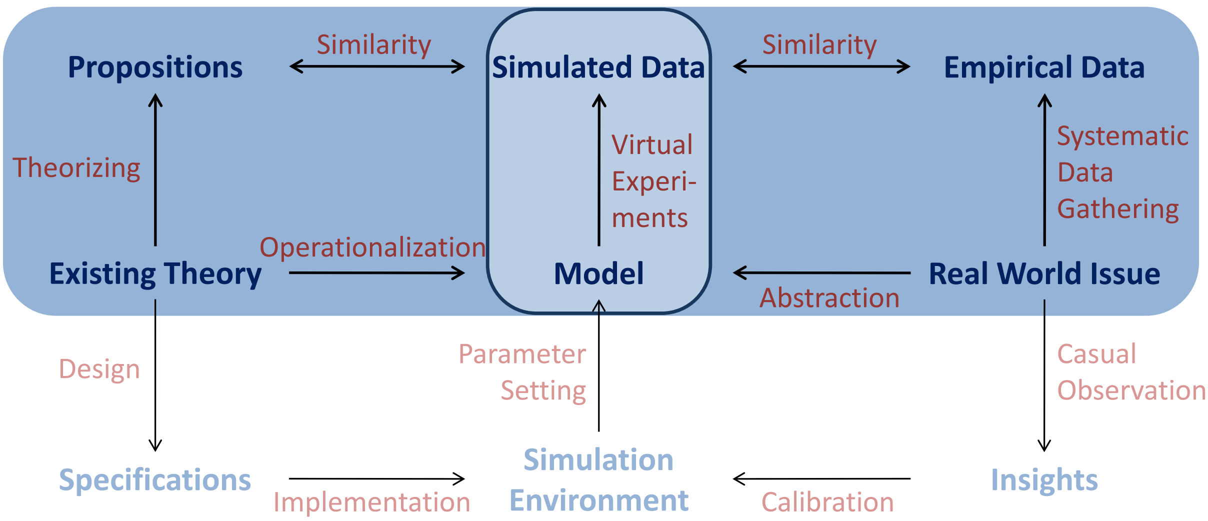

Figure 4: Simulation research in action

the power of computer tools that enable us to study more complex processes by modeling micro behaviors that individually might be straightforward, but may result in unpredictable outcomes when considered together.

VERIFICATION AND VALIDATION

Our framework can also be used to illustrate the important processes of verification and validation, both of which have been explained in detail in the literature (Carley, 2002; Davis, et al., 2007; Gilbert, 2008; Gilbert & Troitzsch, 2005; Harrison, et al., 2007). Gilbert and Troitzsch (2005: 23) give a concise definition:

> While verification concerns whether the program is working as the researcher expects it to, validation concerns whether the simulation is a good model of the target.

Our framework allows us to identify multiple instances where verification and validation come into play (Figure 6). There are three layers of interest: verification, validation, and ascertaining the two.

Verification. Starting at the bottom left, the obvious area where the model needs verifying is in the building of the software, where the researcher has to ensure that the program's technical specifications have been properly implemented. The simulation environment has to perform exactly as described in the technical document, without errors or 'bugs'. The simulation processes, in their abstracted form, also have to work like and be consistent with the real world social processes they represent.

Validation. The middle layer of the framework depicts the areas of interest in validation terms. Most researchers will look to the right, and ensure that the application-specific model fully represents the real world issue or target under study. But there is another important area: since the simulation environment is underpinned by a theory, they also have to make sure the model is a fair representation of the theoretical constructs.

Ascertaining Verification and Validation. The factors noted above will be difficult to ascertain without comparing the simulated data to either empirical data or theoretical predictions, or both. Therefore, simulation researchers have to pay attention to the top layer of the framework. They have to infer and draw conclusions from the actual results of the simulation runs to assess whether the program is working as intended, and represents the actual phenomenon studied. Great care must be taken here because - as Gilbert and Troitzsch (2005) point out - mistakes can occur at any step in the research process.

SIMULATION CAPABILITIES BEYOND A SINGLE RESEARCH PROJECT

Considering our research framework further, the top half of Figure 7 (shaded area, blue) maps the area usually covered by research employing simulation methods. The core research activities classically conducted in scholarly work based on simulation modeling as described by Gilbert and Troitzsch (2005) are those depicted in the top right quadrant. Many recent simulation studies also base their modeling on existing theories, represented by the top left quadrant (a fine example of this would be Csaszar & Siggelkow, 2010). Generally however, simple simulation models are programmed that can only be used for and applied to the particular research topic of the study. For this, an increasing number of researchers turn to preexisting modeling and simulation tools (environments) that support the construction of simulations (e.g., RePast for

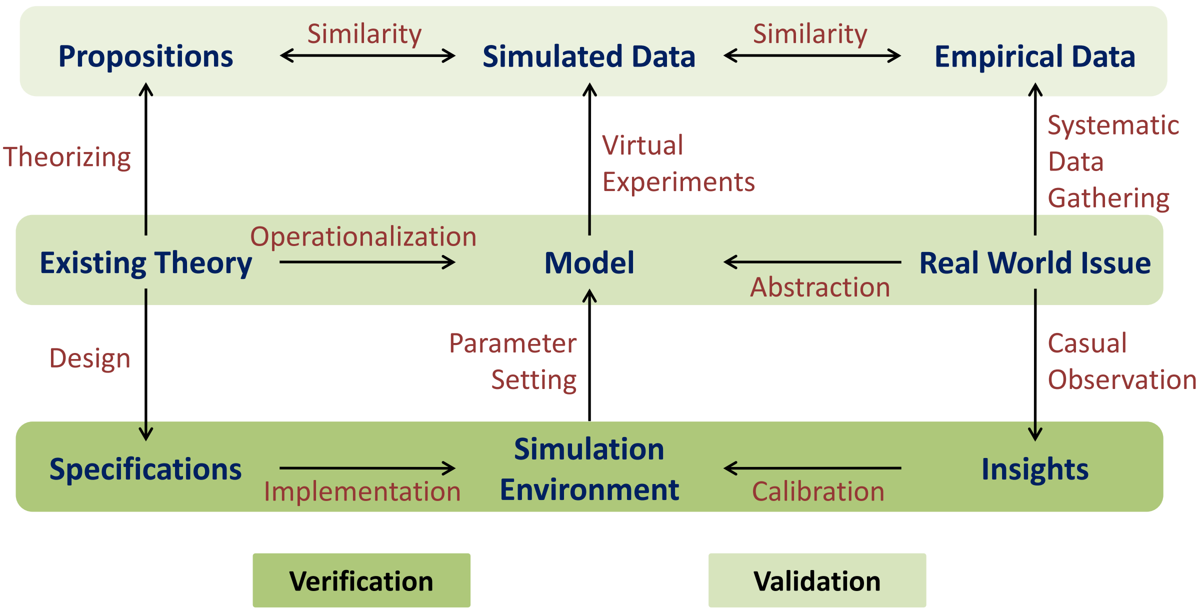

Figure 6: Verification and Validation

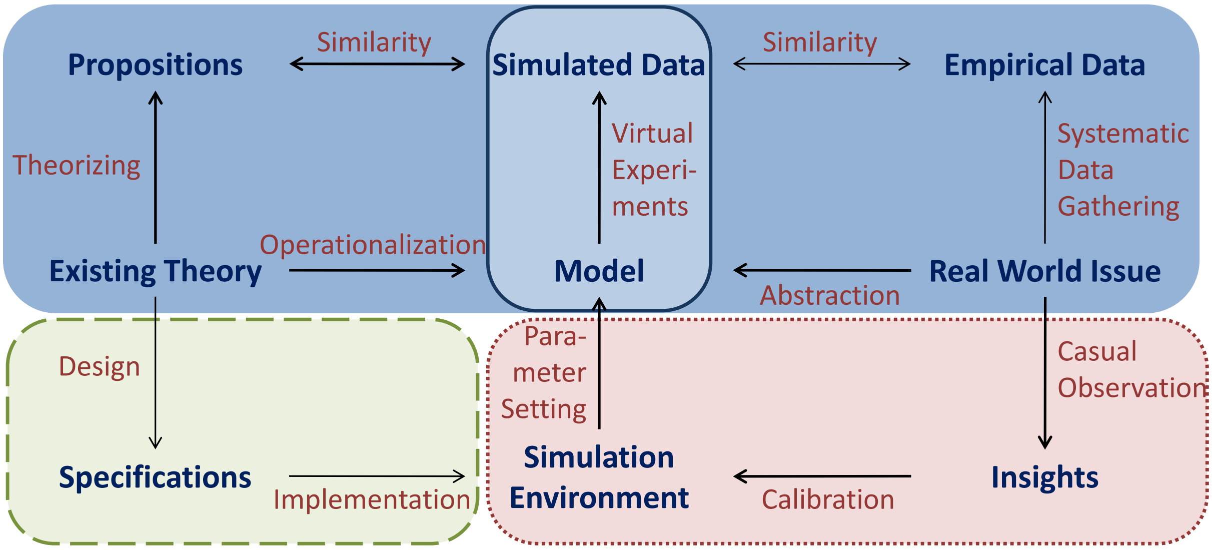

Figure 7: Simulation research activities

agent-based systems (North, Collier, & Vos, 2006)), which is represented by the bottom right quadrant (round dots, red). Very few simulation research projects cover the additional area marked in the dotted line in the bottom left (long dashes, green). Most researchers do not build an entire simulation environment that can be used for many different research topics and purposes, far beyond the subject of the immediate research needs. This is a pity because future research projects studying different topics could also employ the simulation environment created to derive new theory-driven hypotheses that are empirically testable and which range across numerous applications.

Designing and implementing a full simulation environment or platform (including a simulation execution and reporting suite) that can be used for a variety of research projects is a difficult and laborious process (Figure 8). This probably explains why many people shy away from developing them. Problems can occur at various places, especially if researchers are not able to write the actual software code themselves, but have to rely on software engineers who may not necessarily understand the theory fully. In fact, completely debugging a simulation environment can take several years. However, once developed, it benefits the many researchers who are interested in simulation methods but lack the computer science skills that are necessary to build the software. They can use the graphical user interface to easily set up, run, and analyze simulations that model their particular research questions.

To help build a simulation environment, some of the lessons learned from difficulties the author encountered during this effort are listed below. Attending to these four points will help simulation researchers to get closer to the proper computational representation of the theory they want to model and avoid losing too much time on software development.

1. If the software development project is inherited, or there are multiple authors, the researcher will have to revisit the specifications thoroughly and check whether they are all correct and appropriate, a process which involves uncovering and repairing inconsistencies, errors and omission in the spec.
2. If the researcher has to work with successive generations of programmers, and proper documentation is not in place, coherent knowledge about what has been implemented and how it has been implemented may be lacking. The software development process must be either very closely monitored over its entire development cycle, or rigorously and consistently documented.
3. Software bugs are an inevitable part of the development process, but having the program properly designed by a good software architect can at least avoid faulty code and inappropriate architecture.
4. Misrepresenting theory is dangerous: a researcher who is not their own software architect must attend to the specifications and technical document closely to ensure the computer algorithms really implement the theory the researcher wants to model.

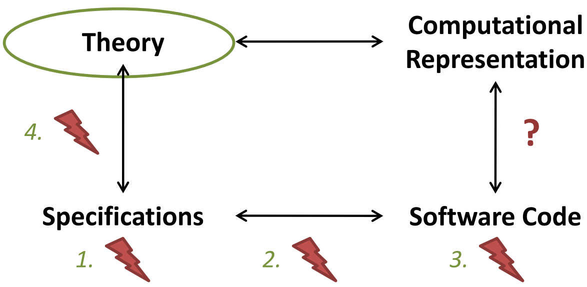

Figure 8: The long road to a simulation environment

CONCLUSION

This paper describes our unique research architecture and discusses the applicability of its framework, and looks ahead to many future research projects that could be structured using this approach. Being able to navigate the research space presented by this framework is a first step in using simulation methods to produce novel theory. We hope that the framework will help other researchers conceptualize the different tasks required to realize good simulation studies.

REFERENCES

Carley, K. M. (2002). Computational organizational science and organizational engineering. *Simulation Modelling Practice and Theory, 10*(5-7), 253-269.

Christensen, C. M., Carlile, P., & Sundahl, D. (2003). The process of theory-building. Retrieved 01.06.2004, 2004, from http://www.innosight.com./template.php?page=research#Theory%20Building.pdf

Csaszar, F. A., & Siggelkow, N. (2010). How Much to Copy? Determinants of Effective Imitation Breadth. *Organization Science, 21*(3), 661-676.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *Academy of Management Review, 32*(2), 480-499.

Gilbert, N. (2008). *Agent-based models*. London: Sage Publications.

Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist*. Maidenhead: Open University Press.

Harrison, J. R., Lin, Z., Carroll, G. R., & Carley, K. M. (2007). Simulation modeling in organizational and management research. *Academy of Management Review, 32*(4), 1229-1245.

North, M. J., Collier, N. T., & Vos, J. R. (2006). Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit. *ACM Transactions on Modeling and Computer Simulation, 16*(1), 1–25.

AUTHOR BIOGRAPHY

MARTIN IHRIG is President of *I-Space Institute, LLC* (USA) and Adjunct Assistant Professor at the *Wharton School* of the *University of Pennsylvania* (USA). He holds a Master of Business Studies from *UCD Michael Smurfit School of Business* (Ireland) and a Doctor of Business Administration (PhD) from *Technische Universität Berlin* (Germany).

The research initiative he manages at Wharton's *Snider Entrepreneurial Research Center* focuses on the strategic and entrepreneurial management of knowledge. In his simulation research, he is studying entrepreneurial opportunity recognition strategies with the help of agent-based models.

His e-mail address is ihrig@wharton.upenn.edu.