Developing Theory Through Simulation Methods

JASON P. DAVIS
Department of Management Science and Engineering
Stanford University
423 Terman Engineering
Stanford, California 94305
(650) 799-1642
jpdavis@stanford.edu

CHRISTOPHER B. BINGHAM
Robert H. Smith School of Business
University of Maryland
4519 Van Munching Hall
(301) 405-3422
cbingham@stanford.edu

KATHLEEN M. EISENHARDT
Department of Management Science and Engineering
Stanford University
415 Terman Engineering
Stanford, California 94305
(650) 723-1887
kme@stanford.edu

January 31, 2006

Forthcoming, *Academy of Management Review*, April 2007

We thank Ron Adner, Peter Glynn, Ashish Goel, Jesper Sorenson, Riitta Katila, and Jan Rivkin for helpful conversations. We also appreciate the partial support of the National Science Foundation (IOC Award #0323176) and the Stanford Technology Ventures Program.

ABSTRACT

Simulation is an important method, but its link to theory development remains unclear and even controversial. Our purpose is to clarify when and how to use simulation methods in theory development. First, we develop a *roadmap* for conducting theory development using simulation methods. It ranges from selecting the research question and simple theory to conducting verification and validation. The primary value of simulation occurs in creative and systematic experimentation to produce novel theory. Second, we position simulation methods within the *broad context* of theory development. Simulation sits in the "sweet spot" between theory creating using methods such as multiple case inductive studies and formal modeling, and theory testing using methods such as multivariate statistical testing of hypotheses. We note the strengths of theory building using simulation including internal validity, experimentation to create new theory, and facility in coping with longitudinal, non-linear, and process phenomena, especially when empirical data are challenging to obtain. We also note simulation's weaknesses such as external validity. We conclude with guidelines for evaluation that focus on the importance of theoretical contribution, strength of method, and the insights of the emergent theory.

Key words: simulation, theory development, methods

Simulation is an increasingly significant methodological approach to theory development in the organizations and strategy literatures (e.g., Lant & Mezias, 1990; Adner, 2002; Repenning, 2002; Rivkin & Siggelkow, 2003; Zott, 2003). Indeed, several influential research efforts (e.g., Cohen et al., 1972; March, 1991) have used simulation as their primary method. Yet, while simulation has become an important methodology, its value for theory development remains clouded and even controversial.

On the one hand, some argue that simulation methods contribute effectively to theory development. For example, simulation can provide superior insight into complex theoretical relationships among constructs, especially when challenging empirical data limitations exist (Zott, 2003). It can also provide an analytically precise means of specifying the specific assumptions and theoretical logic that lie at the heart of verbal theories (Carroll & Harrison, 1998; Kreps, 1990). Simulation can also clearly reveal the outcomes of the interactions among multiple underlying organizational and strategic processes, especially as they unfold over time (Repenning, 2002). From these perspectives, simulation can be a powerful method for sharply specifying and extending extant theory in useful ways.

On the other hand, some researchers maintain that simulation methods often yield very little in terms of actual theory development. They suggest that simulations are simply "toy models" of actual phenomena in that they either replicate the obvious, or strip away so much realism that they are simply too inaccurate to yield valid theoretical insights (Chattoe, 1998; Fine & Elsbach, 2000). For example, simulation research is usually based on at least some clearly unrealistic assumptions such as zero search costs (Rivkin, 2000) and all strategic rules are effective (Davis, Eisenhardt, and Bingham, 2005). In addition, simulation constructs are often "measured" by empirically distant approaches such as "0" and "1" bit strings as representations of organizations (Bruderer & Singh,

1996) and strategies (Rivkin, 2001). The results of research using simulation methods can also be dynamically indeterminate and overly complex (Fichman, 1999). From these perspectives, the value of simulation methods for theoretical development is tenuous.

The controversy surrounding the value of simulation methods for theory development partially arises, in our view, from a lack of clarity about the method and its related link to theory development. There appears to be limited understanding within the broad research community about (1) when simulation is a useful methodological choice for theory development, (2) how to select among the various simulation approaches (e.g., system dynamics vs. genetic algorithms), (3) the appropriate steps for performing simulation research, and (4) the relevant criteria for evaluating simulation research. Most significant, there seems to be limited recognition within the research community of how simulation methods fit into the broader scheme of relating methodological choices to theoretical development.

Our purpose is to address these issues by clarifying when and how to use simulation methods in theory development. Although scholars writing about theory development (e.g., Dubin, 1976; Pfeffer, 1982; Priem & Butler, 2001; Sutton & Staw, 1995; Whetten, 1989) may have different emphases, most agree that theory has four elements: constructs, propositions that link those constructs together, logical arguments that explain the underlying theoretical rationale for the propositions, and assumptions that define the scope or boundary conditions of the theory.[i] Consistent with these views, we define theory as consisting of constructs linked together by propositions that have an underlying, coherent logic and related assumptions.

Broadly, we attempt to make two contributions. First, we offer a *roadmap* for how to use simulation methods to develop theory. This roadmap synthesizes prior work on the design of simulation research (e.g., Sterman, 2000) and extends that work into specific areas such as identifying appropriate research questions, choosing among simulation approaches, developing computational representations, elaborating the fundamental role of verification v. experimentation, and evaluating simulation research. We rely on exemplars from the extant literature to ground our observations. The intended result is a more complete roadmap for executing and evaluating theory development research using simulation methods than has previously existed. This roadmap is summarized in Table 1.

Second, we position simulation methodology within the *broad context* of theoretical development in the organizations and strategy literatures. Relying on related theory development literature (e.g., Priem & Butler, 2001; Sutton & Staw, 1995; Weick, 1989), we explore the rationale for using simulation, its relationship to other methods such as laboratory experiments, simulation's strengths and weaknesses, and guidelines for evaluating theoretical development using simulation. We argue that simulation is especially useful in the "sweet spot" between theory creating research using methods such as inductive case studies (Eisenhardt, 1989) and formal modeling (Freese, 1980), and theory testing research using multivariate, statistical analysis (Pfeffer, 1992). That is, simulation enables the elaboration of rough, basic (or what we term "simple") theory that is often derived from inductive cases or formal modeling into logically precise and comprehensive theory. This theory can then be more effectively tested using deductive logic and empirical evidence. Simulation is particularly effective when the theoretical focus is longitudinal, non-linear, or processual, or when empirical data are challenging to obtain.

We begin by discussing the background of simulation methods. We then turn to the basic steps of conducting simulation research including selecting research questions, choosing a computational approach, and developing experiments. We conclude with a discussion of the strengths and weaknesses of simulation research, and guidelines for its evaluation.

BACKGROUND

We define simulation as a method for using computer software to model the operation of "real-world" processes, systems, or events (Law & Kelton, 1991). This definition is consistent with other definitions that describe simulation models as virtual experiments (Carley, 2001), or as simplified pictures of the world having some, but not all, of the characteristics of that world (Lave & March, 1975). In particular, simulation involves creating a computational representation of the underlying theoretical logic that links constructs together within these simplified worlds. These representations are then coded into software that is run repeatedly under varying experimental conditions (e.g., alternative assumptions, varied construct values) in order to obtain results.

While simulation can be used purely for description or exploration[ii], our focus is on using simulation for theory development when simple theory exists (e.g., Rivkin, 2000; Rudolph & Repenning, 2002). By *simple theory*, we mean undeveloped theory that has only a few constructs and related propositions with modest empirical or analytic grounding such that the propositions are in all likelihood correct, but are currently limited by weak conceptualization of constructs, few propositions linking these constructs together, and/or rough underlying theoretical logic. Simple theory also includes basic processes that may be known (e.g., competition, imitation), but that have interactions that are only vaguely, if at all, understood. Thus, simple theory contrasts with well-developed theory such as institutional and transaction cost theories that have multiple and clearly defined theoretical constructs (e.g. normative structures, mimetic diffusion, asset specificity, uncertainty), well-established theoretical propositions that have received extensive empirical grounding, and well-elaborated theoretical logic. Simple theory also contrasts with situations where there is no real theoretical understanding of the phenomena.

Simulation is particularly suited to the development of simple theory because of its strengths in enhancing theoretical precision and related internal validity, and in enabling theoretical elaboration and exploration through computational experimentation. In particular, simulation relies on some theoretical understanding of the focal phenomena in order to construct a computational representation. Yet, simulation also depends upon an incomplete theoretical understanding such that fresh theoretical insights are possible from the precision that simulation enforces and the experimentation that simulation enables. In contrast, while it is possible to simulate well-developed theories such as institutional and transaction cost theories, they offer fewer opportunities for new theoretical insights using the logical precision and experimentation that simulation enables. These theories are already well-elaborated. So, there is likely to be less payoff to simulation because it is likely just to replicate known theoretical ideas. At the other end of the theory development spectrum are "clean-slate" theoretical situations. Here there is not much to simulate, and other methods such as multiple case induction and formal modeling are typically better methodological choices. In contrast with these extremes, simple theory has enough theoretical development to construct a simulation, but also sufficient room to improve internal validity and to develop novel theoretical insights. In addition, simulation is especially useful for theory development when the focal phenomena involve multiple and interacting processes, time delays, or other non-linear effects such as feedback loops and thresholds. In these situations, simulation is likely to reveal non-intuitive elaborations of simple theory that are difficult to uncover using other methods including simple thought processes.

Overall, the central value of simulation for theory development lies in the exploration, elaboration, and extension of simple theories. Theory development begins with one or several theoretical ideas (what we term "simple theory") that are then represented in computer software, verified, and explored through computational experimentation. These experiments can range from incremental (e.g., adding a moderating relationship) to elaborate involving numerous, wide-ranging experiments that push the theory well beyond its immediate application (e.g., examining alternative theoretical logics) (Lee, Mitchell & Sablynski, 1999)[iii]. Several exemplar studies using simulation in theory development are in Table 2.

ROADMAP FOR DEVELOPING THEORIES WITH SIMULATIONS

Begin with an Intriguing Research Question and Simple Theory

Like all good research, studies that develop theory through simulation should begin with an intriguing research question that reflects deep understanding of the extant literature and relates to a substantial theoretical issue (Weick, 1989). Without such a question, simulation research simply becomes a "fishing expedition" in which the researcher lacks focus and theoretical relevance, and risks becoming overwhelmed by computational complexity.

Research questions can originate from many sources. Sometimes they come from conundrums within basic science. March (1991), for example, relied on complexity theory from the biological and computer sciences to conceptualize a research question that examined the tradeoff between the exploration of new possibilities and the exploitation of old certainties. Sometimes research questions are motivated with intriguing observations from inductive case study research. For example, Rudolph and Repenning (2002) used the counter-intuitive observations from a case study describing the 1977 Tenerife airport disaster (Weick, 1993) as the inspiration for their research question asking how small problems can become major catastrophes. Sometimes research questions emerge from combining process theories such as when Gavetti and Levinthal (2000) asked how cognition influences experiential learning. Research questions may also come from the classic tradition extending formal analytic models (Adner, 2002).

While simulation shares an emphasis on intriguing and theoretically relevant research questions with other methods, simulation is particularly suited to the theoretical development of simple theory. As described earlier, simple theory is undeveloped theory that involves a few constructs and related propositions with some empirical or analytic grounding, but that is limited by weak conceptualization, few propositions, and/or rough underlying theoretical logic. Simple theory may also include concepts and basic processes from well-known theories (e.g., competition, imitation), especially when the research focus is on their vaguely (if at all) understood interactions. Propositions may be formally stated (Davis, Eisenhardt & Bingham, 2005; Rivkin, 2001) or be implicit (Rudolph & Repenning, 2002). The fundamental idea is that theory development using simulation should begin with a simple theory, rather than either an extensive theoretical base or a clean theoretical slate. This simple theory should address a theoretically intriguing research question that focuses on a fundamental phenomenon. Such theory can then be a platform from which powerful theory can be developed through the verification and experimentation that is enabled by simulation (Lave & March, 1975; Stinchcombe, 1968).

Rivkin (2001), for example, focused on a research question that asked what is the optimal level of strategic complexity. He relied on a single proposition linking strategic complexity with performance that was based on case studies and prior theoretical work. This proposition was that a moderate level of strategic complexity leads to the highest performance. Similarly, Rudolph and Repenning (2002) used previous case study research as the basis of a simple theory that described how minor events could create major catastrophes. The theory consisted of several propositions that linked quantity of interruptions, stress, and performance.

Simulation is particularly effective when the simple theory involves several basic processes such as competition and legitimation (Lomi & Larsen, 1996) or imitation and experimentation (Zott, 2003) with only vaguely (if at all) understood longitudinal interactions. These interactions are often difficult to study with traditional statistical methods or to anticipate with thought processes. In contrast, these processes can usually be readily computationally represented, verified, and then explored (separately and in interaction) using simulation. For example, Sastry (1997) investigated the longitudinal interaction between inertial and change processes. Her results included unexpected insights about time-pacing and dynamic markets that were not anticipated by previous theory and evidence (Tushman & Romanelli, 1985).

Simulation is also particularly effective for theory development when the research question involves a fundamental tension or trade-off. The tension may be temporal such as short vs. long run implications (March, 1991; Sterman et al., 1997), structural such as too much structure vs. too little (Davis, Eisenhardt, & Bingham, 2005; Rudolph & Repenning, 2002), or spatial such as near vs. far away (Lomi & Larson, 1996; Schelling, 1971). These tensions often result in nonlinear relationships such as tipping point transitions and steep thresholds. These and other nonlinear relationships are difficult to discover using inductive case methods and to explore with traditional statistical techniques. Yet, they often offer surprising, non-intuitive results.

Choose a Simulation Approach

Assuming that simulation is the best way to proceed (e.g., intriguing research question and simple theory), the next step is to select a simulation approach. This choice should depend upon the fit of the research question, assumptions, and theoretical logic of the simple theory with those of the simulation approach. This choice is crucial because the simulation approach can impose a theoretical logic, type of research question, and related assumptions. Much as the choice of a statistical technique (e.g., OLS regression v. event history) can shape theory-testing empirical research, the choice of simulation approach shapes the subsequent results. In fact, the choice of simulation approach may be closer to choosing a theoretical framework (e.g., resource dependence v.

transaction cost economics) because of its framing of research questions, key assumptions, and theoretical logic.

Several well-known simulation approaches have been used for theory development in the organizations and strategy literatures[iv]. Some are common such as system dynamics (Repenning, 2002; Rudolph & Repenning, 2002) and NK fitness landscape (Rivkin & Siggelkow, 2003; Levinthal, 1997). Others are less frequently used such as genetic algorithms (Bruderer & Singh, 1996) and cellular automata (Lomi & Larsen, 1996). Each of these is a structured approach that constrains the theoretical logic, assumptions, and research questions that can be explored. In contrast, the stochastic process approach (March, 1991; Zott, 2003; Davis, Eisenhardt & Bingham, 2005) is a customized approach that is useful when the structured approaches do not fit with the research at hand (see Table 3 for comparison of these approaches).

System dynamics. System dynamics focuses on how causal relationships among constructs can influence the behavior of a system (Forrester, 1961; Sastry, 1997; Sterman et al., 1997). The approach typically models a system (e.g., organization) as a series of simple processes with circular causality (e.g., variable A influences variable B which influences variable A). These processes have some common constructs, and so intersect in a set of circular *causal loops*. These causal loops can be positive such that feedback is self-reinforcing and amplifying, or negative such that feedback is dampening (Sterman, 2000). While each process may be well understood, their interactions are often difficult to predict. The system typically includes *stocks* that act as buffers (i.e., constructs with values that accumulate and dissipate over time, and so introduce time delays), and *flows* (i.e., constructs specifying temporal rates in the system).

Rudolf and Repenning (2002), for example, used system dynamics to examine why minor interruptions sometimes trigger sudden catastrophes within organizations. The authors used intersecting causal loops to model two different processes (one positive loop, one negative loop) by which stress, quantity of interruptions, and performance were causally related (Yerkes & Dodson, 1908). The number of accumulated interruptions was included as a stock, and the rate of interruptions and the rate of their resolution were flows. By varying these rates, the authors were able to develop theory about when the system would be stable, when it would hit tipping points, and whether (and if so, how fast) those tipping points would lead to catastrophe.

System dynamics is particularly applicable for understanding the behavior of systems with complex causality and timing. Research questions are often framed as asking how specific initial conditions affect the stability of the system. That is, researchers are usually interested in finding the initial conditions that lead to abrupt, non-linear changes such as tipping points, catastrophes, and the emergence of vicious or virtuous cycles. Researchers often begin with two causal loops, and then add successive ones in order to build intuition, understanding, and realism in a structured way. Although the approach can accommodate some stochasticity, it relies on deterministic differential equations, and so is not as useful when there are many or complicated sources of stochasticity.

NK Fitness Landscape. The NK approach was developed in evolutionary biology to study genetic systems (Kauffman, 1993). This approach focuses on how rapidly and effectively a modular system adapts to reach an optimal point, especially when interactions among the system components are crucial. Specifically, the system (e.g., organization, product, strategy) is conceptualized as a set of N *nodes*, and K *interactions* among the nodes. These systems are assumed to use adaptation (sometimes termed "search") strategies such as incremental moves and long jumps to find the optimal point (e.g., best organization, best product). To illustrate, Rivkin (2001) focused on strategy. He defined N as the elements of a strategy (e.g., manufacturing and advertising choices) and K as the degree of interaction among the decisions. The optimal point was the highest performing strategy.

A key concept is the *fitness landscape* (Wright, 1931) that is created by assigning performance values (termed fitness) to every combination of values. The shape of the landscape is dependent on the interaction (K) among the nodes (Kauffman, 1989). When there is little interaction (low K), there is one optimal combination (or perhaps a few). The corresponding fitness landscape is "smooth" with only one or a few hills such that it is easy to find the optimal point. As interaction (K) increases, more combinations become locally optimal. The corresponding fitness landscape is "rugged" with many hills of varying heights that correspond to varying performance levels. This landscape type is hard to traverse to find the optimal point (Kauffman, 1989). The overall theoretical logic of the NK approach emphasizes the adaptation of a modular system using specific search strategies to find an optimal point on a fitness landscape.

For example, Gavetti and Levinthal (2000) examined how experiential learning (alone and with cognition) affected the time to find an optimal policy for an organization[v]. They represented organizational policy as N policy elements and K interactions among them. They then compared the time to find an optimal policy (and its performance) using only experiential learning (i.e., incremental moves) with using both experiential learning and cognition (i.e., a basic "map" of the landscape that enabled long jumps) under varying levels of coupling (K) among elements.

The NK approach is particularly applicable for understanding how the speed and effectiveness of adaptation to an optimum within a modular system is influenced by tight v. loose coupling among its components. Research questions are often framed as "problem solving" or "searching" to find an optimal point. Researchers often ask how long does it take to find an optimal point (e.g., high performing strategy, high performing product) or what is the effectiveness or performance of the optimal point (e.g., local v. global optimum). They are typically interested in how the number of nodes (e.g., number of product features), the interaction among nodes (e.g., tight v. loose coupling), types of adaptation (i.e., long jumps v. incremental moves), or the use of landscape "maps" (e.g., cognition, science) influences the time to find an optimal solution and/or the performance of that solution. Although the approach can be modified to accommodate some environmental dynamism (e.g., environmental jolts) or node intelligence, it is less useful when the effects of market dynamism, the intelligence and/or uniqueness of nodes in a modular system (e.g., differences among business units in an organization), or varied types of interaction (e.g., strong v. weak ties) are a central interest.

Genetic algorithms. Like the NK approach, genetic algorithms are an optimization approach with roots in biology. But unlike the NK approach, genetic algorithms focus on how rapidly and effectively a population of heterogeneous agents composed of genes (e.g., population of organizations composed of routines, population of consumers with preferences) adaptively learns (Goldberg, 1989; Holland, 1975)[vi]. Adaptation occurs through a stochastic evolutionary process (i.e., variation-selection-retention) that favors gradual improvement through accumulated experience (Aldrich, 1999; Holland, 1975). Variations occur in two ways: *mutation* (i.e., random change of one or a few genes that corresponds to mistakes) and *crossover* (i.e., random switching of sets of genes among agents that corresponds to recombination of existing genes). *Selection* of agents (also termed genes) occurs according to their performance (also termed fitness) with respect to some metric. *Retention* (also termed reproduction) is the copying of the selected agents from one generation to the next. Over time, successful variations are more likely to be retained and form the basis of future variations (Andreoni and Miller, 1995; Arifovec et al. 1997). Eventually, only high performing agents remain in the population and ultimately often only a single agent form survives (e.g., dominant design). The theoretical logic emphasizes the evolutionary adaptation of a population towards an optimal form that is path-dependent and combinatorial.

For example, Bruderer and Singh (1996) used a genetic algorithm to examine organizational evolution within a population of organizations. In their model, each of the 250 organizations in the population was composed of 20 routines that could change as a result of both mutation and crossover variation processes. The authors were particularly interested in the effect of learning on organizational evolution. They found that learning accelerated the discovery of an effective organizational form. Genetic algorithms can also be used to examine the evolution of specific types of strategies within an agent. For example, Zott (2002) used a genetic algorithm approach to model the bargaining behavior of a negotiator with private information. The genes were the negotiator's bargaining rules that determined when to offer a contract and when to accept or reject another's offer. Among other insights, the study clarified the effects of complete and incomplete information in producing inefficient (i.e., non-optimal) outcomes such as delays and failure to agree.

Genetic algorithms are particularly applicable for describing how heterogeneous agents (e.g., organizations, consumers) learn improved solutions (e.g., better organizational form, better strategy) through experimentation. As such, it is consistent with some important aspects of human learning such as experience-based action, importance of recent experience, creative synthesis of ideas, and occurrence of mistakes (Zott, 2002). Research questions are typically framed as asking what affects the rate of adaptation (or sometimes learning or change) and does a dominant form emerge. Researchers often experiment with rates and processes of mutation and crossover, and performance metrics. Although the approach can be modified, it is less useful when the value of the performance metrics varies (e.g., as in dynamic markets), the agents can engage in sophisticated cognitive processes that are not well-captured by experiential learning (e.g., foresight), or crossover is unlikely (e.g., technical standards limit reuse of product components across products in a population, internal labor markets limit mobility across organizations in a population).

Cellular automata. Although not widely used in the organizations and strategy literatures, cellular automata have gained traction in the physical sciences as an approach that focuses the emergence of macro-level system patterns from micro-level interactions among spatially related and semi-intelligent agents (Wolfram, 2002). Cellular automata assume a system of agents that are spatially related (e.g., competitors in an industry, cities in a country). *Spatial relatedness* implies that the degree to which agents influence each other depends upon the distance between them. The agents behave according to a few simple rules (i.e., semi-intelligent agents), some of which relate to how the agents influence each other (Langton, 1984). Typically, the rules relate to spatial processes such that nearby agents are influenced more than distant ones (e.g., diffusion, propagation, and competition processes) (Schelling, 1971). Although not required, the rules are usually uniform (i.e., all agents have the same rules) and deterministic (i.e., all rules are fixed). An important concept is the *neighborhood* that defines which agents are local (and so neighbors) and which are not. At least some rules designate behaviors taken in interaction with neighbors, but not with more distant agents.

For example, Lomi and Larsen (1996) used cellular automata to study density-dependence theory and the tension between competition and legitimation processes (Hannan & Freeman, 1989)[vii]. The organizations were spatially arrayed relative to one another in a 2-dimensional grid. The organizations possessed rules for competition (affecting only those in the neighborhood) and legitimation (affecting all organizations). The authors then observed how the competitive and legitimating interactions among organizations (i.e., micro-level interactions) affected population density, founding rates, and failure rates (i.e., macro-level patterns) over time.

Cellular automata are particularly useful for examining how macro-level patterns emerge from spatial processes (e.g., diffusion, competition, propagation, and segregation) that operate at a micro-level. In these processes, the behaviors of agents affect each other, but their influence diminishes with distance. Research questions usually ask how macro-level patterns emerge and change. Researchers typically vary the size of the neighborhood, the relevant processes (e.g., diffusion, segregation) and their related rules, and the spatial array of agents (e.g., density of agents) in order to describe the emergence and change of macro-level phenomena from micro-level interactions among agents. While cellular automata can be modified, the approach is less useful when intelligence resides at the system level (e.g., CEO in an organization of business units) or the interactions among agents are not spatially dependent.

Stochastic processes. Stochastic processes are a flexible approach that enables researchers to custom-design their simulations (Gallager, 1996).[viii] The approach makes no particular assumptions about the system, research question, or theoretical logic. It is especially useful for exploring theories that do not fit with the theoretical logic and assumptions of the structured simulation approaches. For example, the environment may be dynamic (March, 1991; Davis, Eisenhardt, & Bingham, 2005) or the theoretical logic may involve temporal transmission processes (Carroll & Harrison, 1998). Researchers typically piece together distinct *processes* that mirror the theoretical logic. They also build in several *sources of stochasticity* (e.g., environmental input, timing, elements of the process), and endow them with *stochastic distributions* that can be simple (e.g., 50/50 draw) or complex (e.g., normal and Poisson distributions). A key point is that, while the simulation is custom-designed and may have some original processes or constructs, researchers often utilize existing building blocks such as known processes (e.g., Markov chains), and familiar sources of stochasticity and their related distributions (e.g., Poisson distribution for input arrival times) (Law & Kelton, 1991).

An example is research by Carroll and Harrison (1998). Building on their previous work (Harrison & Carroll, 1991), these authors developed several underlying theoretical logics for the simple, well-known proposition positing that heterogeneity of tenure (LOS) is related to heterogeneity of culture. These logics related to temporal transmission of culture, and so did not fit with the theoretical logics of structured simulation approaches such as system dynamics (circular causality) and NK (search for the optimal point). The authors custom-designed their simulation by combining several processes including turnover and socialization, and designating sources of stochasticity such as the hiring and firing processes, and their related distributions.

Stochastic processes are particularly applicable when the research question, assumptions or theoretical logic do not fit with a structured approach. Although it is sometimes possible to modify a structured approach to fit the research at hand, stochastic processes is the preferred approach when these modifications are extensive[ix]. Extensive modifications can become unwieldy, yielding a poor computational representation. When using stochastic processes, researchers typically ask how alternative theoretical logics, different assumptions, or varying sources of stochasticity affect system outcomes. They often hold some sources of stochasticity constant while varying others. The result is a flexible approach to theory development, albeit at the price of a lack of standardization and increased need for modeling ingenuity.

Create the Computational Representation

In the previous sections, we described the importance of a theoretically intriguing research question, a simple theory that is likely to be accurate but incomplete, and an appropriate simulation approach (i.e., fits with research question, assumptions, and theoretical logic). In this section, we turn to computational representation of the theory. This activity is central to theory development using simulation. It involves: 1) operationalizing the theoretical constructs, 2) building the algorithms that mirror the theoretical logic of the focal theory, and 3) specifying assumptions that bound the theory and results. Although we describe them separately, researchers usually engage in these activities interactively because constructs, algorithms, and assumptions are highly interdependent. Creating the computational representation is roughly analogous to the activities reported in the methods section of other types of research.

*Operationalizing theoretical construct*s concerns defining the computational measures for each construct (and if not already done, creating its verbal definition). This is roughly analogous to creating empirical measures for theoretical constructs in other types of research. As described below, effective operationalization involves choosing an appropriate computational measure and range of values for each construct. It also involves using construct definitions and names that fit with the extant literature where possible (i.e., inventing new theoretical constructs only when none exists) in order to build reader intuition and confidence, and to enhance the clarity of the theoretical contributions (Repenning, 2003).

One type of computational measure is a *single measure* (termed parameter) that has a range of possible values. For example, Davis and colleagues (2005) operationalized their ambiguity construct by defining a single computational measure with values ranging from 0 (no ambiguity) to 1 (complete ambiguity). For constructs with a range of values that has no natural bounds (e.g., rates ranging from 0 to infinity), it is appropriate to artificially bound the range, and then test these bounds to guard against the possibility that the values outside the bounds produce qualitatively different or contradictory results (Law & Kelton, 1991). If so, the range of values should be adjusted.

A second type of computational measure is a *multi-dimensional measure* (often termed a bit string). Typically, it is effective to measure the most central constructs as bit strings (not parameters) because bit strings have more precision and dimensionality. These features enable more refined and better experimentation. An example of bit string representation is strategy as a set of 10 decisions (Rivkin, 2000). In this type of representation, the dimensions of a construct are measured by specific values (often "0" and "1", occasionally "?" to indicate an absent value). For example, Bruderer and Singh (1996) operationalized organizational form as a bit string of 20 routines with "1" indicating a correct routine, "0" indicating an incorrect one, and "?" indicating an absent routine that could be learned.

A strength of simulation research is *construct validity* (i.e., accurate specification and measurement of constructs) (Cook and Campbell, 1979). Simulation requires precise specification of constructs and their measures, and so avoids "noisy" measurement that affects construct validity in empirical research (Rosenthal and Rosnow, 1991). In addition, as required by its rigorous, step-by-step logic, simulation involves precise specification of units of analysis (e.g., product, strategy) and intervening constructs that are often poorly conceptualized and unmeasured in empirical research. Since simulation eliminates the measurement errors associated with empirical data, convergent and discriminant validity are not germane (Campbell and Fiske, 1959). Finally, simulation also has the advantage of quick, flexible adjustment of construct measures (and even the constructs themselves) by changing the computer code. Such adjustment is usually challenging in empirical research, particularly after the data are collected.

The computational representation also involves *building algorithms* in software that captures the step-by-step theoretical logic that underlies the simple theory. In other words, the software code should embody the theoretical logic. Specifically, the algorithms should consist of a series of steps for modifying construct values in accordance with the underlying theoretical logic of the simple theory. As noted earlier, structured simulation approaches (e.g., NK, genetic algorithms) offer a standardized approach to encoding theoretical logic that is useful when the research fits the simulation approach (Rivkin, 2001). In contrast, an unstructured approach such as stochastic processes offers flexibility when no structured approach fits the research.

Algorithmic representations vary in complexity. Some algorithms consist of two or three steps involving a few constructs (e.g., March, 1991; Rudolph & Repenning, 2002). Others are multi-step involving several or more constructs (e.g., Carroll & Harrison, 1998; Sterman et al., 1997). Algorithmic complexity should depend upon the complexity of the underlying theoretical logic being represented, and relate to the well-known theoretical tradeoff between parsimony and accuracy (Pfeffer, 1982). Like other methods, simulation research attempts to balance parsimony and accuracy by capturing the central logic, while stripping away the non-essential. This balance is ultimately a judgment call. But unlike other methods, it is often effective to err on the side of simplicity in simulation research in order to enhance the intuition and confidence of readers in the simulation results (Repenning, 2003). Complexity (and thus greater accuracy and realism) can then be added through a series of structured experiments.

Finally, computational representation involves *specifying assumptions*. Some of these assumptions relate to boundary or scope conditions of the theory. But other assumptions are simplifications of the simulation itself that enable the researcher to strip out complexity in order to focus on the central logic and constructs. Thus, these assumptions are intimately related the complexity of the algorithmic representation. For example, both Rivkin's (20001) assumption of zero search costs and Adner's (2002) assumption that all markets are the same size both enable less complicated algorithmic representations. As above, the key judgment is the balance between parsimony and accuracy. This is, of course, a familiar judgment in other research methods where choices about sampling in specific contexts (e.g., single industry studies) or to use the controlled environment of the laboratory, in effect, enable a less complicated theoretical logic (Fine & Elsbach, 2000). This judgment is, however, more readily changed in simulation research because of the relative ease of shifting computer code.

An important strength of theory development using simulation methods is *internal validity* (Campbell & Stanley, 1966). Creating a computational representation involves the precise specification of the theoretical logic that is enforced through the discipline of algorithmic representation in software (Abelson et al., 1996). Coupled with the precise specification of constructs, measures, and assumptions that is also enforced by the software, the discipline of algorithmic representation sharpens loose theoretical arguments about the definition of constructs, relationships among constructs, and underlying logic (Carroll & Harrison, 1998; Sastry, 1997). The result is theory that is more likely to exhibit strong internal validity.

Verify the Computational Representation

A critical step in developing theory using simulation methods is verification of the computational representation. Roughly analogous to manipulation checks in laboratory experiments and examination of correlation matrices in multi-variate analysis, verification helps to ensure that the computational representation accurately embodies the theoretical logic, the theory is internally valid, and the simulation results can be interpreted with confidence.

Researchers should verify their computational representation in several ways. The most important is *comparing simulation results* with the (implicit or explicit) propositions of the simple theory. If the simulation confirms the propositions, then the theoretical logic and its computational representation are likely to be correct hypothesis[x]. Rivkin, for example, (2001) verified his computational representation by confirming his proposition that the advantage of a firm's replicating its own strategy over another firm's imitating the same strategy was greatest at moderate levels of strategic complexity. He did so by running the simulation at various values of strategic complexity and then comparing the simulation results with his inverted U-shaped prediction. In simulations where the simple theory centers on several well-known processes, each process should be verified.

Simulation researchers should further verify their computational representation with *robustness checks* (sometimes termed sensitivity analysis[xi]) that increase confidence that the computational representation is stable. For example, Zott (2003) used alternative starting values of the decision variables to confirm that his computational representation was robust to alternative initial conditions. Davis and colleagues (2005) used two alternative operationalizations of their strategic complexity construct to verify that their simulation results were not dependent upon a specific representation of their central construct. Researchers should also use other techniques to verify that their software coding is correct such as tracking variables at intermediate steps in simulation, running the simulation with extreme values of constructs, and other techniques, including those specific to particular simulation approaches (e.g., pulse tests in system dynamics). These tests often have no theoretical implications, but rather are a means to verify the accuracy of the software code.

When researchers find a mis-match between the propositions of their simple theory and the simulation results, it is often due to software coding errors that are subsequently fixed. But occasionally the mis-match reveals shortcomings in the theoretical logic. In these situations, researchers often have an opportunity to develop fresh theoretical insights. For example, Davis and colleagues (2005) unexpectedly observed that the anticipated inverted U-shaped relationship between strategic complexity and performance was asymmetric (i.e., steep on one side and skewed on the other). This led the authors to alter their simple theory to account for this skew. While Davis and colleagues (2005) modestly improved their initial theory, Sastry (1997) made a significant improvement. In her attempt to verify her computational representation of punctuated equilibrium theory, she found significant errors in the theoretical logic. Sastry returned to the original theory and discovered that Tushman and Romanelli (1985) had failed to account for some aspects of reorientation. This unexpected logical flaw led Sastry to reformulate the theory by adding assumptions and a negative feedback process.

Overall, the key point of verification is to ensure that the computational representation accurately represents the underlying theoretical logic. Thus, given that the researcher attempts to encode the theoretical logic in the software, the simulation results should replicate the simple theory, bolster internal validity, and thereby increase confidence in the results of the simulation. [xii]

Experiment to Build Novel Theory

Experimentation is at the heart of the value of simulation methods for developing theory. While verification of the computational representation is necessary and useful for demonstrating the accuracy of the computational representation, establishing internal validity, and enhancing reader confidence, it involves confirmation of theory that is already known. In contrast, effective experimentation builds new theory by revealing fresh theoretical relationships and novel theoretical logic. Moreover, experimentation is a particular strength of simulation methods. Empirical experimentation is constrained by data limitations and formal modeling experimentation is constrained by mathematical tractability. In contrast, simulation methods enable experimentation across a wide range of conditions simply by changing software code (Bruderer & Singh, 1996; Zott, 2003).

There are several approaches to effective experimentation. A common one is *varying the value* of constructs that were held constant in the initial simple theory. This approach is particularly useful for uncovering moderating and interaction relationships, new main effects, and intervening constructs. Typically, researchers experiment with a wide range of values of a previously fixed construct in order to explore fully the effects of the construct on outcomes, but then only report the most intriguing results. For example, after verifying their initial proposition that experiential learning and cognition would lead to a higher performance than experiential learning alone, Gavetti and Levinthal (2000) varied the coupling among policy elements that was previously held constant. This enabled the authors to investigate how coupling moderated the relationship between cognition and performance, and to extend their simple theory with this interaction effect. Similarly, Rudolph and Repenning (2002) examined the effects of minor interruptions on the emergence of major organizational catastrophes. After verifying the anticipated tipping points, the authors experimented by varying the previously constant variance in the interruption rate. This enabled the authors to extend their original simple theory to include the interaction effects of interruption variance on performance.

A second approach to experimentation is *unpacking* key theoretical constructs. By unpacking we mean breaking a single construct into constituent component constructs. Unpacking a construct is particularly useful when the original construct is imprecise, abstract, or multi-dimensional such that the different dimensions may have distinct effects. Experiments then focus on uncovering these possible effects. For example, after verifying their simple theory, Davis and colleagues (2005) unpacked their market dynamism construct into four granular constructs (i.e., velocity, complexity, ambiguity, and unpredictability) that captured distinct dimensions of market dynamism. They then experimented with each construct by running the simulation with three constructs held constant and varying the fourth. This enabled the authors to extend their theory to the performance implications of different kinds of market dynamism, and to isolate the specific implications of each construct for their theory.

A third approach to experimentation is *varying assumptions*. This approach is particularly useful when fundamentally different processes may reasonably exist. Experiments focus on revealing their possible distinct effects. For example, Rivkin (2000) verified his theory relating the speed with which executives find high-performing strategies and the coupling among elements of strategy. In doing so, he assumed an "incremental" search process in which executives change their current strategy only if altering specific decisions within that strategy leads to increased performance. He then experimented with two alternative search processes: "follow-the-leader" and "hybrid". These experiments enabled him to elaborate his original theory to include the implications of alternative search processes.

A fourth approach to experimentation is *adding new features* to the computational representation. By adding this complexity in successive computational representations, researchers can build theoretical understanding of the phenomenon in a structured way that enhances the understanding of both researchers and readers. Additional complexity (e.g., processes and constructs) is particularly useful when the researcher wants to explore the interaction of multiple processes that are well-known alone but not in combination, and more broadly, when greater realism is desired. For example, Repenning's (2002) simple theory related persistence of managerial commitment to the success of new innovations. After verifying this theory with a single organizational group, he experimented by adding an additional feature (i.e., second organizational group) that made the simulation more complex but also more realistic. In revealing that a second organizational group is unlikely to be successful in adopting an innovation, the author extended his simple theory to include multiple groups, thereby enhancing its theoretical relevance for real organizations.

Several criteria should shape experimentation design. The most important criterion is theoretical contribution. Understanding where theoretical contributions are likely to be usually stems from knowledge of the literature. Although serendipitous discoveries can occur, knowing the literature often enables the researcher to know where theoretical discrepancies are and so where to experiment to improve the likelihood of intriguing theoretical insights. Such insights are reason d'etre of the research. A related criterion for experimentation design is realism. As noted earlier, it is often effective to begin with a simple computational representation in order to build reader intuition and confidence in the simulation. Experimentation in the form of added features and other experimental approaches can then successively add realism that enhances the generalizeability of the resulting theory.

Experimentation is closely associated with building theory using "disciplined imagination" (Weick, 1989). That is, researchers using simulation can readily engage in an evolutionary process of speculative experiments to create alternative versions of theory (imagination) and then selection of the best among them (discipline). Although researchers can engage in disciplined imagination using other methods such as systematic thought experiments, these methods often fail when the theory involves intertwined and longitudinal processes, timing effects, and non-linearities that are the province of simulation. In effect, the simple theory of a simulation becomes a platform for elaborating and extending theory through creative and systematic experimentation (Weick, 1989). The result may be fundamental (i.e., widely applicable and non-obvious) theory (Lave & March, 1975).

Overall, experimentation is at the heart of the value of using simulation methods for theory development. While verification confirms existing theory and bolsters internal validity, experimentation adds the creative and even surprising theoretical contributions that build new theory. Sometimes this new theory is an incremental advance. But sometimes this theory is a fundamental improvement in which frame-breaking insights generate qualitatively better theory.

Validate with Empirical Data

A final step in theory development using simulation methods is validation. Validation involves comparison of simulation results with empirical data. If the results of the simulation match the empirical evidence, then the simulation is validated for that empirical context. This strengthens the *external validity* of the theory (i.e., generalizeability and predictability of the theory) (Campbell & Stanley, 1966).

There is, however, some debate over the value of validation. For example, some theorists argue that the central task of theory development is creating interesting theory and so diminish the importance of validation (e.g., Van Maanen, 1995; Weick, 1989). Others disagree. Our view is contingent – i.e., the importance of validation should depend on the source of the simple theory that is the basis of the simulation. If this theory is based primarily on empirical evidence (e.g., field-based case studies and empirically grounded processes), then validation is less important because the theory already has some external validity. In contrast, if the theory is based primarily on non-empirical argument (e.g., formal analytic modeling) or on evidence from distant scientific disciplines (e.g., physics), then validation is more important.

There are several approaches to effective validation. One is to compare the results of the simulation to statistical results derived from large-scale empirical data. This approach enables a broadbrush validation of the simulation results. A standard technique in economics, for instance, is to use theories developed with simulation to “predict” the results of some well-known dataset. For example, Nelson and Winter (1982) created a simulation of their theory of evolutionary economic change, and then used that simulation to reproduce historical productivity data. Another approach is to compare the simulation results and theoretical logic with case study data. This approach seeks to demonstrate that the simulation is consistent with the specific details of one or a few examples. This enables granular validation. For example, Sterman and colleagues (1997) used interview data from one organization to show that their theory of organizational change was plausible in at least this one organization. The choice between these two approaches usually depends upon data availability.

DISCUSSION

While simulation is an increasingly significant methodological approach in the organizations and strategy literatures (e.g., Lant & Mezias, 1990; Rivkin & Siggelkow, 2001; Repenning, 2002; Zott, 2003), its link to theory development remains unclear and even controversial. Our purpose is to clarify when and how to use simulation methods to develop theory.

Our first contribution is a *roadmap* for conducting effective theory development using simulation. Like all effective research, developing theory from simulation methods depends upon beginning with an intriguing research question that relates to fundamental theoretical issues. Simulation is especially effective when that research question relates to competing tensions (e.g., long vs. short run, structure vs. chaos) and intertwined processes (e.g., inertia and change, competition and legitimation) that are especially well-addressed by simulation. The central activity is developing an accurate computational representation of simple (i.e., undeveloped) theory through appropriate selection of a simulation approach (e.g., genetic algorithm v. stochastic processes), construct operationalization, and algorithmic representation The benefits of simulation emerge in verifying the computational representation which strengthens the internal validity of the simple theory, and more significant, in creatively experimenting to produce novel theoretical insights. Indeed, creative experimentation is at the heart of the value of simulation methods for theory development.

Our second contribution is the positioning of simulation methods within the *broad context* of theoretical development in the organizations and strategy literatures. Simulation is particularly useful for developing theory in the "sweet spot" between theory creating using methods such as multiple case inductive research (Eisenhardt, 1989) and formal modeling (Freese, 1980), and theory testing using empirical evidence and multi-variate statistical techniques (Lattin, 2003; Pfeffer, 1993). On the one hand, theory-creating methods can reveal simple theory, but are limited in their ability to elaborate that theory. The inductive case method is constrained by limited data and formal modeling is constrained by mathematical tractability. Simulation can mitigate these weaknesses by exploring, elaborating and extending simple theory that is produced by these theory-creating methods. On the other hand, as Sutton and Staw (1995) argue, many theory-testing studies lack conceptual precision and logical theoretical argument. Simulation can mitigate these weaknesses in internal validity by tightening the rigor of the underlying theoretical logic, sharpening constructs, and elaborating theoretical propositions prior to empirical test.

Strengths and Weaknesses

These observations suggest several important *strengths* for developing theory with simulation. One is internal validity (Cook and Campbell, 1979. The computational rigor of simulation forces precise specification of constructs, assumptions, and theoretical logic that creates strong internal validity. This emphasis on internal validity is particularly valuable because it addresses a common weakness of empirical research: often poor theoretical logic regarding "why" propositions might be true (Sutton & Staw, 1995; Whetten, 1989). This emphasis also mitigates another common weakness of empirical research: weak specification of boundary conditions. By requiring precise specification of assumptions, simulation typically bounds the scope of the theory and so clarifies boundary conditions. Thus, while other methods emphasize constructs and propositions, simulation puts their underlying theoretical logic and assumptions on center stage.

Another strength is experimentation. Simulation creates a computational laboratory in which researchers can systematically experiment (e.g., unpack constructs, relax assumptions, vary construct values, add new features) in a controlled setting to produce new theoretical insights. This experimentation is particularly valuable when the theory seeks to explain longitudinal and processual phenomena that are challenging to study using empirical methods because of their time and data demands. Simulation is also well-suited to theory development related to non-linear phenomena such as tipping points, feedback loops, thresholds and catastrophes, and asymmetries. These are difficult to uncover using inductive case methods and to examine using standard statistical techniques. Yet significantly, these phenomena are becoming central as theory development moves from cross-sectional and equilibrium perspectives to longitudinal and dynamic ones.

Yet like all methods, theory development using simulation has *weaknesses.* A primary one is external validity. Simulation eliminates complexity in order to focus on the core aspects of phenomena and so uses computational representations that are often stark such as representing an organization by a 0/1 bit string (Bruderer and Singh, 1996) or making clearly false assumptions such as no disadvantages to being a second mover (Rivkin, 2000). The result can be an overly simplistic and distant model that fails to capture critical aspects of reality.

Towards Better Theory Development using Simulation Research

Like all research, theory development using simulation methods should be evaluated according to two fundamental criteria: theoretical contribution and strength of method. *Theoretical contribution* is partially determined by the quality of the research question and its related grounding in the literature. That is, the research question should center on a significant issue within the related literature. In contrast, research with little grounding in the relevant research often has a research question that does not address a useful theoretical gap, or has theoretical constructs and theory that do not fit with their conceptualization and terminology within the literature. This lack of grounding can further lead to weak integration of the simulation results within the literature. While awareness of and grounding in the extant literature seems obvious, our experience indicates that simulation researchers are more likely to misunderstand their research context (i.e., choose poor research questions, fail to understand and use well-known concepts) than others or simply pass over the contributions of others working with different methods. Yet without situating the research question and theory in the current literature and then relating results to that literature, it is difficult to have a theoretical contribution. Thus, the reader should ask whether the research is situated within the relevant research literature and addresses a substantive research question.

Theoretical contribution is also determined by the quality of the experimentation. As noted earlier, experimentation is a key strength of simulation and a primary source of theoretical insight. That is, systematic experimentation focused on unexplored theoretical constructs (e.g., varying construct values, unpacking constructs into more precise constructs), major assumptions (e.g., relaxing them, varying them), and important elaborations (e.g., adding new features) is integral to theoretical contribution. Therefore, when evaluating theory development using simulation, the reader should focus on the experimentation to assess whether and to what extant theoretical contribution has been made. Simulation authors should likewise distinguish between the verification of the simple theory that confirms the computational representation and the internal validity, and the theoretical insights that emerge from systematic experimentation (or occasionally from failed verification) that creatively elaborates and extends the theory. Although this may seem obvious, simulation researchers have a tendency to communicate their findings as if simply simulation of a phenomenon per se is a theoretical contribution. That is, building a simulation is confused with theoretical contribution. Some also stop at verification (e.g., various robustness checks), rather than continuing to genuine experimentation. But, research that only verifies known theory or jumbles verification and experimentation typically fails to create theoretical contribution. Indeed, like all research, the theoretical contribution of simulation research is independent of method, and relates to the addition of new theoretical insight into the literature. Thus, the reader should ask whether the results constitute a theoretical contribution.

The second criterion is the *strength of method*. In the context of the roadmap outlined earlier, a high quality method includes justification for using simulation for the research question at hand and using a simulation approach (e.g., cellular automata, stochastic processes) that fits the research. For example, as is common practice in empirical research, simulation researchers should clearly justify the simulation approach (as one would justify a statistical approach), define constructs, and indicate the computational measures (as one would describe empirical measures). The methods section of other types of research provides a useful template for conveying this information.

In addition, high quality simulation methods should have an accurate computational representation of the theoretical logic that is conveyed to the reader in a way that carefully builds understanding of assumptions and the logical flow of the simulation. Given the opacity of computer coding, this puts a premium on the effective presentation of the theoretical logic. In addition to logical argument (roughly analogous to the theoretical development of hypotheses in empirical research), pictures and flowcharts are helpful in this regard (see Rudolph and Repenning, 2002 and Zott, 2003 for exemplars). High quality simulation methods also include explicit verification of the computational representation that confirms the accuracy and internal validity of the computational representation, and builds reader confidence in the simulation. It is helpful to state explicitly the propositions of the simple theoretical ideas that form the basis of the simulation, and then to verify them clearly by running the simulation. In contrast, when the computational representation is poor, poorly described, or unverified, the simulation results are unreliable and/or not believed.

Finally, a high quality method involves the appropriate statistical design of the verification and experimentation simulation runs. This design should include an appropriate number of time steps in each simulation run and number of simulation runs per experiment (related to the statistical power).[xiii] As in empirical research, high quality simulation methods also include statistical confidence intervals (Law & Kelton, 1991). In contrast, when research that does not justify the basic statistical properties of the simulation (e.g., number of time steps, number of runs per experiment) or provide basic statistical analysis (i.e., confidence intervals), confidence in the statistical results is compromised.

Overall, readers should expect that high quality theory development using simulation methods meets the usual standards of strong theory such as parsimony, internal validity, accuracy, and interest (Davis, 1971; Pfeffer, 1982; Eisenhardt, 1989; Priem & Butler, 2001). Or, more simply, the resulting should "explain, predict and delight" (Weick, 1989). Given the starting point of simple theory, the result of effective simulation can be fundamental theory that is parsimonious, internally consistent, and applicable to a wide-range of situations (Pfeffer, 1982).

CONCLUSION

We began by observing that, while simulation is a significant method for theory development, it usefulness is unclear and even controversial. Our purpose is to clarify how and when to use simulation methods for theory development. We offer two contributions. One is a roadmap for developing theory through simulation (Table 1) including the pivotal role of experimentation in creating theoretical contribution. The second is positioning simulation methods in the "sweetspot" between theory creating using inductive case methods and formal modeling, and theory testing using multivariate statistical techniques. We conclude by observing that simulation is likely to become an increasingly prominent method for theory development. As organizations and strategy scholars move to an emphasis on theory explaining dynamic and longitudinal phenomena, simulation methods will become a natural choice.

REFERENCES

Abelson, H., Sussman, G. J., & Sussman, J. 1996. ***Structure and Interpretation of Computer Programs***. Cambridge, MA: MIT Press.

Adner, R., & Levinthal, D. 2001. Demand Heterogeneity and Technology Evolution: Implications for Product and Process Innovation. ***Management Science***, 47(5): 611-628.

Afuah, A. 1998. Competitive Advantage from Intellectual Capital: The Case of Cholesterol Ethical Drugs: ICOS Seminar.

Aldrich, H. 1999. ***Organizations Evolving***. Thousand Oaks, CA: Sage Publications.

Bruderer, E., & Singh, J. S. 1996. Organizational Evolution, Learning, and Selection: A Genetic-Algorithm-Based Model. ***Academy of Management Journal***, 19(5): 1322-1349.

Campbell, D. T., & Stanley, J. C. 1966. ***Experimental and quasi-experimental designts for research***. Chicago: Rand McNally.

Carley, K. M. 2001. Computational Approaches to Sociological Theorizing. In J. Turner (Ed.), ***Handbook of Sociological Theory***: 69-84. New York, NY: Kluwer Academic/Plenum Publishers.

Carroll, G., & Harrison, J. R. 1998. Organizational Demography and Culture: Insights from a Formal Model and Simulation. ***Administrative Science Quarterly***, 43: 637-667.

Chattoe, E. 1998. Just how (un)realistic are evolutionary algorithms as representations of social processes? ***Journal of Artificial Social Science Simulation***, 1(3): 2.1-2.36.

Cohen, M. D., March, J., & Olsen, J. P. 1972. A Garbage Can Model of Organizational Choice. ***Administrative Science Quarterly***, 17(1): 1-25.

Cook, T. D., & Campbell, D. T. 1979. ***Quasi-Experimentation: Design and Analysis Issues for Field Settings***. Boston: Houghton Mifflin Company.

Davis, J., Eisenhardt, K. & Bingham, C. 2005. Complexity Theory, Market Dynamism, and the Strategy of Simple Rules. ***Stanford Technology Ventures Program -- Working Paper***.

Davis, M. S. 1971. That's Interesting! Towards a Phenomenology of Sociology and a Sociology of Phenomenology. ***Phil. Soc. Sci.***, 1: 309-344.

Dubin, R. 1976. Theory Building in Applied Areas. In M. Dunnette (Ed.), ***Handbook of Industrial and Organizational Psychology***: 17-40. Chicago, IL: Rand McNally.

Eisenhardt, K. M. 1989. Building Theories from Case Study Research. ***Academy of Management Review***, 14: 532-550.

Fichman, M. 1999. Variance Explained: Why Size Doesn't (Always) Matter. ***Research in Organizational Behavior***, 21: 295-331.

Fine, G. A., & Elsbach, K. D. 2000. Ethnography and Experiment in Social Psychological Theory Building. ***Journal of Experimental Social Psychology***, 36: 51-76.

Forrester, J. 1961. ***Industrial Dynamics***. Cambridge, MA: MIT Press.

Freese, L. 1980. Formal theorizing. ***Annual Review of Sociology***, 6: 187-212.

Gallager, R. 1996. ***Discrete Stochastic Processes***. Boston, MA: Kluwer Academic Publishers.

Gavetti, G., & Levinthal, D. 2000. Looking Forward and Looking Backward: Cognitive and Experiential Search. ***Administrative Science Quarterly***, 45: 113-137.

Glaser, B., & Strauss, A. L. 1967. ***The Discovery of Grounded Theory: Strategies for Qualitative Research***. London: Wiedenfeld and Nicholson.

Goldberg, D. E. 1989. ***Genetic algorithms: In search of optimization and machine learning***. Reading, MA: Addison-Wesley.

Harrison, J. R., & Carroll, G. R. 1991. Keeping the faith: A model of cultural transmission in formal organizations. ***Administrative Science Quarterly***, 36: 552-582.

Holland, J. H. 1975. ***Adaptation in natural and artificial systems***. Ann Arbor, MI: The University of Michigan Press.

Kauffman, S. 1989. Adaptation on rugged fitness landscapes. In E. Stein (Ed.), ***Lectures in the Science of Complexity***. Reading, Mass.: Addison-Wesley.

Kauffman, S. 1993. ***The Origins of Order***. New York, NY: Oxford University Press.

Kreps, D. M. 1990. Corporate culture and economic theory. In J. E. Alt, & K. A. Shepsle (Eds.), ***Perspectives on positive political economy***: 90-143. Cambridge [England] ; New York: Cambridge University Press.

Langton, C. G. 1984. Self-Reproduction in Cellular Automata. ***Physica***, 10D: 134-144.

Lant, T., & Mezias, S. 1990. Managing Discontinuous Change: A Simulation Study of Organizational Learning and Entrepreneurship. ***Strategic Management Journal***, 11: 147-179.

Lant, T., & Mezias, S. 1992. An Organizational Learning Model of Convergence and Reorientation. ***Organization Science***, 3(1): 47-71.

Lattin, J. 2003. ***Analyzing Multivariate Data***. Toronto: Brooks/Cole, Thomson Learning, Inc.

Lave, C., & March, J. G. 1975. ***An Introduction to Models in the Social Sciences***. New York, NY: Harper and Row.

Law, A. M., & Kelton, D. W. 1991. ***Simulation Modeling and Analysis*** (2nd ed.). New York, NY: McGraw-Hill.

Lee, T., Mitchell, T. and C. Sablynski. Qualitative Research in Organizational and Vocational Psychology. ***Journal of Vocational Behavior***, 55: 161-187.

Levinthal, D. 1997. Adaptation on Rugged Landscapes. ***Management Science***, 43: 934-950.

Lomi, A., & Larsen, E. 1996. Interacting Locally and Evolving Globally: A Comutational Approach to The Dynamics of Organizational Populations. ***Academy of Management Journal***, 39(4): 1287-1321.

March, J. G. 1991. Exploration and Exploitation in Organizational Learning. ***Organization Science***, 2(1): 71-87.

Nelson, R. R., & Winter, S. G. 1982. ***An Evolutionary Theory of Economic Change***. Cambridge, Massachusetts: Belknap - Harvard University Press.

Pfeffer, J. 1982. ***Organizations and organization theory***. Boston: Pitman.

Pfeffer, J. 1993. Barriers to the Advance of Organizational Science: Paradigm development as a Dependent Variable. ***Academy of Management Review***, 18(4): 599-620.

Priem, R. L., & Butler, J. E. 2001. Is the Resource-based "View" a Useful Perspective for Strategic Management Research? ***Academy of Management Review***, 26(1): 22-41.

Repenning, N. 2002. A Simulation-Based Approach to Understanding the Dynamics of Innovation Implementation. ***Organization Science***, 13(2): 109-127.

Repenning, N. 2003. Selling system dynamics to (other) social scientists. ***System Dynamics Review***, 19(4): 303-327.

Rivkin, J., W. 2000. Imitation of Complex Strategies. ***Management Science***, 46(6): 824-844.

Rivkin, J., W. 2001. Reproducing Knowledge: Replication Without Imitation at Moderate Complexity. ***Organization Science***, 12(3): 274-293.

Rosenthal, R., & Rosenow, R. L. 1991. ***Essentials of Behavioral Research: Methods and Data Analysis*** (2nd ed.). New York: McGraw-Hill.

Rudolph, J., & Repenning, N. 2002. Disaster Dynamics: Understanding the Role of Quantity in Organizational Collapse. ***Administrative Science Quarterly***, 47: 1-30.

Sastry, M. A. 1997. Problems and paradoxes in a model of punctuated organizational change. ***Administrative Science Quarterly***, 42(2): 237-275.

Schelling, T. 1971. Dynamic models of segregation. ***Journal of Mathematical Sociology***, 1: 143-186.

Sterman, J. 2000. ***Business Dynamics: Systems Thinking and Modeling for a Complex World***. New York, NY: Irwin McGraw-Hill.

Sterman, J., Repenning, N., & Kofman, F. 1997. Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement. ***Management Science***, 43(4): 503-521.

Stinchcombe, A. 1968. ***Constructing Social Theories***. Chicago, IL: University of Chicago Press.

Sutton, R. I., & Staw, B. M. 1995. What theory is not. ***Administrative Science Quarterly***, 40(3): 371-384.

Tushman, M., & Romanelli, E. 1985. Organizational Evolution: A metamorphosis model of convergence and reorientation. In L. L. Cummings, & B. M. Staw (Eds.), ***Research in Organizational Behavior***, Vol. 7: 171-222. Greenwich, CT: JAI Press.

Van Maanen, J. 1995. Style as theory. ***Organization Science***, 6(1): 133-143.

Weick, K. E. 1989. Theory Construction as Disciplined Imagination. ***Academy of Management Review***, 14(4): 516-531.

Weick, K. E. 1993. The vulnerable system: An analysis of the Tenerife air disaster. In K. H. Roberts (Ed.), ***New Challenges to Understanding Organizations***: 173-198. New York, NY: Macmillan.

Whetten, D. 1989. What constitutes a theoretical contribution? ***Academy of Management Review***, 14: 490-495.

Wolfram, S. 2002. ***A New Kind of Science***. Champaign, IL: Wolfram Media.

Wright, S. 1931. Evolution in Mendelian populations. ***Genetics***, 16: 97.

Yerkes, R. M., & Dodson, J. D. 1908. The relation of strength of stimulus to rapdity of habit formation. ***Journal of Comparative Neurological Psychology***, 18: 459-482.

Yin, R. K. 1994. ***Case Study Research: Design and Methods*** (Second ed.). Thousand Oaks: Sage Publications.

Zott, C. 2002. When adaptation fails: An agent-based explanation of inefficient bargaining under private information. ***Journal of Conflict Resolution***, 46: 727-753.

Zott, C. 2003. Dynamic Capabilities and the Emergence of Intra-industry Differential Firm Performance: Insights from a Simulation Study. ***Strategic Management Journal***, 24: 97-125.

TABLE 1
Roadmap for Developing Theory Using Simulation Methods

Step	Activities	Rationale
Begin with a research question	• Determine a theoretically intriguing research question • Look for a basic tension like structure v. chaos, long v. short run	• Focuses efforts on a theoretically relevant issue for which simulation is especially effective
Identify simple theory	• Select simple theory that addresses the research question • Look for intertwined processes (e.g., competition and legitimation), non-linearities, and longitudinal effects • Look for theory that requires data that are challenging to obtain	• Forms basis of computational representation by giving shape to theoretical logic, propositions, constructs and assumptions • Focuses efforts on theoretical development for which simulation is especially effective
Choose a simulation approach	• Choose simulation approach that fits with research question, assumptions, and theoretical logic • If the research does not fit an approach or if the approach requires extensive modification, choose stochastic processes	• Ensures that the research uses an appropriate simulation approach given the research at hand
Create computational representation	• Operationalize theoretical constructs • Build computational algorithm that mirrors theoretical logic • Specify assumptions • Ensure that computational representation allows theoretically valuable experimentation	• Embodies theory in software • Provides construct validity • Improves internal validity by requiring precise constructs, logic, and assumptions • Sets the stage for theoretical contributions
Verify computational representation	• Replicate propositions of simple theory with simulation results • Conduct robustness checks of computational representation • If verification fails, correct theory and/or software coding	• Confirms accuracy and robustness of computational representation • Confirms internal validity of the theory
Experiment to build novel theory	• Create experimental design (e.g., vary construct values, unpack constructs, alter assumptions, add new features) based on likely theoretical contribution and realism	• Focuses experimentation on theory development • Builds new theory through exploration, elaboration, and extension of simple theory
Validate with empirical data	• Compare simulation results with empirical data	• Strengthens external validity of the theory

TABLE 2
Recent Examples of Simulation Research

Study	Research Question	Key Processes	Approach	Representative Findings
Rudolph, J., and Repenning, N. (2002)	When do small interruptions create major catastrophes?	Adaptation and selection	System dynamics	Variance in the rate of interruptions affects emergence of tipping points
Sastry (1997)	How do organizations undergo fundamental change?	Change and inertia	System dynamics	An additional negative feedback loop corrects theoretical logic of punctuated equilibrium theory
Rivkin, J. (2002)	What is the optimal strategic complexity?	Replication and imitation	NK Fitness Landscape	Moderate (K) strategic complexity is optimal
Gavetti and Levinthal, (2000)	How does cognition improve experiential learning?	Experiential learning and cognition	NK Fitness Landcape	Cognition is most useful in improving experiential learning at moderate levels of K interactions
Zott, C. (2002)	How does adaptive learning occur within bargaining?	Adaptive learning	Genetic Algorithm	Adaptive failures may occur even with complete information
Bruderer, E., and Singh, J. (1996)	How does organizational learning affect the evolution of a population of organizations?	Variation, adaptation and selection	Genetic Algorithm	Environment influences when organizational learning is most useful for the population's convergence to an optimal form
Lomi, A., and Larsen, E. (1996)	How do competition and legitimation affect density dependence?	Competition and legitimation	Cellular Automata	Neighborhood size moderates the relationship of density with founding and failure rates
March, J. (1991)	What is the relationship between exploration and exploitation?	Exploitation and exploration	Stochastic Processes	Exploitive processes are effective in the short run, but destructive in the long-run
Davis, J., Eisenhardt, K. and Bingham, C. (2005)	What is the optimal degree of structure?	Improvisation	Stochastic Processes	Unpredictability (not ambiguity, complexity or velocity) is the driver of the tension between structure and chaos

TABLE 3
Comparison of Simulation Approaches

Simulation Approach	Focus	Common Research Question(s)	Key Assumptions	Theoretical Logic	Common Experiments
Systems Dynamics • Sterman et al, 1997, Sastry, 1997; Repenning, 2002; Rudolph &Repenning, 2002	• Behavior of a system with complex causality and timing	• What conditions create system instability?	• System of intersecting, circular causal loops • Stocks that accumulate and dissipate over time • Flows that specify rates within system	• Description • Inputs to a system of interconnected causal loops, stocks and flows produces system outcomes	• Add causal loops • Change mean of flow rates • Change variance of flow rates
NK Fitness Landscapes • Levinathal, 1997; Rivkin, 2000; Gavetti & Levinthal, 2000; Rivkin & Siggelkow, 2001	• Speed and effectiveness of adaptation of modular systems with tight v. loose coupling to an optimal point	• How long does it take to find an optimal point (e.g. high performing strategy)? • What is the performance of the optimal point?	• System of N nodes, K coupling btwn nodes • Fitness landscape that maps performance of all combinations • Adaptation via incremental moves and long jumps	• Optimization • Adaptation of a modular system using search strategies (i.e., long jumps, incremental moves) to find an optimal point on a fitness landscape	• Vary N and K • Change adaptation moves • Add a “map” of the landscape • Create an environmental jolt
Genetic Algorithms • Bruderer & Singh, 1996; Zott, 2002	• Adaptation of a population of agents (e.g., organizations) via simple learning to an optimal agent form	• What affects the rate of adaptation (or learning or change)? • When and/or does an optimal form emerge?	• Population of agents with genes • Evolutionary adaptation (v-s-r) • Variation via mutation (mistakes) and crossover (recombinations) • Selection via fitness (performance) • Retention via copying selected agents	• Optimization • Adaptation of a population of agents using an evolutionary process toward an optimal agent form	• Vary mutation probability • Vary crossover probability • Vary length of time of evolution • Create an environmental jolt
Cellular Automata • Lomi & Larsen, 1996	• Emergence of macro patterns from micro interactions via spatial processes (e.g., competition, diffusion) in a population of agents	• How does the pattern emerge and change? • How fast does a pattern emerge?	• Population of spatially arrayed and semi-intelligent agents • Agents use rules (local & global) for interaction, some based on spatial processes • Neighborhood of agents where local rules apply	• Description • Interactions among agents following rules produces macro-level patterns	• Change the rules • Change the neighborhood size
Stochastic Processes • March, 1991; Carroll & Harrison, 1998; Zott 2003; Davis, Eisenhardt & Bingham, 2005	• Flexible approach to a wide variety of research questions, assumptions, and theoretical logics	• No specific research questions beyond asking what are the effects of varying the stochastic sources	• One or more processes by which system operates • One or more stochastic sources (e.g., process elements) • Probabilistic distributions for each stochastic source	• No specific theoretical logic	• Change stochastic sources • Vary levels of stochasticity • Unpack constructs • Change pieces of theoretical logic

ENDNOTES

[i] Scholars writing about theory development sometimes emphasize different elements of a theory. Dubin (1976) and Whetten (1989) explicitly posit these four elements as comprising theory. Weick's (1989) definition of theory draws upon Sutherland's (1975, p. 9) definition of theory as "an ordered set of assertions about a generic behavior or structure" that indicates propositions and theoretical logic. Constructs are implicit and assumptions arise later in Weick's (1989) article. Staw and Sutton (1995) focus on what theory is not. But in so doing they make the case for the necessity of theoretical logic that coherently ties together constructs and propositions, and that answers "why?". Priem and Butler (2001, p. 26) use different language in describing theoretical statements as being generalized conditionals, having empirical content, and exhibiting nomic necessity. These elements relate to a theory's having non-tautological, if/then propositions that relate constructs together and an underlying theoretical logic. Nonetheless despite their different emphases and language, many scholars have converged on many or all of the four elements that we indicate and the related definition of theory that we use.

[ii] There are a number of uses of simulation including description of complicated systems such as bridges and telecommunication networks, and purely exploratory efforts such as observing what emerges from simulating complex relationships. In contrast, our focus is specifically on using simulation for theory development when some simple theory exists. Its application ranges from incremental extensions of that theory (i.e. exploitive) to exploratory work that may rely on only a rough understanding of theoretical mechanisms, may be derived from thought processes, or that pushes the theory "beyond its local neighborhood". This range does not, however, preclude other important uses of simulation. We appreciate the comments of one of our anonymous reviewers in clarifying this scope condition of our efforts and suggesting language.

[iii] We appreciate the comments of one of our anonymous reviewers in assisting us in highlighting this continuum and language.

iv Other researchers (e.g., Dooley, 2002) have used related simulation typologies. We chose our typology because it is a fine-grained mapping that accurately relates to the major approaches that are used in the organizations and strategy literatures. That is, it reflects the most frequently used categorization in the relevant extant research. Given the scope of our paper, our description of these approaches is necessarily limited. Interested readers should turn to technical treatments if they wish to use these approaches.

[v] NK fitness landscape is an appropriate simulation choice in the case of Gavetti and Levinthal (2000) because their research is centered on the effects of coupling on the rate of adaptation of a modular system (i.e., policy with coupled elements) to an optimum, and so fits the approach. In contrast, the closest alternative choice probably is genetic algorithms. But, although genetic algorithms is also an optimization approach involving adaptive learning, it is focused on the evolution of a population to an optimal form and more importantly, does not directly address either coupling or the kinds of deliberate adaptive moves such as long jumps that fit with the authors' interest in cognition. Finally, the authors could have used stochastic processes. But this would have involved time-consuming duplication of algorithmic development. We appreciate the comments of an anonymous reviewer in suggesting a description of the reasoning behind this choice.

[vi] Although there has been only limited use of genetic algorithms in the organizations and strategy literature that is our focus, there is an emerging tradition in the economics literature. Applications include economic growth (e.g., Arifovec et al., 1997), auctions (e.g., Andreoni and Miller, 1995), and forecasting (e.g., Bullard and Duffy, 1998). This work is often framed in juxtaposition with highly rational models of agent behavior such as game theoretic formulations. Readers with particular interest in genetic algorithms should consult these sources as well as more general sources (e.g., Holland, 1975; Goldberg, 1989).

[vii] Genetic algorithms and cellular automata both focus on the evolution of populations. But they are distinct simulation approaches. Genetic algorithms is an optimization approach that models evolutionary

learning (variation-selection-retention) in a population of agents composed of genes. In contrast, cellular automata is a descriptive approach that models agents arrayed in space relative to one another, and in which change occurs according to agent rules that are based (at least partially) on spatial processes like competition and diffusion. In the case of Lomi and Larsen (1996), their research question, theoretical logic, and assumptions about the phenomena make cellular automata (not genetic algorithms) the appropriate simulation approach. We appreciate an anonymous reviewer's suggestion to clarify this choice.

[viii] By stochastic processes, we include a broad class of simulations that are unified by the use of custom designed algorithms or combinations of algorithms. By contrast, although the structured simulation approaches also often involve stochasticity, they have a much more specified structure.

[ix] A structured approach is sometimes easier for readers to accept and for novice simulation researchers to execute, but can be difficult and unwieldy to modify if the changes are extensive. In contrast, stochastic processes are more adaptable to a particular research agenda, but require more modeling ingenuity and writing clarity for readers. The choice between the two depends upon the extent of modification required. That is, a need for greater modification favors stochastic processes while modest modification favors a structured approach.

[x] Both verification and experimentation involve running the simulation, and thus making several specific decisions about those runs. One is the number of time steps. The appropriate number depends upon researcher objectives (e.g., observing process unfold, assessing time to reach an optimal point, and examining steady state processes), and so researchers should justify their choice in light of their objectives. A second is the number of runs per set of specific values of independent constructs. This is roughly analogous to sample size in hypothesis testing research, specifically the sample size in a cell of an ANOVA in a laboratory study. Here the decision depends upon statistical power. More runs raise the statistical power, and so the choice depends upon the power necessary to perform adequate statistical analysis (i.e., creating confidence intervals for dependent variable values). The third is number of sets, roughly analogous to the number of different conditions in a laboratory study – i.e., number of cells in ANOVA of a laboratory study. However, unlike laboratory researchers, simulation researchers are able to make many runs, and so often only report those that are the most significant vis a vis conveying their results. Researchers should also justify these latter decisions in light of their research. Given space limitations and the relationship of these decisions to ideas that are familiar to many readers, we refer interested readers to texts on simulation (e.g., Law and Kelton, 1991) for more information.

[xi] Sensitivity analysis is a widely used term within simulation. It involves exploring the simulation with, for example, various construct values, operationalizations of constructs, and theoretical logics. For many uses of simulation, sensitivity analysis is a useful term. In contrast, however, for using simulation to develop theory in our context, the term includes uses of sensitivity analysis that have very different theory development objectives. Therefore, in this paper, we use more precise language to distinguish three kinds of sensitivity analysis: verification of propositions of simple theory, robustness checks of the computational representation (roughly analogous to the meaning of robustness using empirical methods), and experimentation to create new theory.

[xii] The issue sometimes arises as to whether simulation researchers should make their software code available. Although this is an issue for the field, our view is that a well-characterized computational representation obviates the need for the code because its functionality can be re-created from the computational representation (e.g., construct operationalizations, assumptions, and theoretical logic) that appears in the paper. We appreciate a reviewer's asking us to offer an opinion on this issue.

[xiii] We discussed these decisions in endnote x. In addition, given that simulation creates computer-generated data, other statistical analyses (e.g., regressions, correlations) are possible. But these are usually not central to theory development. This same observation is true of calibration. Thus, given these observations and space constraints, describing these activities is beyond the scope of this paper. We refer the interested reader to texts on simulation (e.g., Law and Kelton, 1991).